JGR Biogeosciences

RESEARCH ARTICLE
10.1029/2024JG008300

Sensitivity of Chlorophyll Vertical Structure to Model Parameters in the Biogeochemical Southern Ocean State Estimate (B-SOSE)

Angela M. Kuhn[1], **Matthew R. Mazloff**[1], **Sarah T. Gille**[1], **and Ariane Verdy**[1]

[1]Scripps Institution of Oceanography, University of California San Diego, San Diego, CA, USA

Key Points:

- Some model parameters may have a large effect on chlorophyll magnitude without affecting vertical structure, while others affect both
- Green's functions informed by a sensitivity analysis are cost-effective to obtain significant reductions in model-observation misfits
- Turbulence, differences in phytoplankton traits, and grazing parameterizations are key areas for model development

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

A. M. Kuhn,
akuhncordova@ucsd.edu

Citation:

Kuhn, A. M., Mazloff, M. R., Gille, S. T., & Verdy, A. (2025). Sensitivity of chlorophyll vertical structure to model parameters in the Biogeochemical Southern Ocean State Estimate (B-SOSE). *Journal of Geophysical Research: Biogeosciences*, *130*, e2024JG008300. https://doi.org/10.1029/2024JG008300

Received 27 JUN 2024
Accepted 7 JAN 2025

Author Contributions:

Conceptualization: Angela M. Kuhn, Matthew R. Mazloff, Sarah T. Gille, Ariane Verdy
Formal analysis: Angela M. Kuhn
Funding acquisition: Matthew R. Mazloff
Investigation: Angela M. Kuhn
Methodology: Angela M. Kuhn, Matthew R. Mazloff, Sarah T. Gille
Resources: Matthew R. Mazloff
Software: Ariane Verdy
Supervision: Matthew R. Mazloff
Visualization: Angela M. Kuhn
Writing – original draft: Angela M. Kuhn
Writing – review & editing: Matthew R. Mazloff, Sarah T. Gille, Ariane Verdy

Abstract The Southern Ocean is a region of intense air–sea exchange that plays a critical role for ocean circulation, global carbon cycling, and climate. Subsurface chlorophyll-a maxima, annually recurrent features throughout the Southern Ocean, may increase the energy flux to higher trophic levels and facilitate downward carbon export. It is important that model parameterizations appropriately represent the chlorophyll vertical structure in the Southern Ocean. Using BGC-Argo chlorophyll profiles and the Biogeochemical Southern Ocean State Estimate (B-SOSE), we investigate the sensitivity of chlorophyll vertical structure to model parameters. Based on the sensitivity analysis results, we estimate optimized parameters, which efficiently improve the model consistency with observations. We characterize chlorophyll vertical structure in terms of Empirical Orthogonal Functions and define metrics to compare model results and observations in a series of parameter perturbation experiments. We show that chlorophyll magnitudes are likely to respond quasi-symmetrically to perturbations in the analyzed parameters, while depth and thickness of the subsurface chlorophyll maximum show an asymmetric response. Perturbing the phytoplankton growth tends to generate more symmetric responses than perturbations in the grazing rate. We identify parameters that affect chlorophyll magnitude, subsurface chlorophyll or both and discuss insights into the processes that determine chlorophyll vertical structure in B-SOSE. We highlight turbulence, differences in phytoplankton traits, and grazing parameterizations as key areas for improvement in models of the Southern Ocean.

Plain Language Summary The Southern Ocean plays a central role in global climate and climate variability. While ocean chlorophyll concentrations are usually expected to be highest near the ocean surface, where the sun's energy is strongest, throughout the Southern Ocean in some parts of the year chlorophyll concentrations are observed to be highest tens of meters below the ocean surface. These features are referred to as "subsurface chlorophyll-a maxima (SCMs)." As chlorophyll concentrations are strongly related to phytoplankton abundance, SCMs are indicative of the amount of food available to other species and the amount of carbon that sinks to the deep ocean. It is thus important that computational ocean models capture the mechanisms that give rise to SCMs. Using Biogeochemical-Argo chlorophyll profiles and the Biogeochemical Southern Ocean State Estimate, we investigate the sensitivity of the simulated Southern Ocean chlorophyll vertical structure to biogeochemical model parameters involved in the estimation of chlorophyll. In a series of model experiments we modify key biogeochemical model parameters one at a time. New parameter values, derived from the perturbation experiments, improve the consistency of the model to observations.

1. Introduction

The Southern Ocean is a region of intense air–sea exchange, with zones where the surface waters are subducted downward, sequestering massive amounts of carbon and heat from the atmosphere (Russell et al., 2006; Wunsch & Heimbach, 2008). It is an important region for ocean circulation, primary production, carbon cycling, and global climate (Frölicher et al., 2015; Gruber et al., 2019; Landschützer et al., 2015). Subsurface chlorophyll-a maxima (SCMs) are an annually recurrent feature throughout the Southern Ocean (Baldry et al., 2020; Carranza et al., 2018). In general, SCMs arise from an interplay between the homogenizing effects of turbulence and factors creating gradients or accumulation of phytoplankton, such as light attenuation and buoyancy of cells (Cullen, 2015). In the Southern Ocean, SCMs are created by a range of processes including photo-acclimation (Rembauville et al., 2016), diatom aggregation (Armand et al., 2008; Gomi et al., 2010; Parslow et al., 2001), eddies (de Villiers et al., 2015; Westwood et al., 2011), and coastal water subduction events (Wright & van den Enden, 2000). On a local scale, SCMs may increase the energy flux to higher trophic levels and facilitate

downward carbon export (Baldry et al., 2020). Future changes in ocean stratification and sea-ice coverage in the Southern Ocean may modify the generation and distribution of SCMs. It is, thus, important to implement model parameterizations capable of appropriately representing chlorophyll vertical structures and their formation mechanisms.

Observational information about the vertical structure of chlorophyll in the Southern Ocean, required for model constraint and validation, is often sparse due to the challenges that the region poses for shipboard oceanography. Satellites provide the longest synoptic record of chlorophyll in the Southern Ocean, but they are limited in their ability to observe deeper layers of the ocean. This intrinsic uncertainty in the satellite-based estimates of integrated chlorophyll and primary production may result in the underestimation of phytoplankton biomass in the Southern Ocean and in biases in the temporal variability (Prend et al., 2019; Uchida et al., 2019). Our understanding of the vertical structure of the ocean in the Southern Ocean has greatly benefited from a network of Biogeochemical-Argo (BGC-Argo) floats deployed by the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project (Johnson et al., 2017; Talley et al., 2019). The assimilation and combined use of BGC-Argo observations and ocean models have also allowed advances in our knowledge of Southern Ocean biogeochemistry. For instance, the Biogeochemical Southern Ocean State Estimate (B-SOSE; Verdy & Mazloff, 2017) and BGC-Argo profiles have been used to produce mapped estimates of monthly pH, revealing Southern Ocean acidification (Mazloff et al., 2023). BGC-Argo profiles and bio-logging profiles have also enabled the exploration of uncertainties in the vertical structure of phytoplankton biomass and chlorophyll in the Southern Ocean (Carranza et al., 2018; Prend et al., 2019; Uchida et al., 2019; Von Berg et al., 2020). Chlorophyll fluorescence measurements from BGC-Argo provide data under clouds, during all seasons, and show good general agreement with satellite-derived surface chlorophyll products (Haëntjens et al., 2017).

One of the major challenges in modeling phytoplankton and chlorophyll is the determination of the biogeochemical model parameters (Franks, 2009). These parameters mathematically affect the biological and chemical rates of change and are poorly known or intrinsically uncertain. Laboratory and in situ data provide an estimate of the range of values these parameters can take; however, they may not be suitable to be applied directly in models (Franks, 2009). For example, observational measurements may be specific to certain plankton species and environmental conditions, while models often combine plankton species together into a single category or just a few groups. In this study, we aim to improve the chlorophyll vertical structure simulated by B-SOSE through the use of BGC-Argo chlorophyll profiles. For this purpose, we first evaluate B-SOSE sensitivity to biogeochemical model parameters directly involved in the estimation of phytoplankton biomass and chlorophyll. We then use results from this sensitivity analysis and a Green's function approach (e.g., Brix et al., 2015; Carroll et al., 2022; Menemenlis et al., 2005) to find a set of optimized parameters that improve B-SOSE performance with respect to assimilated data.

The remainder of this manuscript is organized as follows: in Section 2, we describe the observations (2.1 and 2.2), model (2.3), parameter sensitivity analysis (2.4), and parameter optimization method (2.5); in Section 3, we present the results from the characterization of the Southern Ocean in the observations and model control and experimental runs (3.1) and the results from the parameter optimization contrasted against observations and our control run (3.2); in Section 4, we discuss these results in the context of previous studies and summarize our conclusions.

2. Methods

We use BGC-Argo chlorophyll profiles (Argo, 2000) and B-SOSE chlorophyll results to (a) calculate empirical orthogonal functions (EOFs) characterizing the vertical structure of chlorophyll (Chl-a) in the Southern Ocean, (b) perform a sensitivity analysis of parameters involved in the estimation of phytoplankton biomass and chlorophyll in B-SOSE, and (c) estimate an optimized set of parameter values based on the sensitivity analysis. Using EOFs for the observations–model comparison and for parameter optimization allows us to focus on the vertical structure patterns and minimize the effect of uncertainties in the magnitude of BGC-Argo chlorophyll observations. This is particularly important to avoid biasing the model during the optimization. For the EOF analyses, we consider both the entire model domain south of 30ºS as well as three sub-regions divided based on property characterizations (Figure 1; Orsi et al., 1995). Sub-region 1, north of the Antarctic Circumpolar Current (ACC), comprises the Subtropical Zone (STZ) and Subantarctic Zone (SAZ). Sub-region 2 comprises the Polar Frontal Zone (PFZ) and the Antarctic Zone (AZ) and is found within the extent of the ACC. Finally, sub-region 3

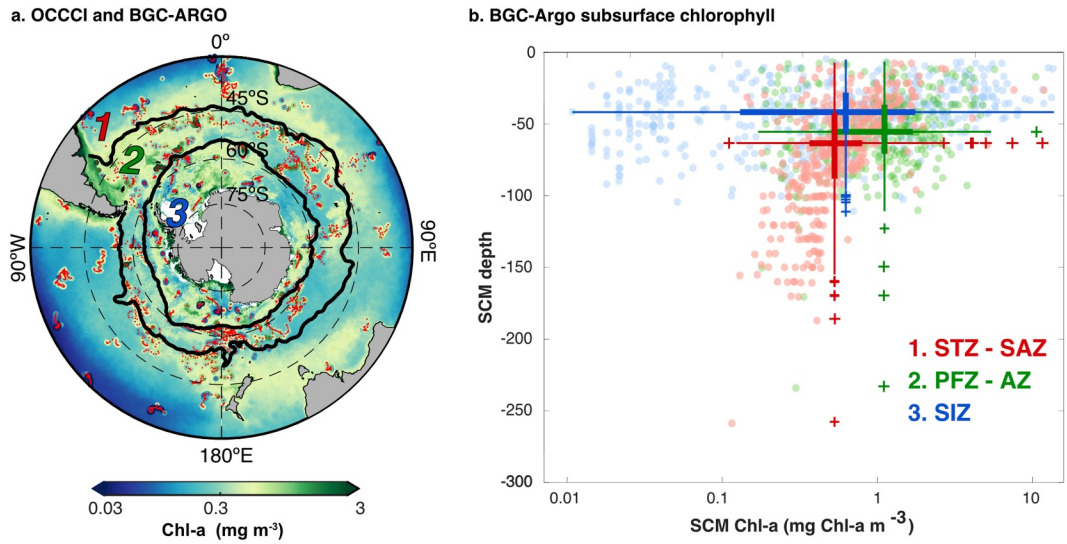

Figure 1. Observational data used in the study. (a) The background map shows average satellite surface chlorophyll from the European Space Agency Ocean Color–Climate Change Initiative (ESA OC-CCI; Sathyendranath et al., 2019), from September 2019 to August 2020. Overlaid and on the same color scale are surface chlorophyll values from BGC-Argo profiles recorded during the same period. Red dots highlight the locations of profiles. Thick black lines delineate the sub-regions used in this study: 1. Subtropical Zone and Subantarctic Zone, 2. Polar Front Zone and Antarctic Zone, and 3. Sea Ice Zone. (b) Subsurface chlorophyll maxima (SCM) magnitude found from the BGC-Argo profiles versus the corresponding SCM depths (SCM_d) in the 3 sub-regions considered in the study. For this plot, SCM magnitudes are defined as the maximum value in the profile and are required to exceed 5% of the surface value. Thick vertical and horizontal bars for each sub-region represent the interquartile range between the 0.25 and 0.75 quartiles and intersect at the median values of SCM magnitude and SCM_d. Thin vertical and horizontal lines represent the range of the data, excluding outliers (values >1.5 times the interquartile range).

comprises the Sea Ice Zone (SIZ). In addition, surface chlorophyll model results prior and after the optimization are compared against satellite chlorophyll observations to evaluate the model skill in the entire domain.

2.1. BGC-Argo Observations

We use quality-controlled (flags 1, 2, 5, and 8) chlorophyll information from BGC-Argo profiles south of 30ºS, measured between September 2019 and August 2020 (Figure 1a). This period contains one complete annual cycle with a higher number of profiles compared to previous years. Each of the 3 sub-regions considered is well represented in the observational data set and presents different characteristics in terms of SCM magnitudes and depths (Figure 1b). Sub-region 1 tends to have deeper SCM depths (>150 m) with magnitudes ranging between 0.1 and 5 mg Chl-a m^{-3}. Sub-region 2 presents a narrower range of SCM depths, with a median value of 55 m and with depths reaching down to 110 m. It also has slightly higher SCM magnitudes, up to 10 mg Chl-a m^{-3}. Sub-region 3 shows the widest range of SCM magnitudes and shallowest SCM depths, with a median value of 40 m depth.

2.2. Satellite Chlorophyll Data

We compare model results to monthly satellite chlorophyll data (Figure 1a) from the 4-km resolution version 5.0 of the European Space Agency Ocean Color–Climate Change Initiative (ESA OC-CCI) from September 2019 to August 2020 (Sathyendranath et al., 2019). While it is not the focus of the present study, daily OC-CCCI was used to evaluate the consistency of the model and BGC-Argo chlorophyll with respect to satellite chlorophyll (Figure S1 in Supporting Information S1). The OC-CCI chlorophyll is a globally and optimally merged product of Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer on the Aqua Earth Observing System (MODIS-Aqua), Visible Infrared Imaging Radiometer Suite (VIIRS), and Medium Resolution Imaging Spectrometer (MERIS) data (Jackson et al., 2017). Previous studies have shown that the OC-CCI product provides higher data coverage and lower bias than individual sensors and previous merged data sets (Belo Couto et al., 2016).

Table 1
Subset of B-SOSE's Biogeochemical Model Parameters Selected for the Sensitivity Analysis and Optimization

Parameter	Control	Optimized	Range	Units	Description
μ_0	1.123	2.1535	0.02–2.4[a]	d^{-1}	Phytoplankton growth rate at $T = 0°C$
k_{Fe}	0.4	0.548	0.02–2.64[b]	μ mol Fe m^{-3}	Half-saturation concentration of Fe
α^{chl}	2.782	3.6915	0.55–8.64[c]	gC $gChl^{-1}$ m^2 W^{-1} d^{-1}	Initial slope of the photosynthetic response
$\theta_{\max lo}$	0.01	0.0192	0.007–0.072[a,c]	gChl gC^{-1}	Maximum Chl:C ratio when iron-limited
$\theta_{\max hi}$	0.04	0.0874	0.007–0.072[a,c]	gChl gC^{-1}	Maximum Chl:C ration when iron-replete
λ_0	0.1892	0.2355	0.2–3.0[d,e,f,g]	d^{-1}	Grazing rate at $T = 0°C$

Note. Parameter values in B-SOSE control and optimized run, ranges from the modeling and observational literature, corresponding units, and description of the parameters are detailed. [a]Sathyendranath et al. (2009). [b]Timmermans et al. (2004). [c]Galbraith et al. (2010). [d]Sarthou et al. (2005). [e]Fahnenstiel et al. (1995). [f]Gifford et al. (1995). [g]Netjstgaard et al. (2001).

2.3. B-SOSE Configuration

The physical framework of B-SOSE is based on the MIT general circulation model and software developed by the consortium for Estimating the Circulation and Climate of the Ocean (ECCO; Wunsch & Heimbach, 2013). B-SOSE assimilates satellite and in situ physical and biogeochemical data, including BGC-Argo NO_3, O_2, and pH data (Verdy & Mazloff, 2017). The model domain has an open boundary at the equator, and the analysis focuses on the solution poleward of 30ºS. It has 52 vertical levels of varying thickness, a zonal grid resolution of 1/3°, and a meridional grid resolution that varies with latitude. Boundaries, initial conditions, and other configuration parameters are described by Verdy and Mazloff (2017).

The B-SOSE biogeochemical module is a nitrogen based adaption of the Biogeochemistry with Light, Iron, Nutrients, and Gases model (N-BLING; Galbraith et al., 2010). The version employed has an intermediate complexity with nine prognostic tracers (iron (Fe), nitrate (NO_3), phosphate (PO_4), dissolved inorganic carbon, alkalinity, dissolved oxygen, dissolved organic nitrogen, dissolved organic phosphorus, and biomass). Due to the simplified ecosystem dynamics, the model employs a relatively small number of prescribed parameters.

Our sensitivity analysis focuses on a subset of the biogeochemical model parameters that directly affect the phytoplankton biomass and chlorophyll parameterizations (Section 2.3; Table 1). In B-SOSE, phytoplankton biomass change is calculated separately for three functional groups: small, large, and diazotrophs. Here, we focus only on non-diazotrophic organisms. In the model, small and large phytoplankton share the same parameter values, but differ in their vulnerability/palatability to grazers. This is expressed as a power law with a size-dependent exponent (Galbraith et al., 2010; Verdy & Mazloff, 2017), which results in increased mortality for the small phytoplankton fraction. To provide context for the sensitivity analysis results, here, we simplify the nomenclature of the model by referring in general terms to phytoplankton biomass (B) and briefly describe its parameterization. B is estimated as a function of phytoplankton growth (μ) minus grazing ($\lambda_0 g$) by herbivores:

$$\frac{dB}{dt} = (\mu - \lambda_0 g)\, B. \tag{1}$$

Phytoplankton growth is equal to nutrient uptake, which is determined as a product of light (I) limitation, temperature (T) dependence, and nutrient (NO_3, PO_4, Fe) limitations:

$$\mu = \mu_{max}\left(1 - e^{-\frac{I}{I_k}}\right) \tag{2}$$

$$\mu_{max} = \mu_0 e^{k_T T}\, min\left(\frac{NO_3}{k_{NO_3} + NO_3}, \frac{PO_4}{k_{PO_4} + PO_4}, \frac{Fe}{k_{Fe} + Fe}\right). \tag{3}$$

The light-saturated phytoplankton maximum growth rate, μ_{max}, depends on the prescribed parameter μ_0, which represents the non-limited reference maximum growth rate at 0°C. Temperature dependence follows Eppley (1992) with $k_T = 0.063°C^{-1}$. Nutrient limitation is given by Monod kinetics and the minimum available nutrient.

The second term in Equation 2 determines the degree of light-limitation (Geider et al., 1997) using a scaling factor, I_k:

$$I_k = \frac{\mu_{max}}{\alpha^{chl} \times \theta_{\max\ Fe}} + \frac{I_{mem}}{2}, \tag{4}$$

where α^{chl} is the initial slope of the chlorophyll-a specific photosynthesis-light response curve, $\theta_{\max\ Fe}$ is a scale factor for the ratio of chlorophyll synthesis to carbon assimilation, and I_{mem} is the irradiance to which the phytoplankton are acclimated. $\theta_{\max\ Fe}$ represents the importance of Fe in forming chlorophyll accessory antennae. $\theta_{\max\ Fe}$ is calculated using the prescribed parameters $\theta_{\max hi}$ and $\theta_{\max lo}$, which represent the maximum chlorophyll-to-carbon (Chl:C) ratio when cells are iron-replete and iron-limited, respectively:

$$\theta_{\max\ Fe} = \theta_{\max lo} + \left(\theta_{\max hi} - \theta_{\max lo}\right) \times \frac{Fe}{k_{Fe} + Fe} \tag{5}$$

A variable Chl:C ratio, θ_{chl}, is then estimated to transform phytoplankton biomass, B (Eq.1), into chlorophyll units:

$$\theta_{chl} = \frac{\theta_{\max\ Fe}}{1 + \alpha^{chl}\theta_{\max\ Fe}\frac{I_{mem}}{2\mu_{max}}} \tag{6}$$

The grazing rate on phytoplankton (g) is also a temperature dependent function. The grazing rate is scaled depending on the type of phytoplankton prey, and it is modified by the prescribed parameter λ_0, which represents the reference grazing rate at $T = 0$°C. This function is discussed thoroughly by Galbraith et al. (2010).

2.4. Parameter Sensitivity Analysis

Our sensitivity analysis focuses on a subset of biogeochemical model parameters that directly affect phytoplankton and chlorophyll calculations, including μ_0, α^{chl}, k_{Fe}, $\theta_{\max hi}$, $\theta_{\max lo}$, and λ_0. The description of these parameters, units, a priori values, and optimized values (see Section 2.3) is summarized in Table 1. We change parameters one at a time, doubling and halving their magnitudes relative to the control values (Table 1), and we re-run the model for each perturbation. We compare observations and model results both in terms of chlorophyll profiles and profiles of empirical orthogonal function decomposition (EOFs). The EOF analyses are carried out for a depth-standardized composite of BGC-Argo profiles, for the control run, and for each parameter perturbation experiment, with sets of EOFs computed for the entire domain and for the three sub-regions in each case. For the comparison, BGC-Argo chlorophyll EOFs were interpolated to model depth levels.

We characterize the sensitivity of the model to each parameter using four metrics based on the first EOF mode (EOF1; Figure 4):

- $EOF1_0$: EOF1 surface magnitude at the top layer (2.1 m depth)
- $EOF1_{SM}$: EOF1 subsurface maximum magnitude, estimated as the maximum value in EOF1 with the criterion that it must be at least 10% higher than the $EOF1_0$ to be considered.
- SM_d: EOF1 subsurface maxima depth, defined as the depth at which $EOF1_{SM}$ occurs.
- SM_h: EOF1 subsurface layer thickness, estimated as the depth difference between the 25th and 95th percentiles of the EOF1 cumulative sum.

2.5. Parameter Optimization

In an attempt to improve the representation of B-SOSE's chlorophyll vertical structure with respect to observations, we optimize the biogeochemical model parameters included in the sensitivity analysis to fit BGC-Argo chlorophyll EOF1. We follow a Green's function approach, which involves estimating the best linear combination of model parameters to fit the target observations. Green's functions have been previously used with ocean general circulation models to find optimal initial and surface boundary conditions, diffusion coefficients, critical Richardson numbers, and relaxation time scales (Menemenlis et al., 2005). In ocean biogeochemical models, the method has been applied to successfully adjust surface pCO_2 and air–sea CO_2 flux, using perturbations to initial

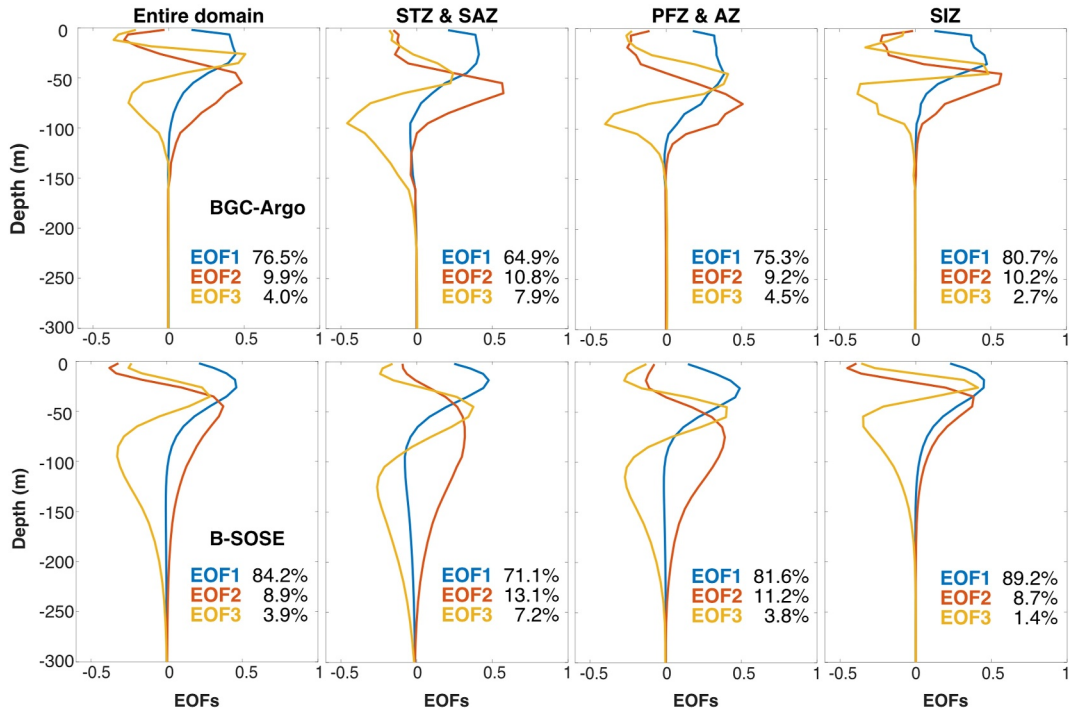

Figure 2. Comparison of empirical orthogonal functions dominant modes of chlorophyll in the BGC-Argo observations (top panels) and in the B-SOSE control run (bottom panels).

conditions, phytoplankton growth rates, plankton palatability, iron scavenging, solubility, and gas exchange coefficients (Brix et al., 2015; Carroll et al., 2022). In contrast with these previous applications, our use of EOFs for the optimization function allows us to reduce the problem dimensionality and focus on the vertical structure of chlorophyll. In our application, we also acknowledge the non-linear nature of the biogeochemical system and consider positive and negative perturbations to each parameter to better constrain the optimized values.

For this optimization, we define vector y as the residual difference between BGC-Argo chlorophyll EOF1 and B-SOSE's control run chlorophyll EOF1. We limit our optimization to the first 300 m of the water column, where the observational EOF1 shows most of its structure (Figure 2). After running all model parameter perturbation experiments and estimating EOFs for each, we construct a matrix H that contains the EOF1 differences from each perturbed run to the control run. That is, H relates each model parameter perturbation run (r) to the EOF1 perturbation response (f):

$$H = \begin{bmatrix} \frac{\partial f(z_1)}{\partial r_1} & \cdots & \frac{\partial f(z_{31})}{\partial r_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(z_1)}{\partial r_{12}} & \cdots & \frac{\partial f(z_{31})}{\partial r_{12}} \end{bmatrix}. \tag{7}$$

H has NP × ND dimensions, which are the number of perturbed runs (here NP = 12) and the number of depth levels (here ND = 31), respectively. The fit to the observational EOF1 differences, y, is estimated as $\hat{y} = Hx$, where x is the optimized perturbation vector solved for by:

$$x = \left(HR^{-1}H^T + P^{-1}\right)^{-1}\left(HR^{-1}\right)y. \tag{8}$$

Here, R is a covariance matrix of the prescribed uncertainty squared in the observed EOF1, and P is the covariance matrix representing the prescribed uncertainty squared in the parameter values. Figure 3 shows that the surface

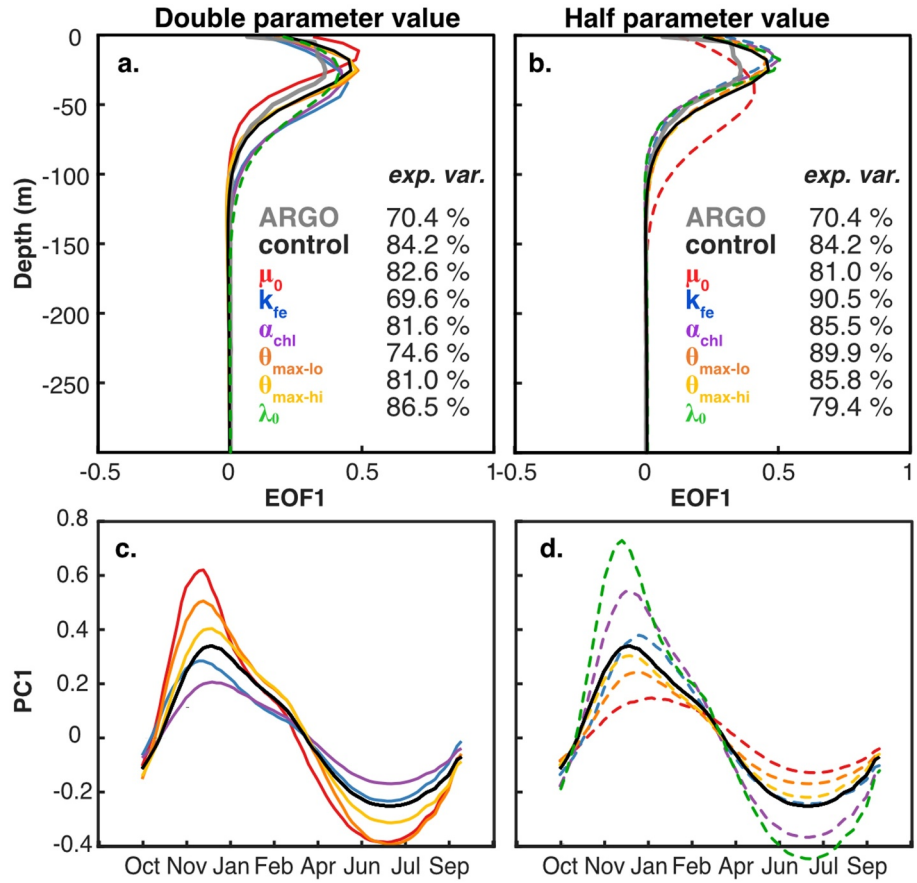

Figure 3. EOF results from the model parameter perturbation experiments compared to observational BGC-Argo EOF1 (a, b) and PC1 (c, d).

chlorophyll BGC-Argo EOF1 value is outside the spread of the perturbation solutions and thus cannot exactly be fit. Hence, based on preliminary results, we prioritized fitting depth layers between 5 and 65 m by assigning a low uncertainty value of 0.01 to the EOF1 differences within our target depth range and a value 1 elsewhere in the R covariance matrix. The prescribed uncertainty in the parameter values, P, was assumed to be 1 for all parameters. Parameter value minimum and maximum ranges from the literature are presented in Table 1 as a reference only. Essentially, we find a least-squared fit with the most suitable combination of parameter perturbations. That is, the vector x represents the fractional adjustment coefficients to the first-guess (control run) parameters p_0. For each parameter, an optimized value ($\hat{p}$) is estimated as:

$$\hat{p} = (1 + x_d - 0.5x_h)\, p_0, \tag{9}$$

where x_d, x_h are the x adjustments coefficients solved for in Equation 7 that correspond to the doubling and halving perturbation runs for each parameter. If a doubling is helpful and a halving unhelpful (e.g., $x_d > 0, x_h < 0$) the parameter will be increased ($\hat{p} > p$). In contrast, if a doubling had no value and a halving was optimal, the optimized parameters may be equivalent to $x_d = 0, x_h = 1, \hat{p} = 0.5p$. A pedagogical example of the derivation of Equation 8 is included in Appendix I, and MATLAB code to replicate the method is available in a public repository (Kuhn, 2024).

The set of optimized parameters (Table 1) that we obtain is then used in a B-SOSE simulation, model EOFs are re-estimated (Figure 5b), and root-mean-square-errors from the EOFs ($RMSE_{EOF}$) and from individual BGC-Argo chlorophyll profiles ($RMSE_{chl}$) are calculated to evaluate the performance of B-SOSE using the optimized parameters.

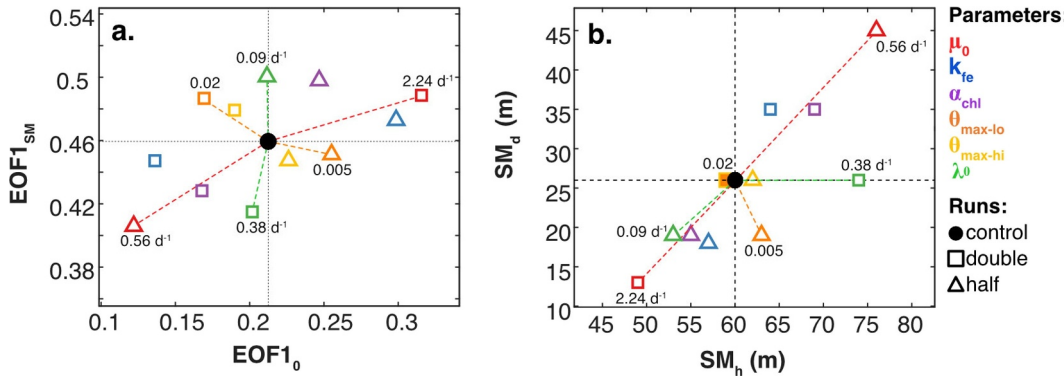

Figure 4. Comparison of metrics in model control run and parameter perturbation experiments estimated from EOF1 for the entire domain: (a) changes in EOF1 surface magnitude and subsurface maximum magnitudes; and (b) EOF1 subsurface maximum depth and subsurface maximum thickness. The filled black circle represents the metrics' values in the control run, and dotted black lines are included for reference. The colored markers represent the metrics' values in each of the perturbation experiments. Squares represent the experiments that double the parameter values, and triangles represent the experiments that halve the parameter values. As reference examples, colored dashed lines connect "doubling" and "halving" experiments for parameters μ_0 (red), λ_0 (green), and $\theta_{\max lo}$ (orange). The corresponding experimental parameter values are shown next to their marker. A description of parameters, units, and values appears in Table 1.

3. Results

3.1. Sensitivity of the Chlorophyll Vertical Structure

The first chlorophyll EOF (EOF1) captures more than 60% of the variability in both BGC-Argo observations and model results for the entire domain, as well as in individual sub-regions (Figure 2, blue lines). The second (EOF2) and third (EOF3) EOFs respectively capture 10.25% ± 1.45% and 4.43% ± 2.16% on average for the entire domain and individual sub-regions (Figure 2, red and yellow lines). In general, EOF1 exhibits pronounced subsurface maxima between 10 and 70 m. EOF1 estimated from BGC-Argo observations (upper panels of Figure 2) shows some variability between sub-regions, with the central sub-region 2 presenting a less pronounced and deeper subsurface maximum. Compared with results from observations, EOF1 profiles estimated from B-SOSE results (lower panels of Figure 2) have higher subsurface maxima magnitudes and less variability among sub-regions. EOF2 profiles estimated from BGC-Argo observations have a more defined peak between 50 and 70 m, while in the model results, they present lower maximum magnitudes with a smoother and deeper distribution.

Perturbing biogeochemical parameters produces changes in both the mean vertical chlorophyll structure and EOF1 vertical structure (Figures 3a–3d). The most significant changes in EOF1 vertical structure are obtained by doubling or halving μ_0, and by doubling k_{Fe}, α^{chl}, or λ_0 (Figures 3a and 3b). Parameters $\theta_{\max lo}$ and $\theta_{\max hi}$ have only small effects on EOF1. None of the doubling and halving perturbations modify EOF1 in a way that best matches the entire observational EOF1 profile. While addressing model sensitivity in terms of temporal variability is not the focus of the present study, we also note that perturbations produce changes in the temporal evolution of the modeled first principal component (PC1). In general, PC1 resembles the evolution of the annual cycle of chlorophyll with a predominant summer phytoplankton bloom (Figures 3c and 3d). Some parameter perturbations that have little impact on the vertical structure of EOF1 appear to have large effects on PC1. This is particularly noticeable for $\theta_{\max lo}$. Temporal dimension PC patterns cannot be properly evaluated from BGC-Argo observations because of the irregular temporal sampling obtained when aggregating multiple profiles. As a consequence, BGC-Argo PC patterns are not a linear function of time but of float ID and profile number, and their temporal interpretation is meaningless.

Changes in the chlorophyll EOF1 profile produced by the parameter perturbations are summarized using the four metrics we describe in Section 2.3. In Figure 4, we focus on the deviations of characteristic properties of EOF1 relative to the control run as a result of perturbing model parameters. A priori, in a system that responds linearly, halving a parameter should produce a change that is of opposite sign and proportional to the effect of doubling the parameter, and the response should be connected via a straight line. In contrast, in a non-linear system, halving or

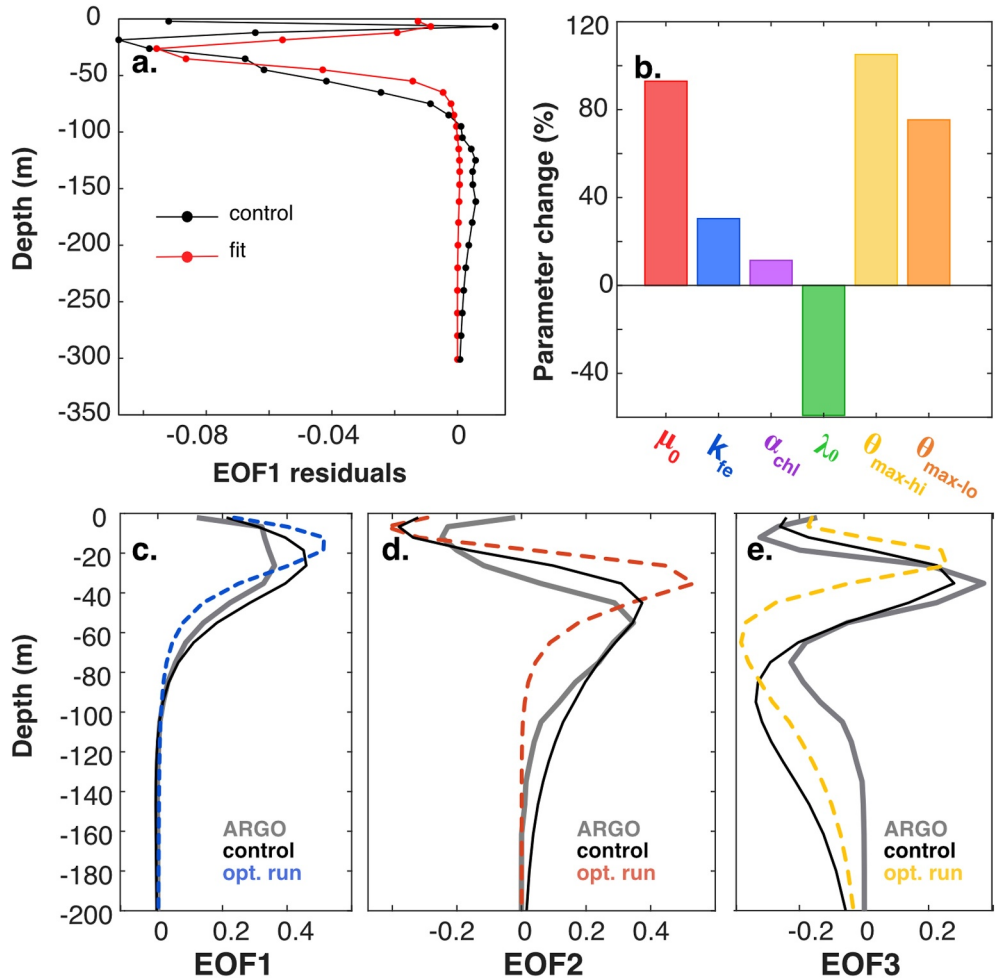

Figure 5. Results from parameter optimization: (a) observational EOF1 differences to the control run (y, BGC-Argo EOF1 minus B-SOSE control run EOF1) and estimated fit $\hat{y}$; (b) Percentage of parameter increase after the optimization $\left(\frac{optimized - control}{control} \times 100\right)$; and (c–e) Argo EOFs and B-SOSE EOFs before (control) and after (opt.) the optimization.

doubling parameters has the potential to yield distinctly different responses. Despite the non-linearity of both the biogeochemical model and the perturbation experiment design (i.e., halving and doubling), the responses of $EOF1_0$ and $EOF1_{SM}$ are quasi-symmetric, as illustrated for a few parameters in Figure 4a using colored dashed lines. Doubling the growth rate parameter (μ_0, red symbols and lines) increases both the EOF1 surface magnitude and subsurface maximum magnitude but decreases the subsurface maximum depth and subsurface layer thickness. The opposite pattern occurs when halving μ_0 or doubling k_{Fe} (blue symbols) and α^{chl} (purple symbols). Modifying the grazing rate (λ_0, green symbols and lines) generates changes in $EOF1_{SM}$, but not $EOF1_0$. Metrics SM_d and SM_h (subsurface maximum depth and thickness) present an asymmetric response to the doubling and halving of parameter values (Figure 4b). Doubling μ_0 or halving k_{Fe}, α^{chl}, and λ_0 deepen SM_d and SM_h, with deviations between 10 and 20 m from the control run values. In contrast, halving μ_0 or doubling k_{Fe}, α^{chl}, and λ_0 generates a shallower SM_d and thinner SM_h with deviations of about 5–10 m.

A number of parameters consistently show asymmetric responses to the examined perturbations. Particularly, perturbations of the grazing mortality λ_0 affect the SM magnitudes $EOF1_{SM}$, but not the surface values (green symbols in Figure 4a). Doubling λ_0 results in a thicker SM_h, but does not affect SM_d, while halving this parameter decreases both metrics (green square vs. green triangle in Figure 4b). As previously observed in Figure 3, the parameter perturbations with the least impact on model EOF1 are due to the Chl:C ratio parameters $\theta_{\max lo}$ (orange symbols) and $\theta_{\max hi}$ (yellow symbols). These parameters also present asymmetric responses to perturbations: doubling $\theta_{\max lo}$ and $\theta_{\max hi}$ increases $EOF1_{SM}$ but decreases the surface values

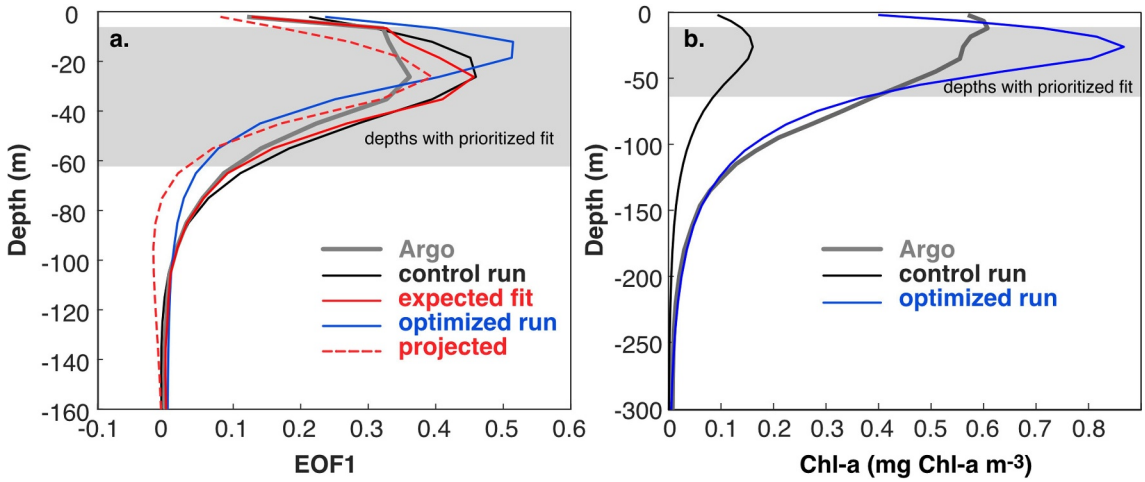

Figure 6. (a) Comparison of chlorophyll EOF1 solutions from BGC-Argo observational profiles (gray), B-SOSE model control run (V_{cont}, black), estimated fit for optimized parameters (EOF1 derived from $\hat{y}$, solid red), B-SOSE model optimized run (V_{opt}, blue), and B-SOSE model optimized run results projected on the control run temporal variability (V_{proj}, dashed red). (b) Comparison of chlorophyll concentrations from BGC-Argo observational profiles (gray), B-SOSE model control run (V_{cont}, black), and B-SOSE model optimized run (V_{opt}, blue). The gray area shows the depth levels where we assigned lower uncertainty during the optimization.

(Figure 4a). Halving the Chl:C ratio parameters slightly decreases the $EOF1_{SM}$ but increases the surface values (Figure 4a). The Chl:C ratio parameters has a negligible effect on the depth and thickness of the SM layer (Figure 4b). The model behaves similarly in all sub-regions in response to parameter perturbations (Figure S1 in Supporting Information S1).

3.2. Optimized Solution

Using the constraints provided by the parameter perturbation experiments, we obtain a solution ($\hat{y}$) that fits observational EOF1 differences to the control run between 6 and 65 m and is loosely constrained up to 300 m (Figure 5a). This solution can be interpreted as a fit to the missing signal in our control model simulation. The model parameters we obtain by solving $\hat{p}$ (i.e., optimized parameters, Equation 9) allow us to integrate this signal into a new optimized model solution. Most of the resulting optimized parameters (5 of 6) show increased values changed between +12% and +100% from their original value (Figure 5b.). The largest parameter value increases are obtained for $\theta_{\max hi}$, $\theta_{\max lo}$, and μ_0. The grazing mortality λ_0 was the only parameter that decreased as a result of optimization. The optimized parameter $\theta_{\max hi}$ exceeds the maximum reference value (Table 1). This was possible because literature reference parameter values were not used during the optimization because of their intrinsic uncertainty.

3.2.1. Effect of the Optimized Solution on Chlorophyll EOFs

After re-running B-SOSE with the optimized parameters and calculating EOFs for the optimized run, we find that optimized parameters appear to degrade the vertical structure of B-SOSE's chlorophyll EOF1, which was the target of the optimization (Figure 5c). Changes to EOF2 and EOF3 profiles are also noticeable (Figures 5d and 5e). This is likely due to the non-linearities of the system affecting both the model chlorophyll vertical structure and the spatio-temporal variability. To properly evaluate the impact of the optimized parameters on the model chlorophyll vertical structure (as defined by EOFs), we need to isolate changes in the vertical structure from changes in the temporal and spatial variability. For this test (Figure 6a), we re-calculated the optimized model EOFs by projecting them onto the control model temporal variability (PCs). That is, consider that matrix A contains the data, V are the eigenvectors (EOFs) for the covariance matrix corresponding to A, Λ are the eigenvalues, and τ are the principal component time series. Then, the EOFs for the control (V_{cont}, black line), and the optimized (V_{opt}, blue line) model runs can be estimated as:

$$V_{cont} = \left(\frac{\tau_{cont}}{\sqrt{\Lambda}_{cont}} A_{cont} \right)^T \quad (10)$$

$$V_{opt} = \left(\frac{\tau_{opt}}{\sqrt{\Lambda}_{opt}} A_{opt} \right)^T \quad (11)$$

The projection of the optimized-run chlorophyll vertical structure onto the control-run temporal variability is (V_{proj}), given by:

$$V_{proj} = \left(\frac{\tau_{cont}}{\sqrt{\Lambda}_{opt}} A_{opt} \right)^T . \quad (12)$$

Figure 6a shows that the projected EOF solution (V_{proj}, dashed red line) effectively reduces the EOF1 misfits with respect to BGC-Argo EOF1 (gray line) at the target depth levels (6–65 m). The projected EOF solution also outperforms the expected EOF1 fit (derived from $\hat{y}$, solid red line) estimated during the optimization. This suggests that, when we isolate the effects of the parameter changes on the B-SOSE chlorophyll vertical structure, we find that the optimized parameters successfully improve chlorophyll vertical structure patterns. The improvement in chlorophyll vertical structure is evident in Figure 6b, where we compare the mean chlorophyll profiles over the entire domain from BGC-Argo (gray line), the control run (black line) and the optimized run (blue line). In the next Section (3.2.2), we further discuss changes in the chlorophyll distribution patterns obtained when running B-SOSE with the optimized parameters.

3.2.2. Effect of the Optimized Solution on Chlorophyll Magnitudes

The performance of the B-SOSE using the optimized parameters significantly improved with respect to BGC-Argo chlorophyll profiles, achieving a ∼60% reduction in total misfits ($RMSE_{control}$ = 0.2622; $RMSE_{optimized}$ = 0.1055). This improvement is largely caused by an overall increase in chlorophyll magnitudes (Figure 6). The increase in surface chlorophyll magnitudes is generalized throughout the domain and prevails the entire year. September-to-December surface chlorophyll magnitudes in the optimized run better fit the highest surface values reported by BGC-Argo (Figures 7a–7f), while overestimating mid-range chlorophyll values (0.3–1.5 mg chl-a m^{-3}) compared to BGC-Argo and satellite chlorophyll. From March to May, the optimized run matches BGC-Argo surface chlorophyll mid-range values better than the control, but low surface chlorophyll values in Sub-regions 2 and 3 are not replicated by the optimized run (Figures 7g–7i). From June to August, the optimized run matches BGC-Argo chlorophyll surface values well in sub-region 3 and overestimates other sub-regions (Figures 7j–7l). In general, the optimized surface chlorophyll solution overestimates chlorophyll with respect to the satellite OC-CCCI chlorophyll product throughout the year. Compared with monthly satellite observations (Figure S1 in Supporting Information S1), the surface chlorophyll in the optimized solution has higher misfits than the control run ($RMSE_{control}$ = 0.49; $RMSE_{optimized}$ = 0.57). We note a small increase in the model–satellite correlation (log10(chl-a): $r_{control}$ = 0.58, $r_{optimized}$ = 0.64) and an improvement in the slope of the relationship (Figures S1a and S1b in Supporting Information S1).

The widespread increase in chlorophyll magnitudes helps improve chlorophyll misfits throughout the water column (Figures 6b and 8), and the optimized solution approaches the floats' observational values through changes in the SCM magnitude and depth (Figures 8b–8c and 8e–8f). In many cases, artificial SCMs (i.e., present in B-SOSE but not in the observations) in the control run are not corrected by the optimization, as the optimized run is unable to replicate observational profiles with well-mixed chlorophyll (Figures 8b and 8c). In some cases, the slope of the chlorophyll profile between 100 and 150 m is significantly improved in the optimized solutions (Figures 8b–8c and 8e–8f). B-SOSE results appear to underestimate chlorophyll below 200 m with respect to BGC-Argo measurements (Figures 8b–8c, 8e–8f). However, the chlorophyll observational noise floor is on the order of 10^{-3}, while B-SOSE can reach concentrations numerically close to 0. As the focus of our optimization was between 6 and 65 m (i.e., lower uncertainty assigned to chlorophyll EOF1 values at these depths), these discrepancies do not affect our optimization results.

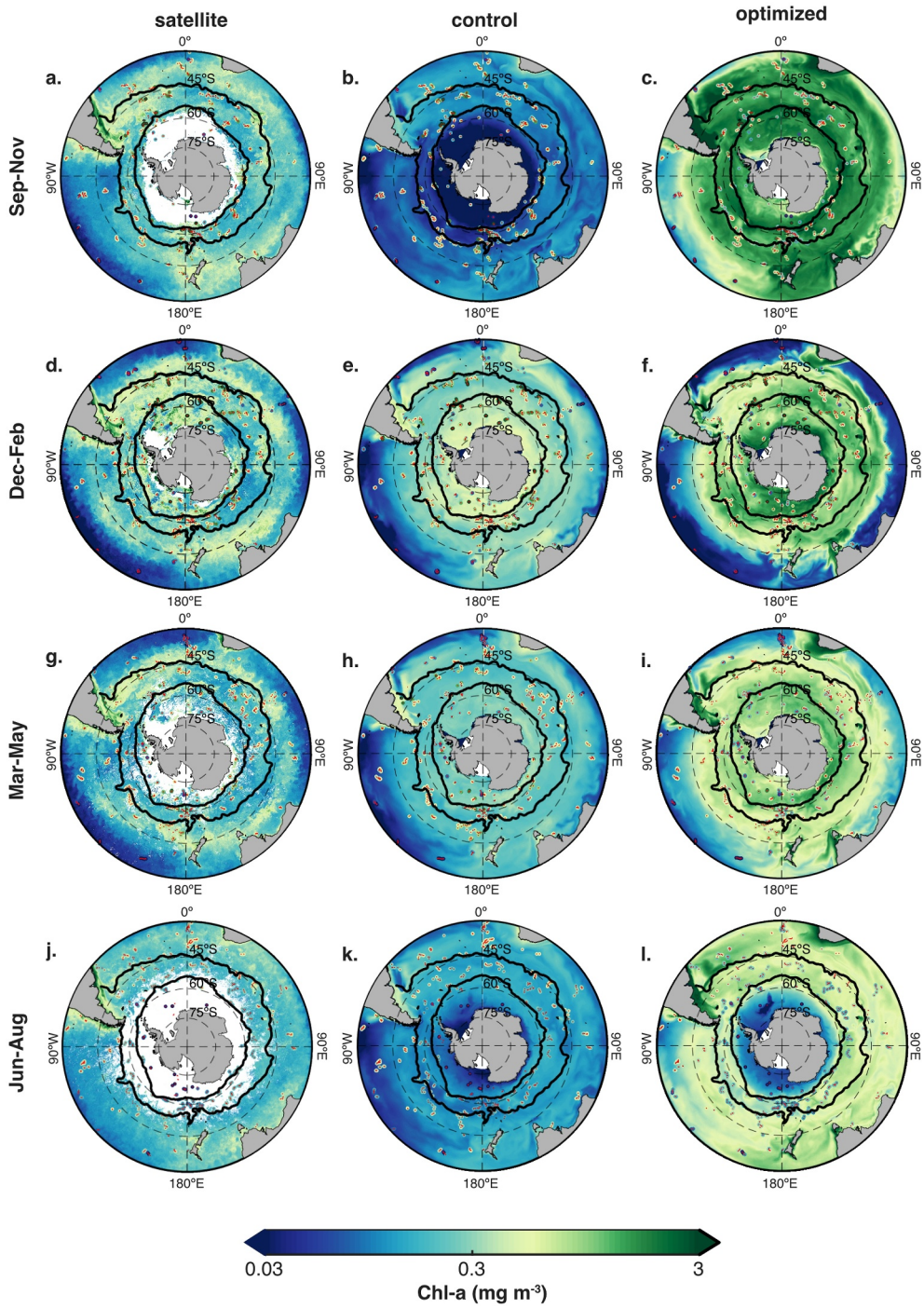

Figure 7. Comparison of average surface chlorophyll from the European Space Agency Ocean Color–Climate Change Initiative (ESA OC-CCI; Sathyendranath et al., 2019) satellite product, B-SOSE's control, and optimized runs and BGC-Argo surface average. In the left column (a, d, g, j), the background map shows average ESA OC-OCCI satellite surface chlorophyll. The middle column (b, e, h, k) shows B-SOSE's top layer average surface chlorophyll from the control run. The right column (c, f, i, l) shows B-SOSE's top layer average surface chlorophyll from the optimized run. Overlaid on all maps and on the same color scale are surface chlorophyll values from BGC-Argo profiles recorded during the same period. Red dots highlight the locations of profiles. Thick black lines delineate the sub-regions used in this study: 1. Subtropical Zone and Subantarctic Zone, 2. Polar Front Zone and Antarctic Zone, and 3. Sea Ice Zone.

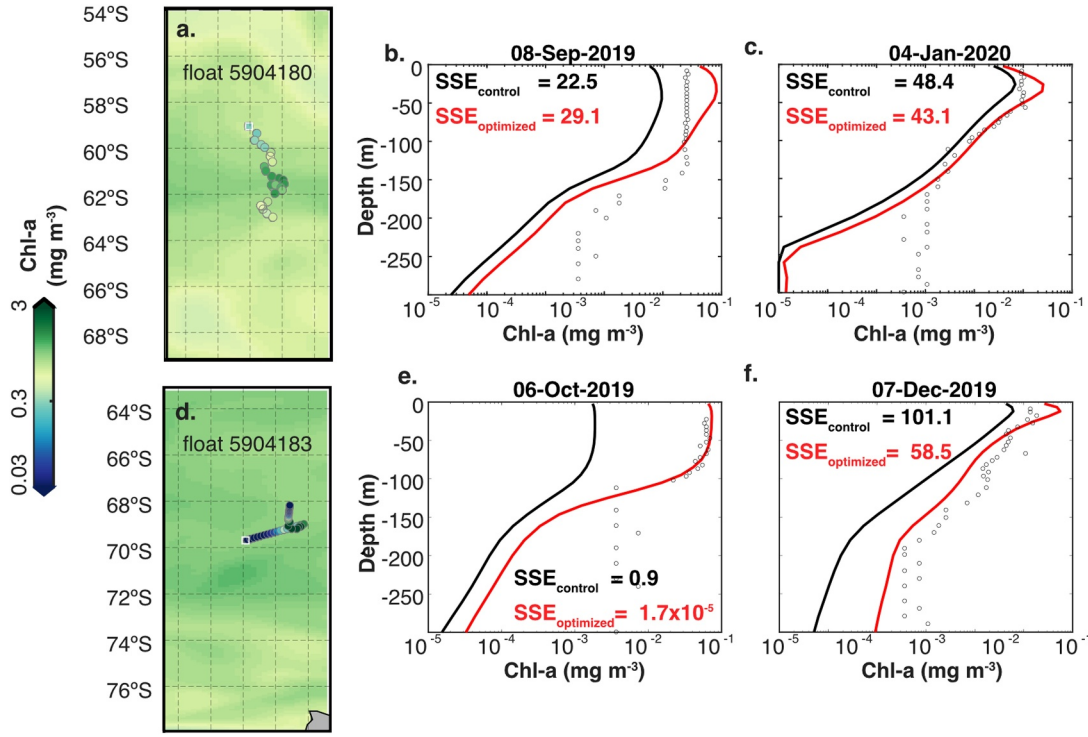

Figure 8. Examples of chlorophyll model performance with respect to observational BGC-Argo chlorophyll profiles. Maps (a) and (d) show the trajectory of each float with their corresponding observational top-30 m average chlorophyll values. The white squares in a. and d. show the beginning of the trajectories. The background color map shows the average model surface chlorophyll during the duration of each trajectory. BGC-Argo chlorophyll profiles (circles) are shown for two floats on different dates: float 5904180 (b, c) and float 5904183 (e, f). The thick lines are B-SOSE chlorophyll results before (black) and after (red) the parameter optimization. The sum of squared errors (SSE) before and after optimization is shown for each case as a reference.

4. Discussion and Conclusions

Based on a relatively small number of model perturbations, we have demonstrated a simple methodology to improve the representation of the chlorophyll vertical structure in B-SOSE. The optimized solution makes use of empirical orthogonal functions of BCG-Argo chlorophyll profiles; this approach allows us to focus on the vertical structure and avoids the dependence on chlorophyll magnitudes derived from chlorophyll fluorescence that may contain systematic uncertainties. Several other methods exist for parameter optimization (Schartau et al., 2017); however, they are computationally expensive when applied to three-dimensional models. The method we apply here is a cost-effective way to achieve a significant reduction in the simulated chlorophyll misfits (~60% of RMSE) with respect to BGC-Argo chlorophyll profiles. The Green's functions approach that we have used has also been found successful in other ocean biogeochemical model implementations, such as the ECCO2-Darwin model (Brix et al., 2015; Carroll et al., 2022). While the goal of reducing chlorophyll vertical structure misfits was achieved, we note a degrading effect on the misfits of simulated surface chlorophyll with respect to the OC-CCCI satellite chlorophyll product (see Section 4.3). In addition, the parameter sensitivity perturbation experiments demonstrate how the vertical structure of simulated chlorophyll varies with changes in the biogeochemical model parameters. These results provide insights for further parameter tuning and development of the ecosystem model, on which we elaborate below.

4.1. Parameter Sensitivity and Optimized Values

The chlorophyll vertical structure in B-SOSE, as represented with vertically resolved EOFs, was found to be very sensitive to the phytoplankton growth rate parameter (μ_0). Increasing this parameter results in higher chlorophyll EOF1 surface values and a shallower, thinner, and more pronounced EOF1 subsurface maximum. The opposite occurs when decreasing this value. The optimized solution generates an increase of almost double (+93%) the control parameter value μ_0 (Figure 5b.). This increase in phytoplankton growth rate is accompanied by increases in k_{Fe}, α^{chl} (+30% and +11% respectively), which partly counteract the increase in growth rate when nutrient and

light levels are low. An increase in k_{Fe} specifically helps regulate phytoplankton growth at low iron levels. The increase in α^{chl} acts both ways (Equations 4 and 5), increasing phytoplankton biomass when light-limited, but decreasing the Chl:C ratio, thus decreasing chlorophyll overall. The grazing parameter λ_0 (−60% change) presents an asymmetric response when doubled, changing the chlorophyll EOF1 subsurface maximum magnitudes and the subsurface layer thickness, but less so the surface magnitude and subsurface maximum depth (Figure 4). The decrease in λ_0 in the optimized solution would, thus, act to enhance the subsurface maximum by increasing the maximum phytoplankton biomass at depth without significantly modifying surface chlorophyll values or the depth of the maximum chlorophyll.

4.2. Insights for Model Development

Despite the improvement in chlorophyll magnitudes and vertical patterns obtained with the optimized parameters, B-SOSE still has limited skill in replicating the seasonal minimum chlorophyll values and well-mixed chlorophyll profiles during the Southern Ocean winter (March–July). Three aspects of the current B-SOSE configuration may explain this mismatch with observations: (a) turbulence, (b) differences in phytoplankton traits, and (c) grazing closure.

First, parameterized turbulence and other physical constraints in the B-SOSE may play a predominant role in establishing chlorophyll vertical structure during winter. BGC-Argo floats and elephant seal profiles from the Southern Ocean have shown that SCMs are less frequent during winter than during summer (Carranza et al., 2018). However, in our results, the optimized run overestimates both chlorophyll magnitudes and the prevalence of SCMs during the winter season. This suggests that parameterized turbulence or the mixing time-scale in the model is not enough to overcome the creation of biogeochemical vertical gradients. Sub-mesoscale processes, which are important for local restratification and biogeochemical distributions (du Plessis et al., 2019; Rosso et al., 2016), may also be underrepresented in the state estimate. Moreover, B-SOSE implicitly assumes that phytoplankton are always buoyant, even though changes in density occur throughout the year and throughout the water column. The interactions between mixing and variable sinking rates of phytoplankton play an important role in determining chlorophyll vertical structure (Cullen, 2015). The sinking rates of diatoms can change abruptly at the bottom of the mixed layer, influencing the number of times that cells return to the euphotic zone (Lande & Wood, 1987). For instance, SCMs in the Southern Ocean are often found at the base of the mixed layer, and density or turbulence gradients have been suggested as an important factor in their development (Carranza et al., 2018).

Second, variations in phytoplankton traits and grazing are among the ecological processes that potentially affect the prevalence of SCMs in the model solution. The B-SOSE used here estimates non-diazotrophic phytoplankton biomass in two groups—small and large—that share the same growth rate and nutrient uptake half-saturation parameters. In reality, different phytoplankton functional groups, and even species, are often characterized by different nutrient uptake characteristics. Both maximum growth rates and half-saturation constants scale with cell size, and species with high maximum growth rates tend to also have high half-saturation constants (Irwin & Finkel, 2017). They are, thus, poor competitors for a given nutrient under low nutrient concentrations (Edwards et al., 2013; Litchman et al., 2010). These different traits facilitate the establishment of a succession of phyto-plankton types during the year (Edwards et al., 2013).

Third, an additional ecological factor to consider is the parameterization of grazing in the model. Using sensitivity experiments in a different global biogeochemical model, Le Queré et al. (2016) concluded that zooplankton grazing was key in setting spatial and seasonal patterns of chlorophyll in the Southern Ocean. In our parameter perturbation experiments, by increasing the grazing parameter λ_0, we can modify high values of chlorophyll EOF1 (e.g., maximum subsurface magnitude) without strongly affecting surface values (Figures 3 and 4). In B-SOSE's current grazing parameterization, zooplankton pressure is implicit, and phytoplankton mortality does not depend on zooplankton densities. Small phytoplankton mortality by grazers increases linearly with phytoplankton biomass, while large phytoplankton mortality by grazers increases weakly and saturates at high prey densities (Dunne et al., 2005; Galbraith et al., 2010). Thus, an explicit representation of mortality dependence on predator density would increase phytoplankton mortality at high prey values and might improve the misfits obtained from March to August. An alternative to phytoplankton functional types and grazing parameterization would be the implementation of time and space varying parameters. This treatment has been shown to be useful in increasing

the similarity of simple models to observations in other systems (El Jarbi et al., 2013; Mattern et al., 2012), as well as in the Southern Ocean (Melbourne-Thomas et al., 2015).

In general, despite the simplicity of the present B-SOSE ecosystem dynamics, the model is able to replicate horizontal and vertical distributions of chlorophyll in the Southern Ocean, as well as some characteristics of seasonal variations. For instance, the absence of planktonic ecosystem detail in the biogeochemical model N-BLING (nitrogen based Biogeochemistry with Light, Iron, Nutrients, and Gases model; Verdy & Mazloff, 2017) has previously been shown to have minimal impact on the transport of nutrients, but can produce significant differences in biomass and export production when compared to more complex models (Schartau et al., 2017). Simple ecosystem models also offer the analytical advantage of fewer uncertain parameters than more complex models, making analyses such as the one we present, easier to interpret. Our detailed sensitivity analysis helps identify the contribution of biogeochemical processes and physical factors setting the vertical structure of chlorophyll in the Southern Ocean throughout the year. We highlight turbulence, species succession, and susceptibility to grazing as key factors in the development of subsurface chlorophyll maxima in the model, which is supported by previous observational studies (Baldry et al., 2020; Gomi et al., 2010). Our findings indicate that these parameterizations and the interconnection between phytoplankton growth and turbulence require further attention in future B-SOSE model developments and, in general, in other Southern Ocean model applications and observational analyses.

4.3. Caveats and Limitations

The cost-effective ocean biogeochemical parameter optimization we have presented uses 13 three-dimensional model simulations to achieve a significant reduction in the model misfits with respect to observations. Three main considerations make this possible: (a) simplified model ecology, (b) reduced problem dimensionality by using EOF's, and (c) linearized optimization via the Green's function method. We elaborate on these aspects below.

First, previous studies have shown that observations tend to be insufficient to constrain the total number of parameters even in simple biogeochemical models (Fennel et al., 2001; Schartau et al., 2017; Ward et al., 2010). The optimization of underconstrained parameters may generate unrealistic parameter values or increase model errors against unassimilated data (Friedrichs et al., 2007). Here, the simplified ecology of the biogeochemical model in B-SOSE (N-BLING) and our focus on chlorophyll vertical structure alone, allowed us to consider only a reduced number of parameters based on the model equations and a priori knowledge. For example, we excluded from the analysis nutrient half-saturation parameters for NO_3 and PO_4, as these are not usually limiting phytoplankton growth in our study region. Likewise, in the analysis we did not consider parameters that are not directly involved in the chlorophyll calculation (e.g., remineralization rates), but we acknowledge they may have an indirect effect on chlorophyll due to the feedbacks and non-linearities of the system. Because the computational cost of the Green's functions approach increases linearly, the method may be best suited for these type of optimizations, focused only on a single or few aspects of the system.

Second, we focused the optimization on EOF1, which captures more than 60% of the explained variance of the chlorophyll vertical structure in both the model and observations. However, the temporal pattern (PC1) cannot be trivially estimated from the observations due to the irregular temporal sampling. Changes in the model parameters affect both the chlorophyll vertical structure and the temporal patterns, which affect the optimized run EOF1 solution, making it appear degraded with respect to the control run EOF1. However, we demonstrated the success of the optimization with respect to EOF1 by isolating the vertical structure of the optimized run (see Section 3.2.1). This highlights the non-linearities of the system and the need to assess temporal variability. The EOF1-based optimization is able to improve model performance relative BGC-Argo chlorophyll profiles but degrades model performance relative to satellite surface chlorophyll product (Figure 7). This degradation of surface values is likely to arise from model limitations (see Section 4.2) in simultaneously representing profiles with low well-mixed chlorophyll and high sub-surface chlorophyll concentrations. Both BGC-Argo chlorophyll and OC-CCCI satellite chlorophyll contain uncertainties in the magnitude of chlorophyll concentrations. Moreover, the relationship between these two data sets is sensitive to the methodology used to compare them (Figures S1c and S1d in Supporting Information S1). A joint cost function to simultaneously optimize B-SOSE relative to both BGC-Argo and satellite chlorophyll magnitudes requires careful examination to consider light attenuation changes that occur through the year and that affect the satellite optical

depth (e.g., Haëntjens et al. (2017)) and/or to assign appropriate weights to each type of data. To optimize B-SOSE's chlorophyll temporal–spatial variability, comparison relative to direct observations may still be needed. This joint temporal–spatial assessment using satellite and floats chlorophyll data is outside the scope of the present study.

Third, the Green's function method assumes that an optimal model parameter set can be obtained as a linear combination of the model control run and the parameter perturbation experiments. However, ocean biogeochemical models often exhibit non-linear responses to changes in parameters and correlations between parameters. Our use of half and double perturbations for each parameter attempts to address this limitation, achieving a 60% reduction in the model—observations RMSE. We acknowledge that we neither explore the entire parameter space nor evaluate second-order parameter interactions as could be accomplished by alternative parameter optimization methods. Other optimal parameter combinations are possible and may result in further reductions to the model—observations misfits. These could be achieved by further iterating the method presented here or by applying alternative methods.

Appendix A: Estimation of Optimal Parameters

Consider the functional form of the observational EOF1 differences to the control run is $y = f(p_1, p_2, \ldots p_n)$, where p are the perturbed model parameters. To optimize p_1 we estimate:

$$y_{true} = y_0 + \frac{\partial y}{\partial p_1}\Delta p_1,$$

where y_0 represents the control run model solution. Because we assume the model response to parameters is not completely linear, we compute two variations of the derivative (y_d : solution obtained by doubling a parameter, y_h : solution obtained by halving a parameter) and find a least-squares fit for the optimal x values to determine Δp_1 (see Equation 7 in the main text). Thus:

$$y_{true} = y_0 + \left(\frac{y_d - y_0}{2p_1 - p_1}\right) x_{1d} p_1 + \left(\frac{y_h - y_0}{\frac{p_1}{2} - p_1}\right) x_{1h} p_1,$$

where, x_{1d} and x_{1h} are the least-squares fitted adjustment coefficients related to the doubling and halving of parameter p_1. Our H matrix carries the numerator of both these perturbation derived derivatives, $(y_d - y_0)$ and $(y_h - y_0)$, when solving for x, but not the denominator, such that the negative perturbation must be scaled by 1/2, and leading to Equation 8 in the main text.

Data Availability Statement

The BGC-Argo data (Argo, 2000) "were collected and made freely available by the International Argo Program and the national programs that contribute to it (https://argo.ucsd.edu, https://www.ocean-ops.org). The Argo Program is part of the Global Ocean Observing System." The B-SOSE simulated chlorophyll EOFs results (Kuhn et al., 2024) discussed in this manuscript are available in a Zenodo public repository https://doi.org/10.5281/zenodo.11396471. Code to replicate the estimation of optimal parameters is available in a public repository https://github.com/Akuhnc/BSOSE_optimal_parameters (Kuhn, 2024).

Acknowledgments
This study is funded through the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project (NSF award OPP-1936222). MM also acknowledges funding from NSF awards OPP-2149501, OCE-1924388, and NASA Grant 80NSSC24K0243.

References

Argo. (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) [Dataset]. *SEANOE*. https://doi.org/10.17882/42182

Armand, L. K., Cornet-Barthaux, V., Mosseri, J., & Queguiner, B. (2008). Late summer diatom biomass and community structure on and around the naturally iron-fertilised Kerguelen Plateau in the Southern Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography*, *55*(5–7), 653–676. https://doi.org/10.1016/j.dsr2.2007.12.031

Baldry, K., Strutton, P. G., Hill, N. A., & Boyd, P. W. (2020). Subsurface Chlorophyll-a maxima in the Southern Ocean. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.00671

Belo Couto, A., Brotas, V., Mélin, F., Groom, S., & Sathyendranath, S. (2016). Inter-comparison of OC-CCI chlorophyll-*a* estimates with precursor data sets. *International Journal of Remote Sensing*, *37*(18), 4337–4355. https://doi.org/10.1080/01431161.2016.1209313

Brix, H., Menemenlis, D., Hill, C., Dutkiewicz, S., Jahn, O., Wang, D., et al. (2015). Using Green's Functions to initialize and adjust a global, eddying ocean biogeochemistry general circulation model. *Ocean Modelling*, *95*, 1–14. https://doi.org/10.1016/j.ocemod.2015.07.008

Carranza, M. M., Gille, S. T., Franks, P. J. S., Johnson, K. S., Pinkel, R., & Girton, J. B. (2018). When mixed layers are not mixed. Storm-driven mixing and bio-optical vertical gradients in mixed layers of the southern ocean. *Journal of Geophysical Research: Oceans*, *123*(10), 7264–7289. https://doi.org/10.1029/2018JC014416

Carroll, D., Menemenlis, D., Dutkiewicz, S., Lauderdale, J. M., Adkins, J. F., Bowman, K. W., et al. (2022). Attribution of space-time variability in global-ocean dissolved inorganic carbon. *Global Biogeochemical Cycles*, *36*(3), e2021GB007162. https://doi.org/10.1029/2021GB007162

Cullen, J. J. (2015). Subsurface chlorophyll maximum layers: Enduring enigma or mystery solved? *Annual Review of Marine Science*, *7*(1), 207–239. https://doi.org/10.1146/annurev-marine-010213-135111

de Villiers, S., Siswana, K., & Vena, K. (2015). In situ measurement of the biogeochemical properties of Southern Ocean mesoscale eddies in the Southwest Indian Ocean, April 2014. *Earth System Science Data*, *7*(2), 415–422. https://doi.org/10.5194/essd-7-415-2015

Dunne, J. P., Armstrong, R. A., Gnanadesikan, A., & Sarmiento, J. L. (2005). Empirical and mechanistic models for the particle export ratio. *Global Biogeochemical Cycles*, *19*(4), GB4026. https://doi.org/10.1029/2004GB002390

du Plessis, M., Swart, S., Ansorge, I. J., Mahadevan, A., & Thompson, A. F. (2019). Southern ocean seasonal restratification delayed by sub-mesoscale wind–front interactions. *Journal of Physical Oceanography*, *49*(4), 1035–1053. https://doi.org/10.1175/JPO-D-18-0136.1

Edwards, K. F., Litchman, E., & Klausmeier, C. A. (2013). Functional traits explain phytoplankton community structure and seasonal dynamics in a marine ecosystem. *Ecology Letters*, *16*(1), 56–63. https://doi.org/10.1111/ele.12012

El Jarbi, M., Rückelt, J., Slawig, T., & Oschlies, A. (2013). Reducing the model-data misfit in a marine ecosystem model using periodic parameters and linear quadratic optimal control. *Biogeosciences*, *10*(2), 1169–1182. https://doi.org/10.5194/bg-10-1169-2013

Eppley, R. W. (1992). Chlorophyll, photosynthesis and new production in the Southern California Bight. *Progress in Oceanography*, *30*(1–4), 117–150. https://doi.org/10.1016/0079-6611(92)90010-w

Fahnenstiel, G. L., McCormick, M. J., Lang, G. A., Redalje, D. G., Lohrenz, S. E., Markowitz, M., et al. (1995). Taxon-specific growth and loss rates for dominant phytoplankton populations from the northern Gulf of Mexico. *Marine Ecology Progress Series*, *117*, 229–239. https://doi.org/10.3354/meps117229

Fennel, K., Losch, M., Schröter, J., & Wenzel, M. (2001). Testing a marine ecosystem model: Sensitivity analysis and parameter optimization. *Journal of Marine Systems*, *28*(1–2), 45–63. https://doi.org/10.1016/s0924-7963(00)00083-x

Franks, P. J. (2009). Planktonic ecosystem models: Perplexing parametrizations and a failure to fail. *Journal of Plankton Research*, *31*(11), 1299–1309.

Friedrichs, M. A. M., Dusenberry, J. A., Anderson, L. A., Armstrong, R. A., Chai, F., Christian, J. R., et al. (2007). Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups. *Journal of Geophysical Research*, *112*(C08001), 1–22. https://doi.org/10.1029/2006jc003852

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*(2), 862–886. https://doi.org/10.1175/jcli-d-14-00117.1

Galbraith, E. D., Gnanadesikan, A., Dunne, J. P., & Hiscock, M. R. (2010). Regional impacts of iron-light colimitation in a global biogeochemical model. *Biogeosciences*, *7*(3), 1043–1064. https://doi.org/10.5194/bg-7-1043-2010

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*, 187–200. https://doi.org/10.3354/meps148187

Gifford, D., Fessenden, L., Garrahan, P. R., & Martin, E. (1995). Grazing by microzooplankton and mesozooplankton in the high-latitude North Atlantic Ocean: Spring versus summer dynamics. *Journal of Geophysical Research*, *10*(C4), 6665–6675. https://doi.org/10.1029/94jc00983

Gomi, Y., Fukuchi, M., & Taniguchi, A. (2010). Diatom assemblages at subsurface chlorophyll maximum layer in the eastern Indian sector of the Southern Ocean in summer. *Journal of Plankton Research*, *32*(7), 1039–1050. https://doi.org/10.1093/plankt/fbq031

Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The variable Southern Ocean carbon sink. *Annual Review of Marine Science*, *11*(1), 159–186. https://doi.org/10.1146/annurev-marine-121916-063407

Haëntjens, N., Boss, E., & Talley, L. D. (2017). Revisiting Ocean Color algorithms for chlorophyll a and particulate organic carbon in the Southern Ocean using biogeochemical floats. *Journal of Geophysical Research: Oceans*, *122*(8), 6583–6593. https://doi.org/10.1002/2017JC012844

Irwin, A. J., & Finkel, Z. V. (2017). Phytoplankton functional types: A trait perspective. *BioRxiv, 148312*.

Jackson, T., Sathyendranath, S., & Mélin, F. (2017). An improved optical classification scheme for the ocean colour essential climate variable and its applications. *Remote Sensing of Environment*, *203*, 152–161. https://doi.org/10.1016/j.rse.2017.03.036

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., et al. (2017). Biogeochemical sensor performance in the SOCCOM profiling float array. *Journal of Geophysical Research: Oceans*, *122*(8), 6416–6436. https://doi.org/10.1002/2017JC012838

Kuhn, A. M. (2024). Parameter estimation MATLAB code [code]. Retrieved from https://github.com/Akuhnc/BSOSE_optimal_parameters

Kuhn, A. M., Mazloff, M. R., Gille, S. T., & Verdy, A. (2024). B-SOSE chlorophyll vertical structure [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.11396471

Lande, R., & Wood, M. (1987). Suspension times of particles in the upper ocean. *Deep-Sea Research, Part A: Oceanographic Research Papers*, *1*(34), 61–72. https://doi.org/10.1016/0198-0149(87)90122-1

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C., Van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224. https://doi.org/10.1126/science.aab2620

Le Quéré, C., Buitenhuis, E. T., Moriarty, R., Alvain, S., Aumont, O., Bopp, L., et al. (2016). Role of zooplankton dynamics for southern ocean phytoplankton biomass and global biogeochemical cycles. *Biogeosciences*, *13*(14), 4111–4133. https://doi.org/10.5194/bg-13-4111-2016

Litchman, E., De Tezanos Pinto, P., Klausmeier, C. A., Thomas, M. K., & Yoshiyama, K. (2010). Linking traits to species diversity and community structure in phytoplankton. In L. Naselli-Flores & G. Rossetti (Eds.), *Fifty years after the "Homage to Santa Rosalia": Old and new paradigms on biodiversity in aquatic ecosystems* (pp. 15–28). Springer. https://doi.org/10.1007/978-90-481-9908-2_3

Mattern, J. P., Fennel, K., & Dowd, M. (2012). Estimating time-dependent parameters for a biological ocean model using an emulator approach (Vol. 96–97, 32–47).

Mazloff, M. R., Verdy, A., Gille, S. T., Johnson, K. S., Cornuelle, B. D., & Sarmiento, J. (2023). Southern ocean acidification revealed by biogeochemical-Argo floats. *Journal of Geophysical Research: Oceans*, *128*(5), e2022JC019530. https://doi.org/10.1029/2022JC019530

Melbourne-Thomas, J., Wotherspoon, S., Corney, S., Molina-Balari, E., Marini, O., & Constable, A. (2015). Optimal control and system limitation in a Southern Ocean ecosystem model. *Deep Sea Research Part II: Topical Studies in Oceanography*, *114*, 64–73. https://doi.org/10.1016/j.dsr2.2013.02.017

Menemenlis, D., Fukumori, I., & Lee, T. (2005). Using Green's functions to calibrate an ocean general circulation model. *Monthly Weather Review*, *133*(5), 1224–1240. https://doi.org/10.1175/MWR2912.1

Netjstgaard, J. C., Hygum, B. H., Naustvoll, L.-J., & Bamstedt, U. (2001). Zooplankton growth, diet and reproductive success compared in simultaneous diatom- and flagellate-microzooplankton-dominated plakton blooms. *Marine Ecology Progress Series*, *221*, 77–91. https://doi.org/10.3354/meps221077

Orsi, A. H., Whitworth, T., & Nowlin, W. D. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers*, *42*(5), 641–673. https://doi.org/10.1016/0967-0637(95)00021-W

Parslow, J. S., Boyd, P. W., Rintoul, S. R., & Griffiths, F. B. (2001). A persistent subsurface chlorophyll maximum in the Interpolar Frontal Zone south of Australia: Seasonal progression and implications for phytoplankton-light-nutrient interactions. *Journal of Geophysical Research*, *106*(C12), 31543–31557. https://doi.org/10.1029/2000jc000322

Prend, C. J., Gille, S. T., Talley, L. D., Mitchell, B. G., Rosso, I., & Mazloff, M. R. (2019). Physical drivers of phytoplankton bloom initiation in the southern ocean's Scotia Sea. *Journal of Geophysical Research: Oceans*, *124*(8), 5811–5826. https://doi.org/10.1029/2019JC015162

Rembauville, M., Blain, S., Caparros, J., & Salter, I. (2016). Particulate matter stoichiometry driven by microplankton community structure in summer in the Indian sector of the Southern Ocean. *Limnology & Oceanography*, *61*(4), 1301–1321. https://doi.org/10.1002/lno.10291

Rosso, I., Hogg, A. M., Matear, R., & Strutton, P. G. (2016). Quantifying the influence of sub-mesoscale dynamics on the supply of iron to Southern Ocean phytoplankton blooms. *Deep Sea Research Part I: Oceanographic Research Papers*, *115*, 199–209. https://doi.org/10.1016/j.dsr.2016.06.009

Russell, J. L., Stouffer, R. J., & Dixon, K. W. (2006). Intercomparison of the Southern Ocean circulations in IPCC coupled model control simulations. *Journal of Climate*, *19*(18), 4560–4575. https://doi.org/10.1175/jcli3869.1

Sarthou, G., Timmermans, K. R., Blain, S., & Tréguer, P. (2005). Growth physiology and fate of diatoms in the ocean: A review. *Journal of Sea Research*, *53*(1–2), 25–42. https://doi.org/10.1016/j.seares.2004.01.007

Sathyendranath, S., Brewin, R. J. W., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., et al. (2019). An ocean-colour time series for use in climate studies: The experience of the Ocean-Colour Climate Change Initiative (OC-CCI). *Sensors*, *19*, 4285. https://doi.org/10.3390/s19194285

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Marine Ecology Progress Series*, *383*, 73–84. https://doi.org/10.3354/meps07998

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., et al. (2017). Reviews and syntheses: Parameter identification in marine planktonic ecosystem modelling. *Biogeosciences*, *14*(6), 1647–1701. https://doi.org/10.5194/bg-14-1647-2017

Talley, L. D., Rosso, I., Kamenkovich, I., Mazloff, M. R., Wang, J., Boss, E., et al. (2019). Southern Ocean biogeochemical float deployment strategy, with example from the Greenwich Meridian line (GO-SHIP A12). *Journal of Geophysical Research: Oceans*, *124*(1), 403–431. https://doi.org/10.1029/2018jc014059

Timmermans, K. R., van der Wagt, B., & de Baar, H. J. W. (2004). Growth rates, half-saturation constants, and silicate, nitrate, and phosphate depletion in relation to iron availability of four large, open-ocean diatoms from the Southern Ocean. *Limnology & Oceanography*, *49*(6), 2141–2151. https://doi.org/10.4319/lo.2004.49.6.2141

Uchida, T., Balwada, D., Abernathey, R., Prend, C. J., Boss, E., & Gille, S. T. (2019). Southern ocean phytoplankton blooms observed by biogeochemical floats. *Journal of Geophysical Research: Oceans*, *124*(11), 7328–7343. https://doi.org/10.1029/2019JC015355

Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, *122*(9), 6968–6988. https://doi.org/10.1002/2016JC012650

Von Berg, L., Prend, C. J., Campbell, E. C., Mazloff, M. R., Talley, L. D., & Gille, S. T. (2020). Weddell sea phytoplankton blooms modulated by sea ice variability and polynya formation. *Geophysical Research Letters*, *47*(11), e2020GL087954. https://doi.org/10.1029/2020GL087954

Ward, B. A., Friedrichs, M. A. M., Anderson, T. R., & Oschlies, A. (2010). Parameter optimization techniques and the problem of under-determination in marine biogeochemical models. *Journal of Marine Systems*, *81*(1–2), 34–43. https://doi.org/10.1016/j.jmarsys.2009.12.005

Westwood, K. J., Griffiths, F. B., Webb, J. P., & Wright, S. W. (2011). Primary production in the Sub-Antarctic and Polar Frontal zones south of Tasmania, Australia; SAZ-Sense survey, 2007. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(21–22), 2162–2178. https://doi.org/10.1016/j.dsr2.2011.05.017

Wright, S. W., & van den Enden, R. L. (2000). Phytoplankton community structure and stocks in the East Antarctic marginal ice zone (BROKE survey, January–March 1996) determined by CHEMTAX analysis of HPLC pigment signatures. *Deep Sea Research Part II: Topical Studies in Oceanography*, *47*(12–13), 2363–2400. https://doi.org/10.1016/s0967-0645(00)00029-1

Wunsch, C., & Heimbach, P. (2008). How long to oceanic tracer and proxy equilibrium? *Quaternary Science Reviews*, *27*(7–8), 637–651. https://doi.org/10.1016/j.quascirev.2008.01.006

Wunsch, C., & Heimbach, P. (2013). Dynamically and kinematically consistent global ocean circulation and ice state estimates. In *International geophysics* (Vol. 103, pp. 553–579). Elsevier. https://doi.org/10.1016/b978-0-12-391851-2.00021-0