

Spatially Varying Biogeochemical Parameter Estimation in a Global Ocean Model

N. Mamnun^{1,2} , C. Völker¹ , M. Vrekoussis^{3,4,5} , and L. Nerger¹ 

¹Alfred-Wegener-Institut (AWI), Helmholtz Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany, ²Now at Department of Physics, Imperial College London, London, UK, ³Institute of Environmental Physics (IUP), University of Bremen, Bremen, Germany, ⁴Center of Marine Environmental Sciences (MARUM), University of Bremen, Bremen, Germany, ⁵Climate and Atmosphere Research Center (CARE-C), The Cyprus Institute, Nicosia, Cyprus

Key Points:

- Spatially varying ocean biogeochemical model parameters are estimated by assimilation of ocean color data with an ensemble Kalman filter
- The spatial variation of the resulting parameter estimates is consistent with that reported from the observations
- Simulations conducted using the estimated spatially varying parameter values perform better than simulations using uniform values

Correspondence to:

N. Mamnun,
nabir.mamnun@awi.de;
n.mamnun@imperial.ac.uk

Citation:

Mamnun, N., Völker, C., Vrekoussis, M., & Nerger, L. (2025). Spatially varying biogeochemical parameter estimation in a global ocean model. *Journal of Geophysical Research: Oceans*, 130, e2025JC022752. <https://doi.org/10.1029/2025JC022752>

Received 15 APR 2025

Accepted 1 NOV 2025

Author Contributions:

Conceptualization: N. Mamnun, C. Völker, M. Vrekoussis, L. Nerger
Formal analysis: N. Mamnun
Funding acquisition: L. Nerger
Methodology: N. Mamnun, C. Völker, L. Nerger
Software: N. Mamnun, L. Nerger
Supervision: C. Völker, M. Vrekoussis, L. Nerger
Visualization: N. Mamnun
Writing – original draft: N. Mamnun
Writing – review & editing: N. Mamnun, C. Völker, M. Vrekoussis, L. Nerger

Abstract Ocean biogeochemical (BGC) models are key tools for investigating ocean biogeochemistry and the global carbon cycle. These models contain many uncertain and often poorly known process parameters that are treated as constant values. This study addresses this limitation by estimating spatially and temporally varying parameters in the Regulated Ecosystem Model 2 (REcoM2) through the assimilation of satellite-derived chlorophyll-a data using an ensemble Kalman filter. Nine key BGC parameters were optimized, significantly improving the model's performance. Utilizing the optimized parameters in the model results in a 26% reduction in root mean square error for surface chlorophyll-a concentrations compared to simulations with uniform parameters, with the spatial patterns of parameter estimates aligning well with observed distributions. These findings underscore the benefits of incorporating spatially and temporally varying parameters for enhancing model accuracy and understanding BGC variability.

Plain Language Summary Ocean biogeochemical (BGC) models are key tools for studying the global carbon cycle and BGC processes. These models often rely on uncertain parameters that are not precisely known, thus reducing the accuracy of simulations. Usually, these parameters are kept constant in time and space in the models, but in reality, they can vary over time and across different ocean regions. In this study, we combined satellite data with an ocean BGC model called Regulated Ecosystem Model 2, to estimate more realistic spatially and temporally varying parameter values. To do this, we applied a technique called data assimilation, which objectively searches for the best values for the uncertain parameters by optimally combining the simulated output and observational data. The updated model scheme reduced errors in surface chlorophyll-a concentration predictions by 26%, making it more aligned with observations and increasing confidence in its use for understanding ocean biogeochemistry and the carbon cycle.

1. Introduction

Ocean biogeochemical (BGC) models are powerful tools for studying ocean BGC processes and understanding their role in the global carbon cycle. BGC models are an essential component of Earth system models used to compute climate projections (Orr et al., 2017). They play a central role in quantifying the patterns and rates of ocean anthropogenic CO₂ uptake (see Crisp et al., 2022) and in estimating the global carbon budget (see Hauck et al., 2020). The latter is important because the global ocean absorbs more than a quarter of the anthropogenic emissions of CO₂ (Friedlingstein et al., 2022). Toward this direction, ocean BGC models are pivotal for (a) characterizing future ocean CO₂ uptake and its sensitivity to climate change under different policy scenarios (see Crisp et al., 2022), (b) assessing the predictability of global-scale atmosphere-ocean CO₂ flux relevant to carbon policy and management (Ilyina et al., 2021), and (c) investigating potential CO₂ removal (see Gattuso et al., 2018), ranging from the efficacy of net CO₂ uptake to the permanence of carbon storage, method verification, or carbon accounting, scalability, and environmental impacts in a local, regional, and global scale (National Academies of Sciences, Engineering and Medicine, 2022).

Beyond their role in global carbon cycle research, BGC models are used to investigate ocean deoxygenation (e.g., Andrews et al., 2017; Bopp et al., 2017) and ocean acidification (e.g., Gehlen et al., 2007; Ilyina et al., 2009; Krumhardt et al., 2019), and to study compound events with overlapping extremes of acidification, marine heat waves, and deoxygenation (e.g., Gruber et al., 2021; Hauri et al., 2021). These models also assess the economic impact of climate change on fisheries (e.g., Loukos et al., 2003) and project changes in fish catch potential (e.g., Cheung et al., 2010; Lam et al., 2016). BGC models are further used to develop marine environmental

© 2025. The Author(s).

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

applications and services (e.g., Gutknecht et al., 2019; Jones et al., 2016) and to generate reanalysis data sets (e.g., Carroll et al., 2020; Ciavatta et al., 2016).

Ocean BGC models describe the transformations of BGC constituents by ecosystem growth and interactions, incorporating the spatial distributions of various elements they represent. BGC constituents include nutrients, functional plankton groups, nonliving organic matter, dissolved gases, and variables of the inorganic carbon system contained in seawater. Whether by choice or necessity, each BGC transformation in the model is described by a simplified process formulation known as parameterizations. These parameterizations require process parameters to make the equations complete and solvable. Ocean BGC models include various biophysical processes and therefore involve numerous process parameters (see Fennel et al., 2022). The uncertainty of these parameter values is quite large (see Schartau et al., 2017), leading to possibly significant uncertainty in the model outputs (Leles et al., 2018; Mamnun et al., 2023; Prieur et al., 2019).

Ocean BGC models describe the diversity of numerous plankton species by a limited number of functional types. The available reference parameter values related to the parameterizations describing the dynamics of these functional types are usually taken from laboratory experiments targeting single species, while in the model, they are applied broadly to describe whole classes of organisms. The values of the parameters depend on the physical and BGC context (see Follows et al., 2007), which influences the distribution of ecosystem species and the acclimation of individual species. Thus, in reality, the parameter values vary spatially and temporally, while in practice, they are used as constants across space and time in the model simulations.

In this context, data assimilation (DA) helps in estimating the uncertain values for BGC parameters. DA combines the model variables with related observational data in a quantitative way, achieving an optimal match between simulation output and observations while accounting for their uncertainty. This process also leverages correlations between model parameters and observed variables to estimate optimal parameter values.

DA requires evaluating the model multiple times—variational methods necessitate multiple iterations to minimize the cost function and achieve an optimal fit between the model and observations, while sequential methods require an ensemble of model runs to capture model uncertainties. Due to the high computational expense of running a data assimilative model multiple times over a large three-dimensional (3-D) domain, parameter optimization is often carried out in one-dimensional (1-D) BGC model configurations. Parameter values estimated from a 1-D assimilative application are then used in a 3-D implementation (e.g., McDonald et al., 2012; Oschlies & Schartau, 2005; St-Laurent et al., 2017).

Studies estimating BGC parameters in multiple locations (e.g., Friedrichs et al., 2007; Gharamti et al., 2017; Mamnun et al., 2022; Schartau & Oschlies, 2003; Tjiputra et al., 2007) found different estimated parameter values (EPVs) across locations, reflecting that BGC parameters vary spatially depending on the real-world physical and BGC context. Treating these parameters as uniform across space and time in simulations, as is commonly done, likely introduces significant uncertainties and reduces the accuracy of the model results. As such, spatially varying parameters can likely improve the model performance and realism. However, it is important to note that significant spatial variations in parameter values partly come from the simple representation of plankton diversity in these models (T. R. Anderson, 2005; Friedrichs et al., 2007). Using a small number of functional groups to represent diverse microbial communities in globally different biomes naturally results in greater spatial variability of parameter values. More complex models with more plankton functional groups might show smaller spatial variations in parameter values, but increasing the number of parameters can lead to more uncertainty (weak identifiability) unless observational data are sufficiently dense to constrain reliable estimation (Turner & Gardner, 2015; Ward et al., 2012).

BGC parameter estimation using 2-D and 3-D models reveals considerable spatial variation in the EPVs (Doron et al., 2013; Losa et al., 2004; Simon et al., 2015; Singh et al., 2022; Tjiputra et al., 2007; Xu et al., 2022). Losa et al. (2004) estimated six BGC parameters in an array of simple box models in the North Atlantic Ocean and obtained varying parameters in different cells. Tjiputra et al. (2007) estimated spatially varying BGC parameters using an adjoint method by assimilating satellite chlorophyll-a concentration. They found that using estimated spatially variable parameters improved the global simulation of net primary production (NPP). Doron et al. (2013) estimated five spatially varying BGC parameters by assimilating ocean color-derived chlorophyll-a data into a 3-D regional model. They found better model-data agreement using spatially varying estimated parameters than the reference simulation using uniform ones. Estimating four spatially varying BGC parameters, Simon

et al. (2015) found regional patterns of estimated parameters similar to the Longhurst provinces (Longhurst, 2007) in the regions where the model used performs reasonably. They demonstrated that BGC predictions generally benefit from spatially varying parameter estimates. Xu et al. (2022) estimated spatially varying BGC parameters in the Bohai, Yellow, and East China Seas, assimilating satellite chlorophyll-a data using an adjoint method and obtaining reasonable parameter values. Using an idealized twin (identical twin) experiment, Singh et al. (2022) showed that estimating spatially varying ocean BGC parameters is feasible using ensemble-based DA techniques in global-scale models.

Incorporating temporally varying parameters can significantly improve the agreement between models and observations (e.g., Mattern et al., 2013, 2014; Roy et al., 2012; Simon et al., 2015). Simon et al. (2015) specifically identified seasonal patterns in the estimated parameters and advocated using time-dependent parameters in ocean BGC models. However, they also highlighted that in regions with substantial model errors, the parameter values either converge to extreme values, resulting in larger model errors, or may diverge toward a high ensemble spread. Singh et al. (2022) also noted that even in an ideal model setting, certain BGC parameters do not converge to their true values when significant model errors occur.

The high parameter uncertainty of BGC models, combined with sparse and error-prone BGC observations, poses significant challenges in establishing relationships among BGC parameters, model state variables, and observations. In a high-dimensional BGC model, the number of unknown state variables and parameters of the model exceeds the available observations, creating an underdetermined inverse problem that the DA seeks to solve using a small number of observations to estimate a large set of unknowns. Despite the benefits of using satellite-derived surface observations for BGC state estimations (e.g., Ford & Barciela, 2017; Goodliff et al., 2019; Gregg, 2008; Nerger & Gregg, 2007, 2008; Pradhan et al., 2019, 2020), it remains unclear how effectively they can constrain uncertain BGC parameters in a 3-D global ocean model and estimate their spatially varying values. Additionally, the response of the joint state-parameter estimation (JSPE) algorithm to highly nonlinear relationships and non-Gaussian error statistics, such as those found in ocean BGC models, is poorly understood.

The objective of this study is to estimate selected spatially and temporally varying parameters in a global ocean BGC model by assimilating chlorophyll-a concentration retrieved from satellite ocean color measurements and to study the effect of spatially varying parameters on BGC processes and modeling. We estimate nine spatially and temporally varying parameters, chosen based on a global sensitivity analysis (Mamnun et al., 2023), of the BGC model Regulated Ecosystem Model Version 2 (REcoM2, Hauck et al., 2013; Schourup-Kristensen et al., 2014). We further assess the skill of REcoM2 in a simulation using the estimated spatially and temporally varying parameters. We discuss the effect of estimated spatially and temporally varying parameters on BGC processes and modeling.

2. Materials and Methods

2.1. The Coupled Physical-Biogeochemical Model

In this study, we used the coupled physical-biogeochemical model consisting of the Massachusetts Institute of Technology general circulation model (MITgcm, Marshall, Adcroft, et al., 1997; Marshall, Hill, et al., 1997) to simulate the ocean dynamics and tracer transports and the REcoM2 (Hauck et al., 2013; Schourup-Kristensen et al., 2014) simulating the BGC processes.

MITgcm is a 3-D, finite-volume, general circulation model. It solves the time-dependent, Boussinesq-approximated Navier-Stokes equations with or without hydrostatic approximation and conservation equations for mass and energy. The nonhydrostatic capabilities allow the users to use the model to study small-scale and global processes. A sea ice model (Losch et al., 2010) is integrated with MITgcm.

REcoM2 describes two phytoplankton groups—nanophytoplankton and diatoms—and a generic heterotrophic zooplankton class. The nanophytoplankton has an implicit representation of calcifiers. REcoM2 further has one class of organic sinking particles whose sinking speed increases with depth. REcoM2 describes the carbon cycle; the nutrients nitrogen, silicon and iron; and chlorophyll. REcoM2 is a so-called quota model (Droop, 1983). As such, the intracellular stoichiometry of carbon, nitrogen, and chlorophyll ($C:N:Chl$) pools for nanophytoplankton and carbon, nitrogen, silicate, and chlorophyll ($C:N:Si:Chl$) pools for diatoms are allowed to respond dynamically to the environmental conditions following Geider et al. (1998) and Hohn (2009) for the Si quota. The intracellular iron pool is a function of the intracellular nitrogen concentration (fixed iron to nitrogen

ratio), as iron is physiologically mainly linked to nitrogen metabolism and the photosynthetic electron transport chain (Behrenfeld & Milligan, 2013; Geider & La Roche, 1994). Dead organic matter is transferred to detritus by aggregation and grazing by the zooplankton group, and the sinking and advection of detritus are represented explicitly. REcoM2 has two external iron sources: atmospheric dust deposition and sedimentary input. The iron cycle in the model is driven by biological uptake, remineralization, and scavenging onto biogenic and lithogenic particles.

2.2. Model Setup

This study employs a global model configuration of the so-called Lat-Lon-Cap 90 grid (LLC90, Forget et al., 2015). It represents the Earth in a Cartesian curvilinear coordinate system using a cubed-sphere structure in the Northern Hemisphere and a dipolar grid arrangement in the Southern Hemisphere.

The horizontal model grid resolution varies spatially, ranging from 22 to 110 km, with the highest resolutions at high latitudes and the lowest resolution at midlatitudes. There are 50 vertical levels. The vertical grid spacing increases with depth from 10 m at the surface layer to 456.5 m at the bottom layer.

We initialized the temperature, salinity, and dissolved oxygen (DO_2) fields using the winter mean data from the World Ocean Atlas 2018 (WOA-18, Boyer et al., 2018). For the dissolved inorganic nitrogen and silica (DSi) fields, we used annual climatology from WOA-18 (García et al., 2019). The total alkalinity (ALK) and the dissolved inorganic carbon fields were initialized with mapped climatological data from the GLObal Ocean Data Analysis Project (GLODAPv2, Lauvset et al., 2016). To initialize the dissolved iron (DFe) field, we relied on concentrations obtained from a previous Pelagic Interactions Scheme for Carbon and Ecosystem Studies (PISCES, Aumont et al., 2015) model run by Aumont et al. (2015). We used the monthly dust deposition field from the present-day simulation of Albani et al. (2014) to compute DFe input flux from the atmosphere, assuming 3.5% iron content in dust particles and 2% iron solubility. All other passive tracers were initialized with small arbitrary values.

We utilized the interannually varying atmospheric forcing (2 m air temperature, specific humidity, downward radiation, and 10 m wind and precipitation) from the 6-hourly ERA-Interim reanalysis fields (Dee et al., 2011) from 1992 to 2018 and from ERA-5 hourly reanalysis fields (Hersbach et al., 2020) for 2019 to 2021. For consistency, hourly ERA-5 data were interpolated to 6-hourly data. Freshwater runoff is forced using the river and ice-sheet melting runoff data from the JRA55-do data set (Tsujino et al., 2018).

2.3. Data Assimilation Methods

We employ an ensemble-based Kalman Filter (EnKF, see review by Vetra-Carvalho et al., 2018). EnKFs utilize an ensemble of model state realizations to represent the state estimate and to account for model uncertainties and covariances between different variables contained in the so-called state vector. We implemented the ensemble DA with the Parallel Data Assimilation Framework (PDAF, Nerger & Hiller, 2013), an open-source software (accessible at <http://pdaf.awi.de>; Nerger, 2023). PDAF offers comprehensive and parallelized ensemble filter algorithms and support for parallel ensemble integrations. It can be coupled into an existing model code as a library. PDAF currently provides not only many variants of EnKFs but also particle filters and variational methods.

2.3.1. Error Subspace Transform Kalman Filter

We used the localized (Nerger et al., 2012a) error subspace transform Kalman filter (ESTKF, Nerger et al., 2012b) for all of our DA experiments in this study. The ESTKF is a highly efficient filter for high-dimensional models. As an ensemble square root filter, the ESTKF computes the weights for the ensemble transformation directly in the error subspace represented by the ensemble and can be used with a deterministic minimum transformation, allowing the use of small ensembles.

Domain localization (Nerger et al., 2012a) with observation weighting is used in this study. In the domain localization approach, each vertical column of the model grid is considered a disjoint local domain. Only observations with a distance smaller than a cutoff radius are used in the analysis step for a given local domain. The weight of the observations is decreased with increasing distance. The localization radius was tested with different

Table 1

The Nine Selected BGC Parameters With Their Symbol, Unit, and Default Value

	Parameter	Symbol	Unit	Default value
1	Nanophytoplankton initial slope of the photosynthesis-irradiance curve	α_{Nano}	$\text{mmol C (mg Chl)}^{-1} (\text{Wm}^{-2}\text{d})^{-1a}$	0.14
2	Diatom initial slope of the Photosynthesis-irradiance curve	α_{Dia}	$\text{mmol C(mg Chl)}^{-1} (\text{Wm}^{-2}\text{d})^{-1}$	0.19
3	Nanophytoplankton maximum photosynthesis rate	$\mu_{Nano}^{\max}$	d^{-1}	3.00
4	Diatom maximum photosynthesis rate	$\mu_{Dia}^{\max}$	d^{-1}	3.50
5	Maximum chlorophyll to nitrogen ratio of nanophytoplankton	$q_{Nano}^{Chl:N\max}$	$\text{mg Chl (mmol N)}^{-1}$	3.15
6	Maximum chlorophyll to nitrogen ratio of diatom	$q_{Dia}^{Chl:N\max}$	$\text{mg Chl (mmol N)}^{-1}$	4.2
7	Nanophytoplankton chlorophyll degradation rate	d_{Nano}^{Chl}	d^{-1}	0.10
8	Diatom chlorophyll degradation rate	d_{Dia}^{Chl}	d^{-1}	0.10
9	Maximum grazing rate of zooplankton	ξ	$\text{mmol N m}^{-3} \text{d}^{-1}$	2.4

^aThe unit indicates millimoles of carbon produced per milligram of chlorophyll, adjusted for the amount of light energy received daily per square meter.

values and set to 5 grid cells based on trial and error to improve the assimilation results. No ensemble inflation is used in this study.

2.3.2. Ensemble Generation

We apply perturbations to nine BGC process parameters of the REcoM2 model to generate an ensemble of BGC model states. Our goal is to minimize the uncertainty associated with the initial parameter values through stochastic estimation using satellite-derived surface chlorophyll-a concentration data. Accordingly, we targeted the perturbation to parameters showing high sensitivity to the model outputs of surface chlorophyll-a concentration. The selection of these nine parameters was based on the sensitivity analysis conducted by Mamnun et al. (2023).

Among the nine selected parameters, there are four phytoplankton photosynthesis-irradiance parameters (H. A. Bouman et al., 2018), two cell quotas—the maximum chlorophyll to nitrogen ratios, two parameters for chlorophyll degradation, and one parameter related to zooplankton grazing. Table 1 lists the selected parameters, including their symbols, units, and default values, which are the values used by Hauck et al. (2013).

We utilized a lognormal distribution function to generate random perturbations for each of the selected parameters. The respective default values of the parameters were considered the expected value of the distribution, with a standard deviation of 25% of the default value. These perturbations induce subtle differences in the BGC processes across different ensemble members, consequently generating a diverse range of model outcomes. We defined each selected parameter as a two-dimensional (2-D) field within the model and established that each ensemble member has a different set of parameter values. In each ensemble member, the initial values of these parameters were identical across all 2-D grid points. Running these models generates an ensemble of simulations with spread in the BGC fields, allowing us to observe the variability in outcomes induced by the perturbations.

2.3.3. The (Augmented) State Vectors

From the BGC model, we included eight model state variables that describe the two phytoplankton groups in the state vector of variables that are directly updated by the filter method. The selected state variables are as follows:

1. Nanophytoplankton biomass content of carbon
2. Diatom biomass content of carbon
3. Nanophytoplankton biomass content of nitrogen
4. Diatom biomass content of nitrogen
5. Nanophytoplankton calcium carbonate concentration
6. Biogenic silica concentration of diatoms
7. Nanophytoplankton chlorophyll-a concentration
8. Diatom chlorophyll-a concentration

The assimilated observation type is the total surface chlorophyll-a concentration. Since the chlorophyll is represented in the model by two group-specific concentrations, we include a variable “total chlorophyll-a concentration” in the state vector by summing up the chlorophyll-a concentration of nanophytoplankton and diatoms. The observation operator for this variable, total chlorophyll-a concentration, selects corresponding values from the state vector. Subsequently, the ESTKF updates the model state variables and the parameters through cross-covariances with the total chlorophyll-a concentration.

We applied the state augmentation approach (J. L. Anderson, 2001), which merges state variables and parameters into an augmented state vector, treating the parameters as time-varying variables. This augmented state vector method facilitates the estimation of parameters, given the observation of specific variables and the multivariate covariances between model parameters and model variables. The analysis increments were computed within each 5-day assimilation window, corresponding to the temporal resolution of the assimilated observations (see Section 2.4). The model state and parameters were then instantaneously updated at the corresponding model time step by directly applying the increments, without imposing any bounds or prior limits on their values.

EnKFs are only optimal if the errors have a Gaussian distribution. This is not fulfilled in an ocean BGC model due to non-Gaussian state distributions of concentration variables. To address this limitation, we transformed the chlorophyll-a concentrations, both from the model and observations, into a logarithmic scale, based on the assumption that their distribution is lognormal (Campbell, 1995). This logarithmic transformation was applied also to the other BGC state variables in the state vector, assuming a similar probability distribution as the chlorophyll-a concentration. Moreover, since the parameters were perturbed following a lognormal distribution, we also log-transformed their values before each assimilation cycle for analysis.

An advantage of the log-transformation approach is that it prevents estimating unrealistic negative concentrations or parameters during the assimilation process, which could otherwise arise due to the Kalman filter's linear combination of model estimates and observations in combination with the Gaussian assumption. By employing this technique, we ensure that the assimilation maintains the integrity of BGC variables and parameters as positive quantities, aligning with their inherent natural constraints.

2.4. Observational Data

The assimilated observations are remotely sensed surface chlorophyll-a concentrations obtained from the European Space Agency's Ocean Colour Climate Change Initiative (OC-CCI; Sathyendranath et al., 2019) project product, Version 6.0 (last access 03 March 2023). This product was created by merging satellite data from multiple sensors. The assimilated data set consists of a 5-day mean level 3 binned data presented on a global sinusoidal grid with a resolution of 4 km.

The data set includes per-pixel error statistics estimated by analyzing matchups between in situ data and ocean color. We used these error statistics as observation uncertainty in the DA analysis. Specifically, we computed and assimilated unbiased values of chlorophyll-a concentration interpolated on the model grid analogous to Pradhan et al. (2019).

2.5. Experiments

We performed four experiments. The experiments are as follows.

1. Reference single forward run (REF_FOR): We conducted a single 30-year model run from 1992 to 2021, using the initial conditions specified in Section 2.2 and the default parameter values (DPVs) in Table 1. This simulation served as the basis for initializing the ensemble reference simulations.
2. Reference 40-member ensemble free run (REF_FRE): We performed a 40-member ensemble simulation spanning three years from 2018 to 2020. The ensemble models were initialized using the restart files from REF_FOR. For each ensemble state, perturbations to the DPVs of the nine selected parameters were applied. The hydrodynamic model states were kept unchanged.
3. Joint state-parameter estimation (JSPE): We conducted a joint BGC state and parameter estimation experiment by assimilating satellite-derived chlorophyll-a concentration data using a 40-member ensemble. In this experiment, we augmented the state vector by 2-D fields of the selected BGC parameters and updated them in each assimilation cycle together with the state variables. The initial model states were identical to the initial states of REF_FRE. The year 2018 was considered a spin-up without applying DA. In the next year (2019),

- only state estimation was performed to minimize the model uncertainties sourced from other than parameter perturbations. Subsequently, the BGC state variables and parameters were updated simultaneously by the DA in 2020.
4. Adjusted simulation (ADJ_FOR): We run a single forward model with the EPVs resulting from the experiment JSPE. This is a single 30-year model run identical to the experiment REF_FOR but using the estimated spatially varying parameter values.

3. Results

3.1. Impact of the Joint State-Parameter Estimation on the State Variables

For accurate parameter estimation, it is essential that the assimilation effectively constrains the model states. The value of the model state variables, especially at the surface, can significantly impact near-surface BGC processes in the model. In this study, all assimilation experiments improved the simulation of surface chlorophyll-a concentration compared to the free run (REF_FOR), as measured by root mean square errors (RMSE) of the log-transformed model state relative to the assimilated observations. Specifically, the JSPE reduced the RMSE between observations and model output by 51% in log space, compared to the RMSE from the REF_FOR.

Figure 1 shows the area-weighted RMSE for log-transformed surface chlorophyll-a for the experiments REF_FOR and JSPE when compared to satellite observations. During spring in both hemispheres, the model tends to overestimate surface chlorophyll-a concentrations. This overestimation contributes to elevated RMSEs from March to May and September to November on a global scale (Figure 1a). Pradhan et al. (2019) observed a similar temporal evolution of RMSEs using the same model with a coarser grid resolution. The model (REF_FOR) is more skillful in the equatorial region (Figure 1b) than the higher latitudes. The joint parameter-state estimation was more effective in the midlatitudes than in the low and high latitudes (Figures 1c and 1d). However, RMSEs are higher at high latitudes during the spring in each respective hemisphere for both the free-run model and the JSPE (Figures 1e and 1f). This is most likely because of the low number of observations available in these regions. Although JSPE significantly reduces RMSEs during the northern spring bloom, residual errors remain substantial.

We evaluated the spatial distributions of the logarithm of mean surface chlorophyll-a concentration for April and September 2020 (Figure 2). We have chosen these 2 months because the global area-weighted RMSEs are higher than in the other months. We compare the ensemble mean of surface chlorophyll-a concentration simulated without DA (REF_FRE) and with DA (Analysis of JSPE) to observations. REF_FRE performed poorly with high positive bias in high latitudes. The JSPE experiment shows a better field with spatial patterns closer to the observations than the REF_FRE. The model still overestimates the chlorophyll-a concentration compared to the observations, particularly in September. However, the model underestimates the observed values in the subtropical South Pacific Ocean.

3.2. Estimated Parameter Values

In this section, we examine the spatial variability of the estimated BGC model process parameters. The DA estimates the parameters using their correlation with the total chlorophyll-a field. The enhanced performance of the DA simulation compared to the free run for state estimation (Section 3.1) is partly due to the simultaneous optimization of the model parameters. Figure 3 shows EPVs in the global ocean for the nine selected parameters at the end of the year 2020 after 1 year of JSPE. The global average, minimum, maximum, and standard deviation of spatially varying estimated parameters are summarized in Table 2. The values of the parameters vary regionally depending on the physical and BGC conditions. All nine estimated parameters exhibit values larger and smaller than their default value, showing that DA optimizes the model's parameterization regionally by utilizing the correlation to the observations. In the following subsection, we assess the effect of the parameter estimation separately for groups of related parameters.

3.2.1. Initial Slope of the Photosynthesis-Irradiance (P-I) Curve of Nanophytoplankton (α_{Nano}) and Diatoms (α_{Dia})

The initial slope of the P-I curve (α) expresses the efficiency with which an organism conducts photosynthesis under light-limited conditions by characterizing the relationship between photosynthetic rate and light intensity

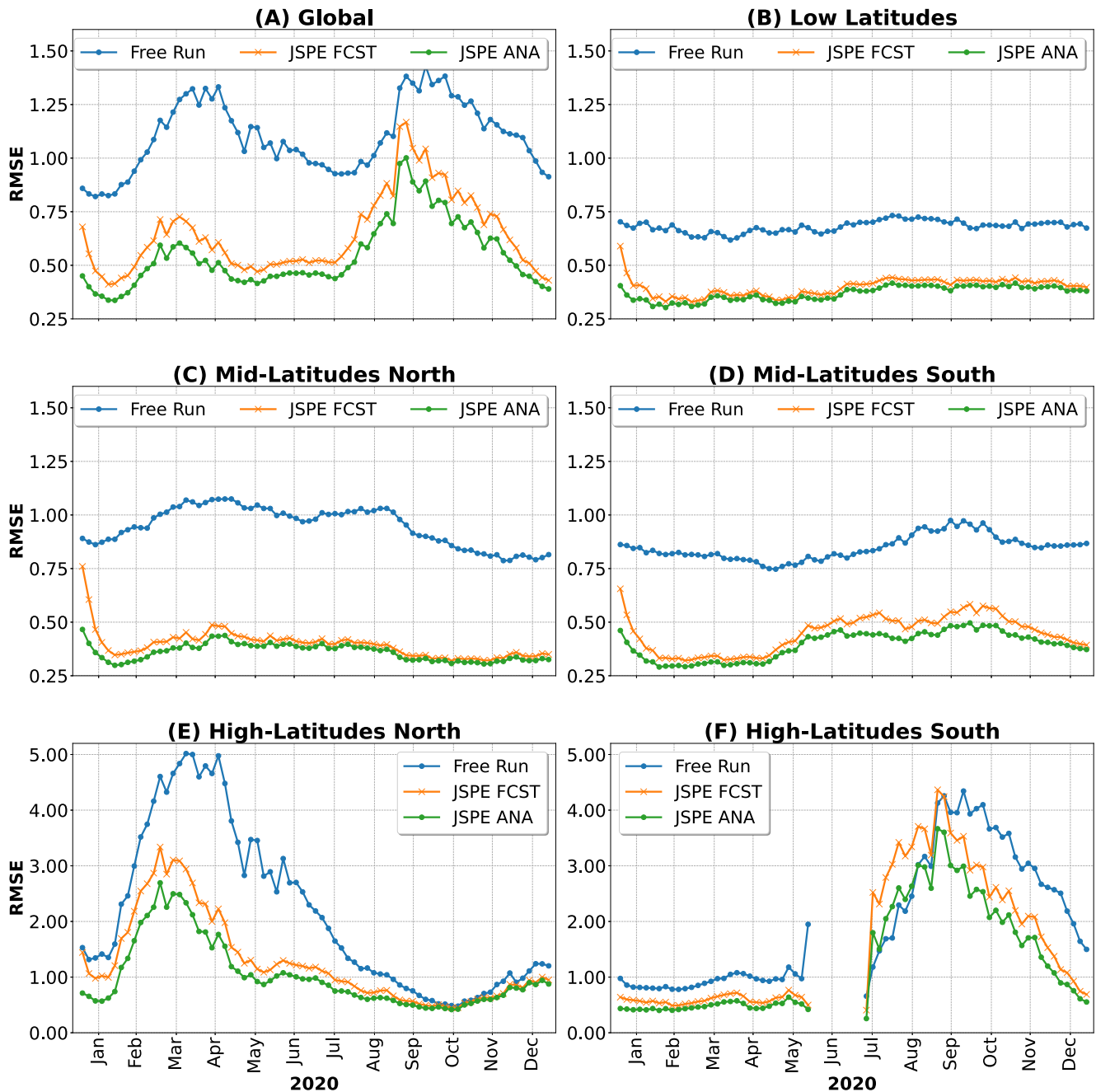


Figure 1. Comparison of area-weighted root mean square errors of log-transformed surface chlorophyll-a concentration of the experiments REF_FRE and joint state-parameter estimation relative to the OC-CCI data (a) for all available observations (global), (b) from the 10°S–10°N latitude (low latitude), (c) 10°N–50°N (midlatitude north), (d) 10°S–50°S (midlatitude south), (e) north of 50°N (high latitude north), and (f) south of 50°S (high latitude south).

(Denman, 2003). The estimated values of α_{Nano} and α_{Dia} are larger than their default values in most parts of the globe (Figures 3a and 3d). Accordingly, the global averages of both parameters increase (Table 2). A higher value of α indicates that a lower chlorophyll-a concentration is sufficient to achieve an equivalent primary production under light-limited conditions. Consequently, the model predicts adequate phytoplankton production with reduced chlorophyll-a during winter when light is limited. However, to compensate for the negative model bias in the subtropical South Pacific Ocean, the assimilation increased the values of α_{Nano} even though the light is not limited here. In contrast, α_{Dia} was only slightly increased in this region. Perhaps, diatom growth is limited here due to iron limitation and thus not sensitive in this region. The Barents Sea is another region where the values of

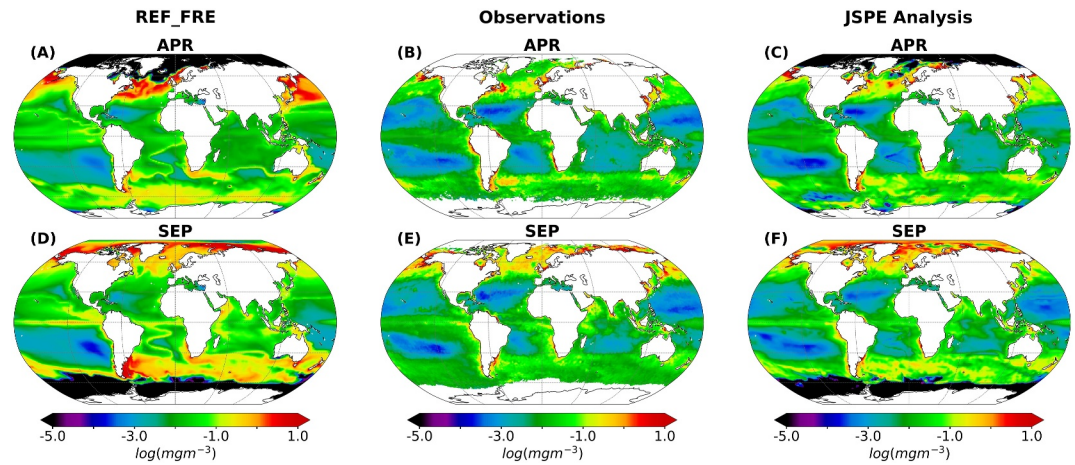


Figure 2. Monthly mean log-transformed surface chlorophyll-a concentrations for April 2020 (a–c) and September 2020 (d–f). From left to right: (REF_FRE); ESA OC-CCI data; and joint state-parameter estimation (JSPE) analysis results and difference of JSPE analysis results from REF_FRE.

both α_{Nano} and α_{Dia} are increased. This compensates for the underestimation of chlorophyll-a in REF_FRE (Figure 2). In the subantarctic zone of the Antarctic Circumpolar Current, the values of α_{Nano} decrease while the values of

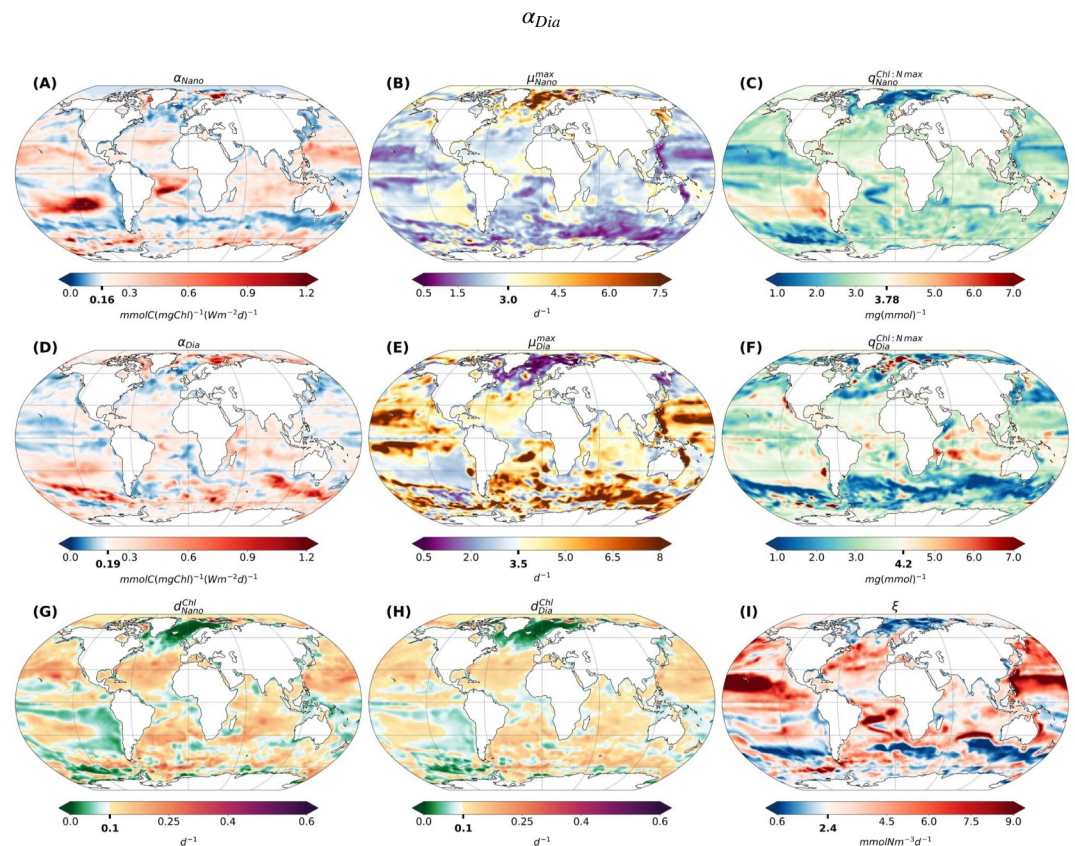


Figure 3. Estimated parameters values for the initial slope of the photosynthesis-irradiance curve of nanophytoplankton (a) and diatoms (d), maximum photosynthesis rate of nanophytoplankton (b) and diatoms (e), maximum chlorophyll to nitrogen ratio of nanophytoplankton (c) and diatoms (f), chlorophyll degradation rate of nanophytoplankton (g) and diatoms (h), and maximum grazing rate (i). The default value is indicated by the bold value indicated by the major tick mark on the color bar.

Table 2

The Default Value and Estimated Values (Global Average, Minimum, and Standard Deviation as a Percentage Relative to Global Average) of Spatially Varying Estimated Parameters

	Parameter	Default value	Average of estimate	Estimate minimum	Estimate maximum	Standard deviation
1	α_{Nano}	0.14	0.23	0.009	2.4	0.17 (73%)
2	α_{Dia}	0.19	0.26	0.01	1.5	0.19 (73%)
3	μ_{Nano}^{max}	3.00	2.67	0.38	17.72	1.14 (43%)
4	μ_{Dia}^{max}	3.50	4.20	0.20	29.84	2.06 (49%)
5	$q_{Nano}^{Chl:N max}$	3.15	3.02	0.44	7.87	0.74 (24%)
6	$q_{Dia}^{Chl:N max}$	4.2	3.32	0.23	40.43	1.31 (31%)
7	d_{Nano}^{Chl}	0.10	0.107	0.001	0.98	0.055 (51%)
8	d_{Dia}^{Chl}	0.10	0.1007	0.001	0.43	0.04 (40%)
9	ξ	2.4	3.18	0.12	20.34	0.24 (62%)

increase. This induces a shift from nanophytoplankton to diatoms. In the coastal areas, the values of α_{Nano} increase in general, whereas the values α_{Dia} show no clear pattern of changes in coastal areas. Overall, the patterns of the changes in the parameter values are different for α_{Nano} and α_{Dia} .

Compared to observations (e.g., H. A. Bouman et al., 2018; Marañón & Holligan, 1999), the estimated values of α_{Nano} are higher than the maximum observed values in the subtropical South Pacific gyre, the South Atlantic, and the Barents Sea. The values of α_{Dia} show no such extreme values. Elevated values of α_{Dia} are observed in the iron-limited South Pacific Ocean and the Barents Sea, though.

3.2.2. Maximum Photosynthesis Rate of Nanophytoplankton (μ_{Nano}^{max}) and Diatoms (μ_{Dia}^{max})

The maximum photosynthesis rate of phytoplankton (μ^{max}) defines the peak rate at which phytoplankton can transform inorganic carbon into organic matter through photosynthesis under optimal light and nutrient conditions (Denman, 2003). Unlike α_{Nano} and α_{Dia} , the parameter estimation changes the values of μ_{Nano}^{max} and μ_{Dia}^{max} in opposite directions—the global average of μ_{Nano}^{max} decreases from its default value, whereas that of μ_{Dia}^{max} increases (Table 2). Similarly, the spatial patterns of the estimated values of μ_{Nano}^{max} and μ_{Dia}^{max} show opposite signs (Figures 3b and 3e). Both nanophytoplankton and diatoms compete for similar resources, such as light and nutrients. The model has a competition term between nanophytoplankton and diatoms for a shared resource (Hauck et al., 2013), which might lead to an inverse relationship between their maximum photosynthesis rates to ensemble dynamics, where adjustments in these two parameters compensate each other. In general, the values of μ_{Nano}^{max} increase where the model strongly underestimates the chlorophyll-a, for example, the Arctic Atlantic Ocean and the subarctic Atlantic Ocean. In these regions, diatom productivity is inherently low, limiting the effectiveness of increasing μ_{Dia}^{max} at correcting model biases.

3.2.3. The Maximum Chlorophyll to Nitrogen Ratio in Nanophytoplankton ($q_{Nano}^{Chl:N max}$) and Diatoms ($q_{Dia}^{Chl:N max}$)

The maximum chlorophyll to nitrogen ratio ($q^{Chl:N max}$) defines the maximum of how much chlorophyll-a can be synthesized per unit of phytoplankton nitrogen (Geider et al., 1998; Omta et al., 2017). The global spatial average values of the estimated $q_{Nano}^{Chl:N max}$ and $q_{Dia}^{Chl:N max}$ are lower than their default values (Table 2). A lower value of $q^{Chl:N max}$ implies that a smaller maximum amount of chlorophyll can be produced by the cell per amount of cellular nitrogen under light-limiting conditions. The DA reduces the values of $q_{Nano}^{Chl:N max}$ and $q_{Dia}^{Chl:N max}$ over the large part of the global ocean (Figures 3c and 3f). This compensates for the overall overestimation of the surface chlorophyll-a without having a direct influence on nitrogen biomass. In contrast, the values of $q_{Nano}^{Chl:N max}$ are increased in the subtropical Pacific Ocean, where the model underestimates the surface chlorophyll-a concentrations, whereas the values of $q_{Dia}^{Chl:N max}$ are not changed much in this region. The latter is due to the fact that nanophytoplankton dominates in this region so that changes in diatoms have little effect on the total chlorophyll. The values $q_{Dia}^{Chl:N max}$ show locally extreme values in some regions, for example, the Norwegian Sea, the

Bellinghausen Sea, and the Amundsen Sea. These localized extreme values likely reflect the DA compensating for strong regional underestimations of chlorophyll-a, especially in areas where diatoms occasionally become important contributors to total chlorophyll-a in these high-latitude regions.

3.2.4. Chlorophyll Degradation Rate of Nanophytoplanktons (d_{Nano}^{Chl}) and Diatoms (d_{Dia}^{Chl})

The chlorophyll degradation rate (d^{Chl}) represents the rate at which chlorophyll is degraded. This affects the amount of chlorophyll available in phytoplankton, especially under long periods of light limitation or at the base of the euphotic zone. By reducing the amount of cellular nitrogen, it influences phytoplankton production. Though the global spatial average values of d_{Nano}^{Chl} and d_{Dia}^{Chl} are close to their default values (Table 2), the regional values vary spatially by two orders of magnitude (with 51% and 40% standard deviations of the global mean, respectively). The patterns of spatial variation of these two parameters are similar (Figures 3g and 3h) and directed by the model-data misfit. The particularly low values for both d_{Nano}^{Chl} and d_{Dia}^{Chl} in the northern Atlantic are another compensation for the seasonal underestimation of chlorophyll in this region.

3.2.5. The Maximum Grazing Rate of Zooplankton (ξ)

ξ represents the maximum possible rate at which zooplankton can consume phytoplankton under ideal conditions—a measure of the grazing pressure zooplankton can exert on phytoplankton populations. The global average value of ξ is higher than its default value (Table 2). A higher value of ξ can lead to a faster removal of phytoplankton from the system, thus decreasing the surface chlorophyll-a concentration, and a lowering of ξ usually leads to an increase in the surface chlorophyll-a concentration since more phytoplankton exists. The values of this parameter increase in a large part of the world's oceans to compensate for the overestimation of surface chlorophyll-a by the model. The parameter value decreased in the Arctic and subarctic Atlantic ocean, the south subtropical convergence zone, and some parts of the subantarctic water ring. The estimated ξ values vary strongly with a range from 0.12 to 20.34 mmol N m⁻³ d⁻¹. One possible reason for this extensive range is that we did not include the grazing efficiency (γ), because its effect is correlated to that of ξ , so that both parameters cannot be estimated independently. However, γ can also be highly sensitive for the simulation of the surface chlorophyll-a simulation (Mamnun et al., 2023), in the parameter estimation. Thus, a change in ξ could compensate for the constant γ .

The JSPE reduces the ensemble spread relative to the initial spread across all parameters (not shown). A spatial consistency is observed in the retrieved pattern values that align well with the spatial distribution of chlorophyll-a, though discrepancies persist. Small-scale noise is likely due to spurious correlations in our finite ensemble of 40 members. Regions, where the estimation does not converge, correspond to regions where the model exhibits deficiencies in simulating surface chlorophyll-a.

3.3. Temporal Evaluation Along the Atlantic Ocean

We monitored the time evolution of the EPVs in 12 Longhurst provinces (Longhurst, 2007) in the Atlantic Ocean (see Figure 4). Most changes in the parameters occur during the bloom periods. In the low and midlatitudes, most parameters reach approximately stable values by 30 DA cycles (May) and show minor variability over time. As expected, the parameters show stronger temporal variability in the high latitudes than in the low and midlatitudes.

The photosynthesis-irradiance parameters α_{Nano} and α_{Dia} exhibit large temporal variability in the polar provinces (Figures 4a and 4d), predominantly due to the latitudinal distribution of incident irradiance. In the Northern Hemisphere, increasing irradiance during spring enhances phytoplankton growth, resulting in elevated simulated chlorophyll-a concentrations. To correct this overestimation, the parameter estimation reduces the values of α_{Nano} and α_{Dia} in the ARCT and SARC provinces. Conversely, during summer, the values of both parameters continuously increase, surpassing their default values. They remain relatively stable during the low-light conditions from September until the end of the year. In the ANTA and SANT provinces, α_{Nano} and α_{Dia} remain above their default values. In midlatitudes and low latitudes, most variations in these parameters occur during the initial cycles of the parameter estimation process, after which they either stabilize or exhibit a discernible trend. Notably, α_{Nano} undergoes more pronounced changes than α_{Dia} in these midlatitudinal and low-latitudinal provinces. This is because diatoms are less abundant in these regions than nanophytoplankton, so that the DA has a more pronounced effect on α_{Nano} in midlatitudes and low latitudes.

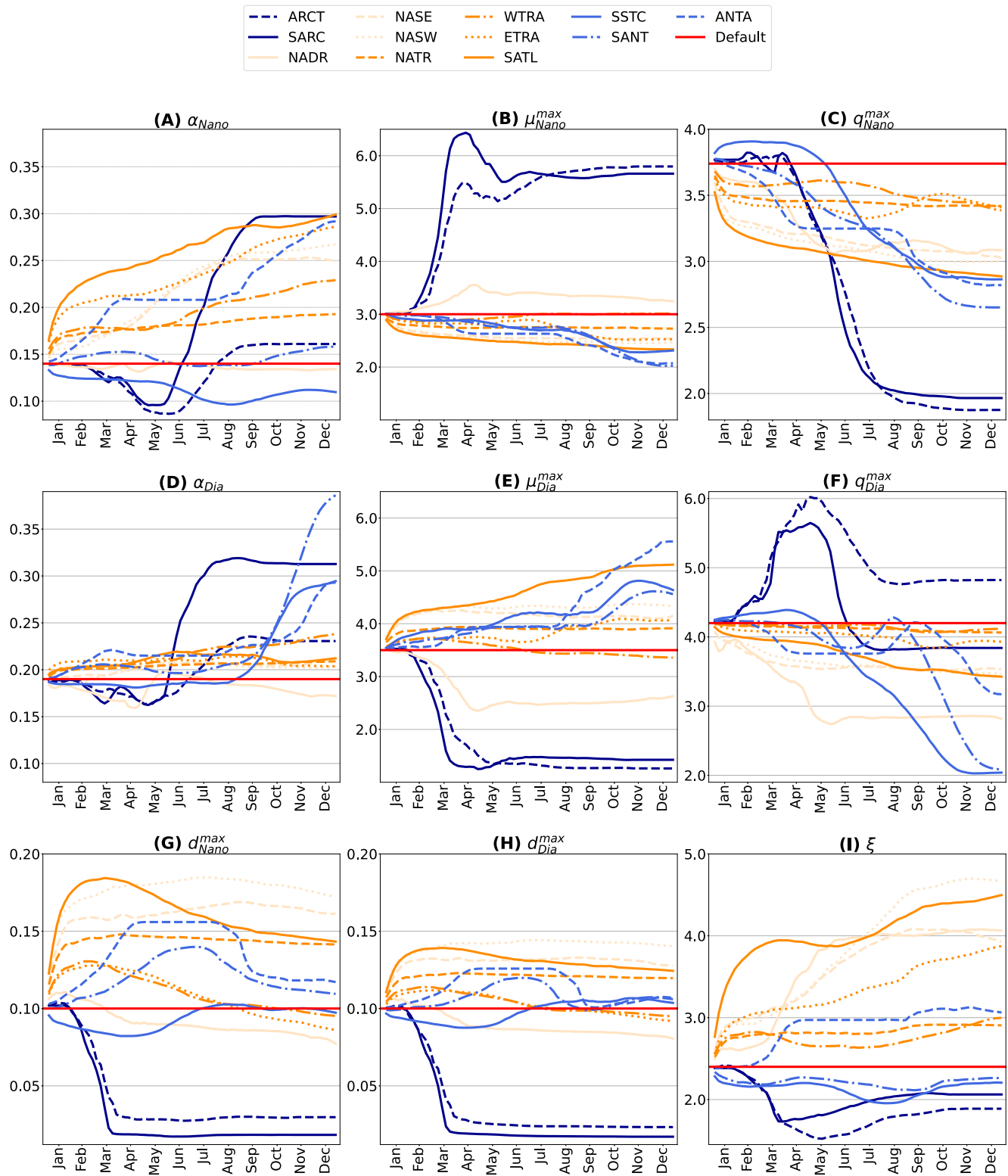


Figure 4. Temporal evaluation of average estimated parameter values across the 12 Longhurst provinces as listed in Table 3. The numbering of the subplots corresponds to that in Figure 3.

Table 3
Longhurst Province (Longhurst, 2007) Included in the Analysis

Code	Province	Code	Province
ARCT	Atlantic Arctic	WTRA	Western tropical Atlantic
SARC	Atlantic subarctic	ETRA	Eastern tropical Atlantic
NADR	North Atlantic drift	SATL	South Atlantic gyre
NASE	North Atlantic subtropical gyre east	SSTC	South subtropical convergence
NASW	North Atlantic Subtropical gyre west	SANT	Subantarctic
NATR	North Atlantic tropical gyre	ANTA	Antarctic

The opposite signs of the changes in photosynthetic parameters $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$ discussed above are also evident in their temporal evolution (Figures 4a and 4d). Large changes in these parameters occur in ARCT, SARC, and NADR provinces where the values of $\mu_{\text{Nano}}^{\text{max}}$ continuously increase and the values of $\mu_{\text{Dia}}^{\text{max}}$ decrease from the beginning of the parameter estimation process until the spring bloom. Afterward, the values of both parameters remain nearly constant. In most of the provinces, the values of these parameters converge and no clear temporal variabilities are visible.

The ratio parameters $q_{\text{Nano}}^{\text{Chl:N max}}$ and $q_{\text{Dia}}^{\text{Chl:N max}}$ exhibit a pronounced temporal variability, especially in the polar provinces (Figures 4c and 4f). In most provinces, the estimated values are lower than their default levels. The changes over time are distinct for both parameters, which leads to the distinct spatial patterns shown in Figures 4c and 4f.

The values of the degradation parameters $d_{\text{Nano}}^{\text{Chl}}$ and $d_{\text{Dia}}^{\text{Chl}}$ exhibit similar trends over time. In most provinces, both parameters increase from their default values (Figures 4g and 4h). However, in the ARCT and SARC provinces, there is a significant reduction in their values, which is nearly equal for both parameters. Provinces where $d_{\text{Nano}}^{\text{Chl}}$ and $d_{\text{Dia}}^{\text{Chl}}$ are increased also exhibit a temporal variability with particularly elevated values during the spring and summer and smaller elevations afterward.

The grazing parameter ξ undergoes substantial changes during the initial DA cycles. Increased values of ξ imply increased grazing and hence reduced concentrations of phytoplankton biomass and chlorophyll-a. Thus, to counteract the general overestimation of surface chlorophyll-a concentration in the model, the values of ξ increase in most provinces. Notable exceptions are the ARCT, SARC, SSTC, and SANT provinces in which ξ is reduced. Despite these large updates, no significant seasonal variability is visible.

3.4. Effect of Estimated Parameters on Model Skill

We verify the state accuracy of a single model forward run that use the final parameter's estimates of experiments JSPE, referred to as ADJ_FOR. We compare the performance against the reference runs.

We first compare the annual average of chlorophyll-a concentrations obtained from both simulations averaged over the year 2019 with OC-CCI data. The Taylor diagrams (K. E. Taylor, 2001) in Figure 5 show the correlations and root mean square differences (RMSD). The surface chlorophyll-a concentration simulated by the experiment ADJ_FOR outperforms the reference run (REF_FOR) with a 26% reduction of RMSD for the annual average of the surface chlorophyll-a concentrations in 2019. The correlation of ADJ_FOR with OC-CCI data is much higher with 0.73 than the correlation of 0.52 of REF_FOR.

We further compare the monthly average for April and October in 2019 also with reference to OC-CCI monthly averaged data to distinguish the spring in the Northern and Southern Hemispheres, respectively. For both months, stronger reductions of the RMSD and increases of the correlation coefficient are visible in ADJ_FOR than REF_FOR. Thus, using the estimated parameters improves the model simulation in these months. In the Southern Hemisphere, although the RMSD decreases substantially, the improvement in the correlation coefficient is smaller than the Northern Hemisphere. This indicates that while parameter estimation effectively reduces overall biases in chlorophyll-a concentration, it does not fully capture the spatial variability and dynamics in the Southern Hemisphere.

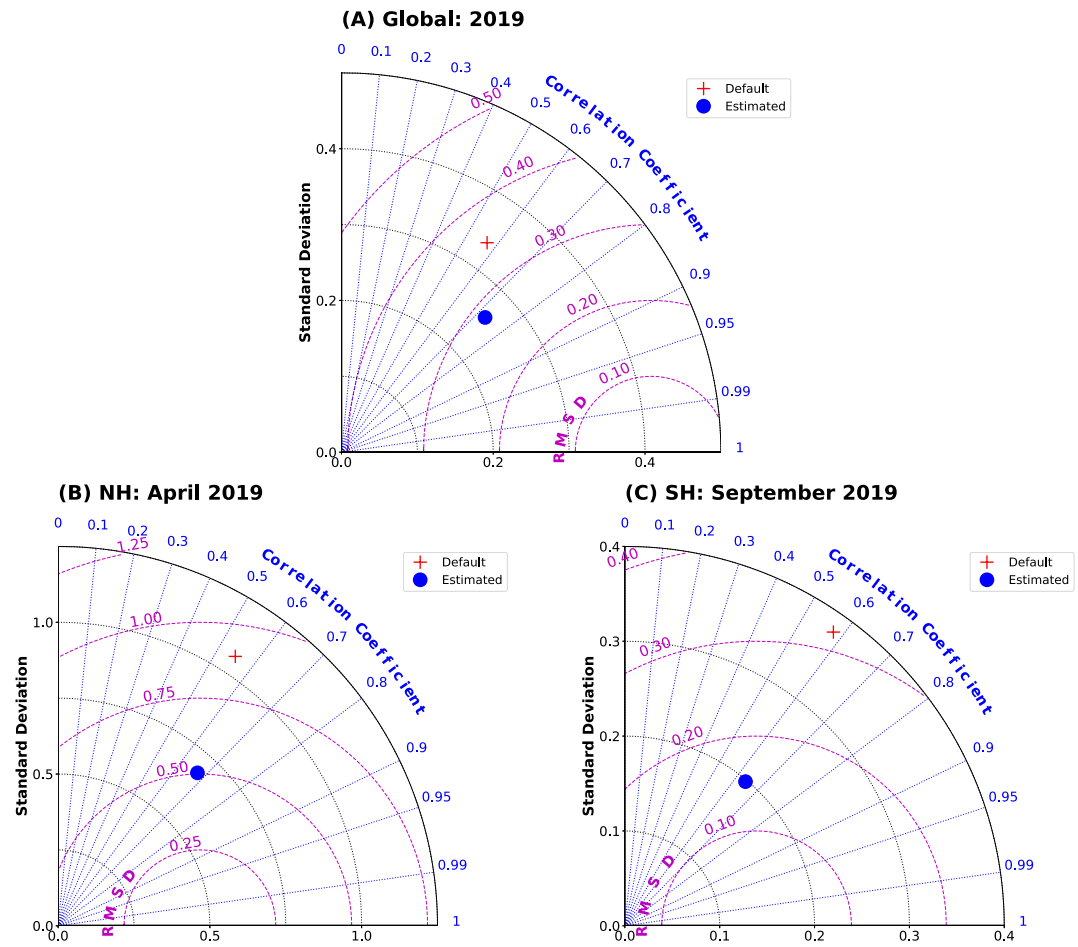


Figure 5. Taylor diagrams for the comparison of surface chlorophyll-a concentration from model simulations with default parameters and estimated parameters against satellite observations for the period 2019. The diagrams represent (a) global annual mean, (b) monthly mean for April in the Northern Hemisphere, and (c) monthly mean for September in the Southern Hemisphere.

Figure 6 illustrates that the simulation with estimated spatially varying parameters (ADJ_FOR) produces spatial patterns of surface chlorophyll-a concentrations closer to observations than to simulations using uniform default parameters (REF_FOR). In April (Figures 6d–6f), the spatial variability of the spring bloom in the Northern Hemisphere is well reproduced by the estimated parameters. Similarly, in October (Figures 6g–6i), the spatial structure of the bloom dynamics in the Southern Ocean is significantly improved; however, notable over-estimations remain in several regions.

To evaluate an independent variable, we compare here the spatial distributions of the modeled vertically integrated NPP for the runs with default and estimated parameters with estimates based on satellite data. The satellite data product is computed from the updated carbon-based productivity model (CbPM, Westberry et al., 2008). CbPM uses spectrally resolved light attenuation based on the semianalytical Garver-Siegel-Maritorena algorithm (GSM). As visible in Figure 7, the vertically integrated NPP from REF_FOR and ADJ_FOR agree reasonably with the NPP obtained using the CbPM. Nevertheless, there are regional differences. Large differences are particularly evident in coastal regions, which could be linked to model parameter estimation deficiencies or high uncertainty of satellite data-based NPP estimates in coastal water (see Westberry et al., 2008). The GSM algorithm used to estimate NPP in the CbPM product tries to distinguish the optical signatures from phytoplankton, particles, and dissolved organic matter but still requires regional tuning in coastal regions, where nonbiotic optically active material makes chlorophyll-a retrieval challenging.

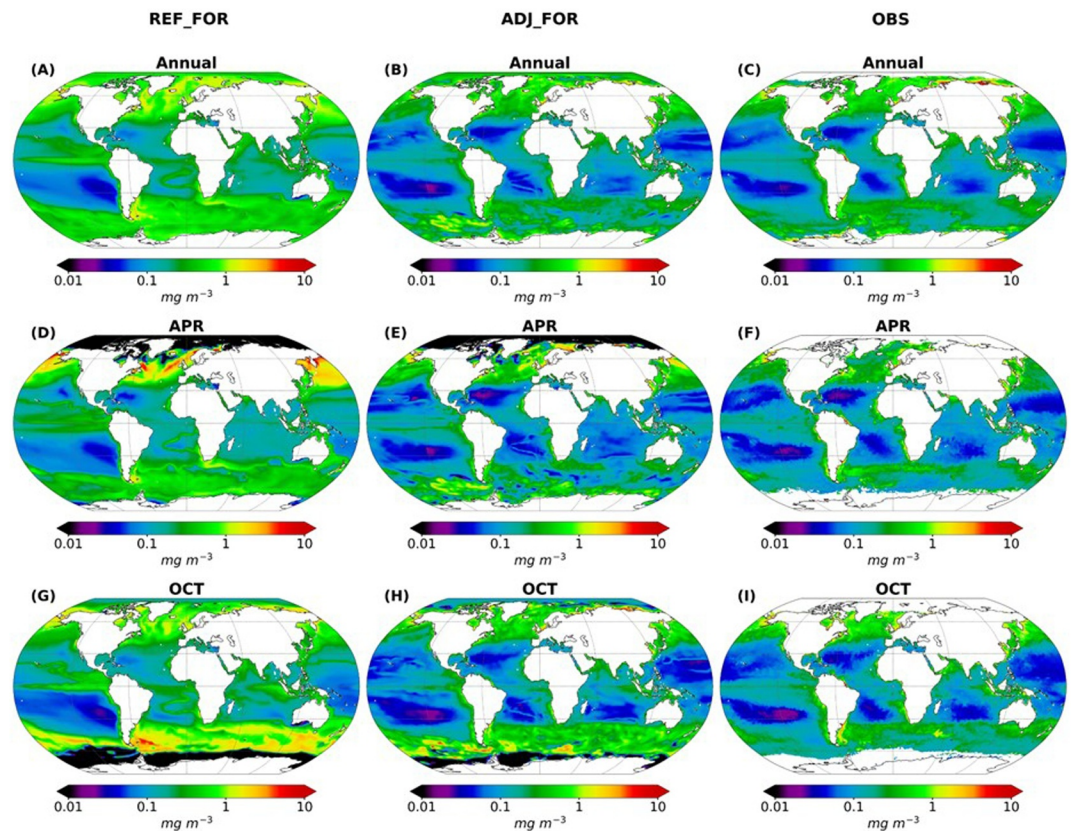


Figure 6. Surface chlorophyll-a concentration simulated with spatially varying (a, d, and g) and uniform default biogeochemical parameters (b, e, and h). Annual mean concentrations for 2019 are shown in panels a–c, with (c) representing observations, while concentrations for April 2019 (d–f) and October 2019 (g–i) highlight seasonal patterns, with F and I showing observational data.

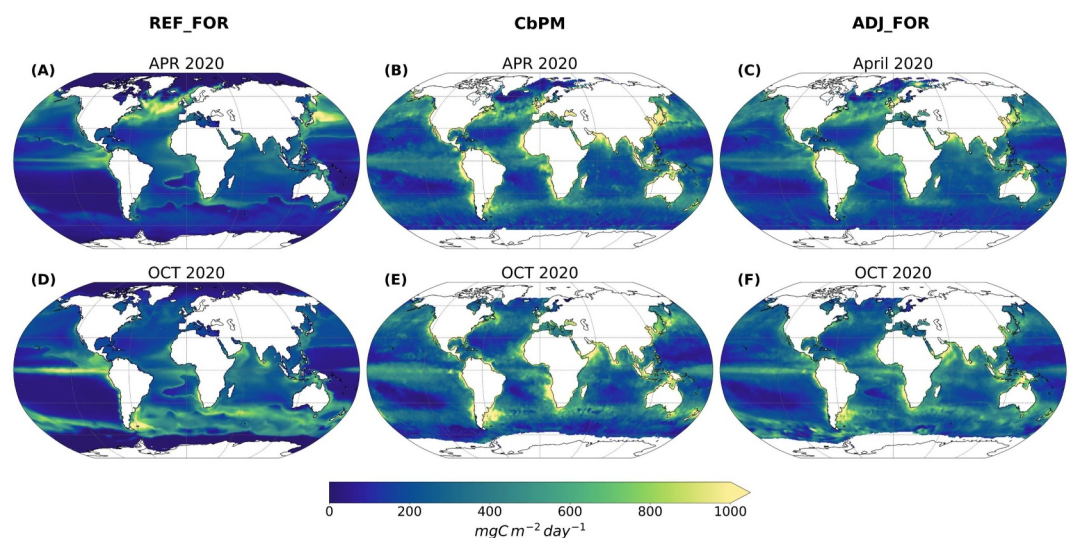


Figure 7. Monthly mean vertically integrated net primary production. The top row (a–c) is for April 2020, and the bottom row (d–f) is for October 2020. The left column (a and d) is the reference simulation of a single model with default parameter values (REF_FOR), the middle column (b and e) is for satellite-based estimation, and the right column (c and f) is from the simulation of a single model with estimated parameter values (ADJ_FOR).

4. Discussion

4.1. Subset of Parameters

In this study, we estimate the spatially varying values of nine parameters on a global scale using ocean color DA. The number of parameters considered is comparable to other studies (e.g., Gharamti et al., 2017; Kim et al., 2021; Losa et al., 2004; Mamnun et al., 2022; Singh et al., 2022), which typically ranges from 3 to 15. Utilizing DA to estimate many BGC parameters may result in low predictive skills due to overfitting to observational noise (Friedrichs et al., 2007). The efficacy of parameter estimation depends on the sufficiency of available observations to accurately constrain the chosen parameters (Thacker, 1989). Directly correlating surface chlorophyll-a with phytoplankton biomass is challenging, given the variability and often ambiguous nature of the chlorophyll-to-biomass ratio. Even with a known chlorophyll-to-biomass ratio, satellite chlorophyll-a observations predominantly constrain parameters that are directly sensitive to surface chlorophyll-a. Nevertheless, several state variables other than surface chlorophyll-a might exhibit high sensitivity to the same parameters (Mamnun et al., 2023). Consequently, relying solely on satellite chlorophyll-a data for parameter estimation may leave these variables inadequately constrained. Relying on a single observation type might be insufficient to differentiate between multiple viable parameter combinations, exemplifying the underdetermination issue frequently encountered in BGC modeling (Ward et al., 2010).

Analogous to the underdetermination problem, correlations between parameters can hinder the identification of a single optimal set of parameter values (Fiechter et al., 2011; Mamnun et al., 2022; Mattern et al., 2017). A notable manifestation of these parameter codependencies is the cancellation of uncertainties. In such cases, the model may align with available observations, not because each parameter value is optimal but because the uncertainties in correlated parameters offset each other. Therefore, individual parameter values may not be portable to other model configurations.

We reduced this underdetermination problem by selecting only nine parameters, excluding parameters that have been shown to be significantly correlated in a previous sensitivity analysis and a DA study using a 1-D setup of the model (Mamnun et al., 2023). For the chosen parameter set, it is assumed that other BGC parameters do not significantly influence the model uncertainty as was demonstrated by the sensitivity analysis by Mamnun et al. (2023). Nonetheless, the current understanding of BGC parameter uncertainties and their interrelationships is inadequate for determining an optimal parameter subset. The outcomes of our analysis are likely influenced by the specific parameters we have chosen to estimate. Thus, it is a priori unknown whether other parameter combinations yield more accurate model predictions.

4.2. Spatial Variation of Estimated Parameters

Allowing spatial variations in parameter values reduced the model-data misfit for both assimilated and independent data (Section 3.4). However, this does not inherently validate the significance of these variations in relation to the foundational BGC processes. In this section, we aim to explore the spatial patterns of the estimated parameters concerning the primary environmental factors that influence the variability in these parameters and chlorophyll-a concentration.

4.2.1. Initial Slope of the P-I Curve

The photosynthesis-irradiance parameters vary in response to various factors, from abiotic factors to community composition (H. A. Bouman et al., 2018). Nutrient availability modulates the mean irradiance in the surface mixed layer and critically influences the physiological performance of phytoplankton cells (Carvalho et al., 2020). The estimated values of the photosynthesis-irradiance parameters α_{Nano} and α_{Dia} exhibit large spatial differences varying globally by two orders of magnitude. The magnitude and spatial variability of these estimated parameters agree well with global observation-based studies (see H. A. Bouman et al., 2018).

The parameter values were generally increased in high-latitude regions (Figures 3a and 3d), which implies enhanced photosynthetic efficiency of phytoplankton in a low-light environment. The parameter α is influenced by environmental variables that exhibit significant latitudinal variations. Particularly in polar and temperate regions, upper ocean physical dynamics affect temperature and light conditions, critically shaping the photosynthetic efficiency of marine phytoplankton (Harrison & Platt, 1986; H. Bouman et al., 2005). Although large-

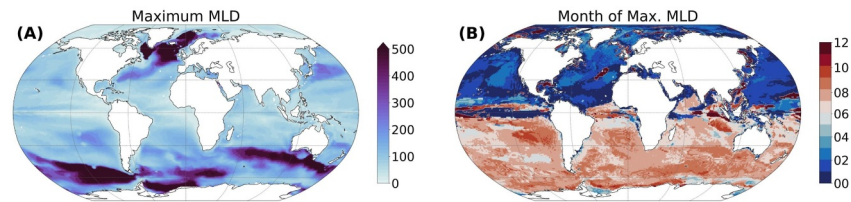


Figure 8. (a) Annual maximum mixed layer depth of Massachusetts Institute of Technology general circulation model (MITgcm) for 2020 and (b) the month of the year when maximum MLDs of MITgcm were found in 2020.

scale spatial variations in the estimated α_{Nano} and α_{Dia} values are observed, temporal differences between seasons are only observed at higher latitudes.

In the oligotrophic subtropical Pacific Ocean, the values α_{Nano} and α_{Dia} show increasing trends but never converge in the DA process. In these regions, the model consistently produces too low nanophytoplankton chlorophyll-*a* concentration, mostly likely because of a limited supply of nutrients. To compensate for the consistent negative bias, the DA raised the values of both α_{Nano} and α_{Dia} in this ocean basin. The extreme values of α_{Nano} in the subtropical region of the basin indicate compensation for the uncertainty produced by other components of the modeling system, such as, insufficient amount of iron.

The magnitude and variation of the estimated values of α_{Nano} and α_{Dia} in the oligotrophic North Atlantic are similar to observations (see H. A. Bouman et al., 2018). The spatial pattern obtained in this study is also similar to that in Losa et al. (2004). The values of α_{Nano} decrease with increasing latitude. Although the gradient for α_{Dia} is not prominent, it exhibits a similar pattern. The values around the Bermuda Atlantic Time-series Study (BATS) agree with the value obtained by Mamnun et al. (2022) and Spitz et al. (2001).

4.2.2. Maximum Photosynthesis Rate

A higher value of the maximum photosynthesis rate increases photosynthesis, thus increasing phytoplankton biomass. A higher phytoplankton growth rate will initially increase the productivity in the oligotrophic subtropical Pacific and subtropical Atlantic regions. However, the nutrients may be depleted over more extended periods (5 or more years) because there would be less supply of new nutrients below the euphotic zone. Subsequently, this condition can lead to a reduction in overall biological productivity over longer periods.

There is generally a low correlation between biomass and phytoplankton production in high latitudes (Platt et al., 1991). The relative uncoupling between the chlorophyll-*a* and the production distributions allowed the filter to make the parameters $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$ highly variable over space. $\mu_{\text{Nano}}^{\text{max}}$ varied globally by a factor of ~ 46 (range: 0.38 to 17.72 d^{-1}) similar to observations (Marañón & Holligan, 1999). However, the variation of $\mu_{\text{Dia}}^{\text{max}}$ is two orders of magnitude, which is larger than reported from observations (see H. A. Bouman et al., 2018).

Spatiotemporal differences in the photosynthesis parameters $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$ are likely driven by changes in oceanographic conditions, for example, temperature, stratification, and macronutrient and micronutrient availability (Geider et al., 1996) and by the community structure and other biological processes that may consume cellular energy at the expense of carbon fixation (Puxty et al., 2016). Cold water at high latitudes limits phytoplankton growth, resulting in lower values of the photosynthesis parameters $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$ (Smith & Donaldson, 2015). The generally low estimates of these parameters in the boreal and austral polar regions (Figures 3b and 3e) are consistent with this effect. Further, sea surface temperature (SST) can govern variations in photosynthesis parameters (Behrenfeld & Falkowski, 1997a, 1997b; Harrison & Platt, 1980; Zaiss et al., 2021). However, in this study, the warm temperatures encountered in tropical latitudes were not accompanied by consistently elevated values of $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$. Overall, we did not find a significant dependence of these parameters on SST, which justifies the absence of a clear latitudinal pattern in the EPVs in this study.

Nutrient availability is the main factor controlling the large variability of photosynthesis parameters (Marañón & Holligan, 1999). The estimated photosynthesis parameters $\mu_{\text{Nano}}^{\text{max}}$ and $\mu_{\text{Dia}}^{\text{max}}$ appear to have some dependence on the spatial variations in the annual maximum mixed layer depth (MLD), which can be seen as a proxy for nutrient flux into the upper mixed layer. To assess the relationship between nutrient availability and photosynthesis parameters, we examined the model's annual maximum MLD for 2020. Figure 8a shows the maximum MLD and

Figure 8b the month in which this maximum MLD occurred. In the Southern Ocean, a deep MLD is associated with higher μ_{Dia}^{max} , and in the North Atlantic, a deep MLD corresponds to higher μ_{Nano}^{max} .

Moreover, in the open ocean, higher temperatures are generally associated with lower nutrient availability (Sathyendranath et al., 1991), which in turn influences photosynthesis parameters. As upwelling enhances nutrient supply, an increase in μ_{Nano}^{max} was estimated in these regions. In contrast, lower values of μ_{Nano}^{max} were estimated in the southern Pacific gyre than the northern Pacific gyre, likely reflecting the stronger water column stability in the former. In the arctic and subarctic North Atlantic Ocean, deep MLD and strong convective mixing (Figure 8) lead to high surface nutrient concentrations. Consistent with this, μ_{Nano}^{max} was elevated, contributing to increased phytoplankton biomass and surface chlorophyll-a. This suggests that environmental conditions in these regions were less favorable for diatoms, leading to a negative correlation with μ_{Dia}^{max} as the filter compensates by reducing diatom productivity. Consequently, the estimated values of μ_{Nano}^{max} and μ_{Dia}^{max} exhibit an inverse pattern, similar to what was found at the BATS station by Mammun et al. (2022).

In a large part of the Southern Ocean, the modeled annual maximum MLD goes deeper than 500 m (Figure 8) before austral summer (Figure 8b). This deep MLD causes high production, leading to high chlorophyll-a concentrations in the model simulations. Consequently, the filter estimates higher values of μ_{Dia}^{max} . Since the filter only utilizes the covariances between chlorophyll-a and the parameters, rather than direct covariances between different phytoplankton groups, the reduction in μ_{Nano}^{max} does not necessarily result from the increase in μ_{Dia}^{max} . Instead, the filter lowers chlorophyll-a concentrations through a combination of adjustments, including increased grazing and reduced chlorophyll-to-nitrogen ratios. Moreover, it is known that satellite observations underestimate chlorophyll-a concentrations in the Southern Ocean (Johnson et al., 2013). In the regions of model deficits (e.g., the subarctic Atlantic), the filter estimated rather extreme values of these two parameters to achieve a reasonable agreement between observation and simulation. These values are commonly considered unrealistic (see H. A. Bouman et al., 2018), despite the exact range of realistic values not being well defined.

Spatial variation in surface irradiance may also influence the latitudinal variation in the values of μ_{Nano}^{max} and μ_{Dia}^{max} . The combination of lower surface irradiances and deep convective mixing in high latitudes results in markedly lower average light levels within the mixed layer. This can result in photoacclimation to lower light levels by modulating the pigment content per cell and hence the maximum photosynthesis rate (Cullen, 1982; Sathyendranath et al., 2009). However, the influences of the irradiance on the variability of parameters controlling the maximum photosynthesis rate are poorly understood (Marañón & Holligan, 1999). Here, we have not found any indication that irradiance influences the spatial variation of μ_{Nano}^{max} and μ_{Dia}^{max} , which also agrees with the findings of Marañón and Holligan (1999).

4.2.3. Maximum Chlorophyll-to-Nitrogen Ratio

Irradiance significantly regulates the values of the maximum chlorophyll-to-nitrogen ratio ($q^{Chl:N max}$) (A. H. Taylor et al., 1997). The dependence of photoacclimation on light is pivotal for accurately predicting the stoichiometry of phytoplankton within light gradients (Álvarez et al., 2018). The photoacclimation term in the original model by Geider et al. (1998) ties chlorophyll synthesis to the light saturation level of the photosynthetic apparatus. Specifically, when pigments absorb light in excess of what is utilized for photosynthesis, there is a downregulation in the synthesis of chlorophyll-a. According to this model, the reduction in light-harvesting complexes arises mainly from dilution, given that the rate of chlorophyll synthesis decreases compared to carbon fixation. Notably, in REcoM2, in addition to the downregulation of chlorophyll synthesis with $q_{Nano}^{Chl:N max}$ and $q_{Dia}^{Chl:N max}$, the loss of chlorophyll from functional cells is described by a chlorophyll degradation rate that is also present in Geider et al. (1998) but assumed to be very low.

One cause for variations of the ratio $q^{Chl:N max}$ in the open ocean is an imbalance between the light absorption and energy demands for photosynthesis and biosynthesis in phytoplankton cells (Geider et al., 1996). $q^{Chl:N max}$ can also change because of variations in phytoplankton photoacclimation and can depend on physiological differences across phytoplankton groups, from a lower value for smaller species to a higher value for larger diatom cells (Geider et al., 1998). In REcoM2, the carbon-specific assimilation of nitrogen is converted to chlorophyll units by multiplying by $q^{Chl:N max}$ for each phytoplankton class (Hauck et al., 2013). In our estimates, the values of $q_{Nano}^{Chl:N max}$ and $q_{Dia}^{Chl:N max}$ were generally increased to minimize the overall positive bias in the simulated

chlorophyll-a concentrations. However, the estimates of $q_{Nano}^{Chl:N\max}$ and $q_{Dia}^{Chl:N\max}$ show a smaller degree of spatial variability than other parameters considered in this study (Figures 3c and 3f) (Table 2).

In Geider et al. (1998), the ratio of chlorophyll synthesis to nitrogen assimilation is described as being highest under low irradiance, where photosynthesis is proportional to light absorption, and decreasing when photosynthesis becomes light-saturated or nutrient-limited. The maximum of this ratio under extreme low-light conditions, $q^{Chl:N\max}$, can be interpreted as a base physiological property regulating photoacclimation—specifically, the allocation of nitrogen to the light-harvesting apparatus when nutrient supply is sufficient but growth is limited by light. As discussed by MacIntyre et al. (2002), physiological traits such as maximum pigment-to-biomass ratios vary widely among taxa, influenced by factors including cell size and the energetic or material costs of pigment synthesis. We therefore interpret $q^{Chl:N\max}$ as a taxa-dependent photoacclimation parameter that can reasonably exhibit spatial variability reflecting differences in community composition.

4.2.4. Chlorophyll Degradation Rates

Constraining the chlorophyll degradation rates (d_{Nano}^{Chl} and d_{Dia}^{Chl}) presents a challenge in quota-based BGC models. The original model by Geider et al. (1998) described photoacclimation over daylong time scales, but while it included a term accounting for chlorophyll degradation, it assumed that term to be very small, based on lab experiments with diatoms. The inclusion of the chlorophyll degradation rate becomes particularly significant during low-growth periods in winter and at the lower reaches of the euphotic zone (Sasai et al., 2022). Without a term for chlorophyll loss, which represents complex processes in aging of photo-stressed cells, the carbon to chlorophyll ratios (C:Chl) in phytoplankton can become skewed.

The chlorophyll degradation parameter is often subjectively adjusted until a satisfactory alignment between observational data and model simulations is achieved. However, this approach might not be universally applicable across different BGC models. In most regions, the parameter values change by a factor of up to 2. In the Arctic and subarctic Atlantic, both parameters are reduced to values close to zero. This avoids loss of chlorophyll during the winter months as described by Joli et al. (2024). Replacing the rudimentary chlorophyll degradation model with a more detailed process-based representation of photosystem functionality degradation can likely refine the modeled C:Chl ratios (Álvarez et al., 2018).

4.2.5. Maximum Grazing Rate of Zooplankton

The estimated maximum grazing rate of zooplankton (ξ) exhibits spatial variability by a factor of 170. Despite the diversity of zooplankton, ranging from unicellular flagellates to multicellular organisms, the REcoM2 version used here represents a single generic zooplankton group, a common approach in many ocean BGC models. However, zooplankton grazing is one of the largest sources of uncertainty in ocean BGC models used for climate projections (Laufkötter et al., 2015; Rohr et al., 2023). Introducing multiple zooplankton types could help reduce biases in surface chlorophyll-a concentrations (Karakuş et al., 2022). In our estimation, we only included the maximum grazing rate but did not consider the grazing efficiency (γ), which is also sensitive (Mamnun et al., 2023). This choice was motivated from the fact that ξ and γ are highly correlated (Mamnun et al., 2022) and hence cannot be estimated independently. However, with only allowing ξ to vary, its variability most likely compensates for other grazing-related parameters.

The parameter estimation reduced the values of ξ in the Arctic and subarctic Atlantic Ocean (Figure 3i). During a spring bloom, the biomass of zooplankton trails behind the growth of phytoplankton due to the effects of temperature on zooplankton development (Daase et al., 2013; Søreide et al., 2010). As a result, there can be situations where phytoplankton and zooplankton exhibit either a negative correlation (e.g., when phytoplankton increases while zooplankton biomass is low or vice versa) or a positive correlation (when both populations are increasing). These dynamics were evident in our study. In REcoM2, surface ocean phytoplankton biomass typically diminishes due to aggregated sinking or grazing, as the model does not include physiological mortality. During the spring bloom, the growth interplay between zooplankton and phytoplankton was more synchronized, mainly as much of the frontal structure had dissipated by then (Dong et al., 2021). Given the spatiotemporal fluctuations observed throughout the study, comprehending the intricate relationship between phytoplankton and zooplankton remains challenging.

The reduced value of ξ in the oligotrophic subtropical South Pacific Ocean (Figure 3i) substantially increases the phytoplankton production in the simulations. This compensates that the model underestimates chlorophyll-a (Figure 2) in this region when the DPV is used. Saito et al. (2005) also found that grazing parameters in this region are highly sensitive. A different effect is found in regions with strong upwelling and convective mixing, such as the Southern Ocean and the tropical North Pacific. Here, increasing the grazing rate substantially reduces surface chlorophyll-a concentrations, which minimizes the misfit with observations. The sensitivity studies with 1-D model configurations in the tropical North Pacific Ocean (Chai et al., 2002; Dugdale et al., 2002) also agree. The parameter estimation also increased the value of ξ around the BATS site, which agrees with the study with a 1-D model configuration of REcoM2 (Mamnun et al., 2022) and other previous studies (Doron et al., 2013; Losa et al., 2004). In situ measurements (Evelyn & Michael, 1998) also show that the grazer community consumes most of the production at this location. In addition, increases of ξ in the high-nutrient, low-chlorophyll regions of the Southern Ocean led to suppressed phytoplankton mass on the surface and compensate for overestimated surface chlorophyll-a by the model.

4.3. Parameter Compensation for Other Model Deficiencies

Our parameter estimation focused exclusively on uncertainties arising from BGC parameters. By treating only nine biological parameters as stochastic and updating them through the assimilation of chlorophyll observations, we did not account for other sources of model error. These include additional biological parameters, errors in the underlying physical model, physical forcing, boundary and initial conditions, and even the functional form of model equations. Moreover, model discrepancies may also stem from coupled system components, such as sea ice dynamics, atmospheric forcing, or inherent structural inadequacies in the model framework. While the model agrees reasonably well with existing observations, this agreement may result from BGC parameter adjustments compensating for limitations in other model components or structural uncertainties. Such compensatory errors pose significant challenges for future climate projections. While they may yield plausible simulations for the present state, they undermine confidence in the model's predictive capability for future scenarios (Löptien & Dietze, 2019). This limitation also implies that the parameter estimates are model-specific.

A particular REcoM2-specific effect is that the free run consistently produces a thin line of elevated chlorophyll-a in the southern subtropical gyres, most visible in the subtropical South Atlantic and Indian Ocean (Figure 2). This feature is an artifact of the formulation of growth limitation in REcoM2. Directly at the transition from iron to nitrogen limitation, neither of the two limitation terms is as low as in the centers of the limited regions. In the REcoM2 code, the colimitation is calculated as a minimum of two limitation terms, so there is less limitation directly at the transition than in the equilibrium regions. The DA reduced this line of elevated chlorophyll-a. Although the values of chlorophyll-a in these features are smaller and probably do not affect global BGC cycles much, the parameter estimation process responded to this model deficiency. This is particularly visible for the maximum chlorophyll to nitrogen ratio of nanophytoplankton and the maximum grazing rate in Figures 3c and 3i. These parameters had larger changes in the transition regions to compensate for the elevated chlorophyll-a. These parameter estimates resulted in a reduced line of elevated chlorophyll in the adjusted rerun (ADJ_FOR). However, the parameter estimations are obviously specific for this colimitation effect and do not fully resolve it.

Whenever uncertainties from different sources compensate for each other, the ocean BGC model may yield reasonable outputs for the period for which the parameter estimation was conducted. However, this compensatory behavior compromises the model's utility for gaining mechanistic insights and for being a reliable predictive instrument beyond that specific timeframe. This phenomenon is particularly concerning when considering slow climate BGC feedback mechanisms, which are inherently challenging to probe with current observations.

In practice, when uncertainties beyond just model parameter uncertainties influence the parameter estimations, the outcomes of parameter estimates become indeterminate in their effectiveness. While compensating for these uncertainties can mitigate some deficits, it does not address all of them. This phenomenon was evident in Simon et al. (2015), where parameter estimation was conducted on a regional BGC model for the North Atlantic and Arctic Oceans. Furthermore, parameter estimates optimized for a specific location or regional scale can deteriorate model outcomes for other regions, as highlighted by Friedrichs et al. (2007). This underscores the need for estimating spatially varying BGC parameters.

5. Summary and Conclusion

We employed ensemble DA to estimate spatiotemporally varying values of nine uncertain process parameters within a 3-D global ocean BGC model by assimilating satellite surface chlorophyll-a concentrations. The EPVs are (a) the initial slope of the P-I curve for nanophytoplankton and (b) diatoms, (c) the maximum photosynthesis rate of nanophytoplankton and (d) diatoms, (e) the maximum chlorophyll-to-nitrogen ratio for nanophytoplankton and (f) diatoms, (g) the chlorophyll degradation rate of nanophytoplankton and (h) diatoms, and (i) the maximum grazing rate of zooplankton. Applying an EnKF, we adopt the augmented state vector approach, which allows us to use multivariate correlations between BGC parameters and observed state variables, that is, surface chlorophyll-a concentration to estimate the parameter values. In addition to the chosen nine parameters, the DA updated eight model state variables, namely the biomass content of carbon, nitrogen, and calcium carbonate of nanophytoplankton; biomass content of carbon, nitrogen, and silicate of diatoms; and chlorophyll-a concentration in nanophytoplankton and diatoms, which were used as the initial conditions for the next forecast cycle.

The resulting parameter estimates span a range above and below the default values, underscoring the efficacy of DA in enhancing the model's regional parameterization. Notably, simulations with the optimized spatially varying parameters align more closely with observations than those using uniform defaults, with a 26% reduction in RMSD for annual surface chlorophyll-a concentrations and a higher correlation (0.73 vs. 0.52) with OC-CCI data. Seasonal improvements are also evident capturing spring bloom dynamics more accurately. Comparisons with satellite-based NPP estimates further confirm improved agreement, though regional discrepancies remain, especially in coastal areas where parameter estimation uncertainties or retrieval challenges may play a role. In some regions, the EPVs change markedly over shorter spatial scales. Such sharp gradients in parameter values over short distances are unlikely to be ecologically plausible. Future work could therefore explore methods for spatial regularization, for example, by constraining parameters to vary smoothly within regions of relatively uniform BGC properties.

A notable, albeit anticipated, consequence of augmented state parameter estimation is the filter's unpredictable use of the additional degrees of freedom. Although our primary objective is to achieve the best parameter estimates by minimizing their uncertainty, the DA process may inadvertently offset the uncertainty from other sources. This can result in suboptimal parameter estimates yet improved state estimates. Such parameter estimates, resulting from the compensation of other uncertainties, restrict the portability of these estimates to a different model configuration or a different model.

While our results demonstrate the efficacy of the estimation method assimilating ocean color data, possible extensions and improvements of the method are numerous, and further developments of DA for different types of observations should be explored. The methodology evaluated in this study is not exclusive to the MITgcm-REcoM2 model and does not necessitate extensive inverse model developments. Thus, it can be adapted to other models, provided that the ensemble simulations, which describe the model's response to parameter uncertainty, are computationally viable. Our approach offers a valuable alternative for consideration in a research environment with multiple coexisting models and limited computational resources for expanding variable counts and calibrating parameters. An additional avenue for future work is to investigate the relative impact of allowing parameters to vary spatially alone versus both spatially and temporally, as incorporating temporal variability may further refine model calibration and skill.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The forward and data assimilative model simulation data used in this study are archived in Mamnun (2025a), and the corresponding Jupyter Notebooks that produced the figures are available in Mamnun (2025b).

Access to the European Space Agency's Ocean Colour-Climate Change Initiative (OC-CCI; Sathyendranath et al., 2019) data sets is available via <https://www.oceancolour.org/>. Satellite data-based updated carbon-based productivity model (CbPM, Westberry et al., 2008) net primary production data are available <https://orca.science.oregonstate.edu/>.

Acknowledgments

This study was supported by the Helmholtz Initiative and Networking Fund pilot project Uncertainty Quantification—From Data to Reliable Knowledge (Helmholtz-UQ). Open Access funding enabled and organized by Projekt DEAL.

References

- Albani, S., Mahowald, N. M., Perry, A. T., Scanza, R. A., Zender, C. S., Heavens, N. G., et al. (2014). Improved dust representation in the community atmosphere model. *Journal of Advances in Modeling Earth Systems*, 6(3), 541–570. <https://doi.org/10.1002/2013ms000279>
- Álvarez, E., Thoms, S., & Völker, C. (2018). Chlorophyll to carbon ratio derived from a global ecosystem model with photodamage. *Global Biogeochemical Cycles*, 32(5), 799–816. <https://doi.org/10.1029/2017gb005850>
- Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, 129(12), 2884–2903. [https://doi.org/10.1175/1520-0493\(2001\)129<2884:AEAKFF>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2)
- Anderson, T. R. (2005). Plankton functional type modelling: Running before we can walk? *Journal of Plankton Research*, 27(11), 1073–1081. <https://doi.org/10.1093/plankt/fbi076>
- Andrews, O., Buitenhuis, E., Le Quere, C., & Suntharalingam, P. (2017). Biogeochemical modelling of dissolved oxygen in a changing ocean. *Philosophical Transactions of the Royal Society A*, 375(2102), 20160328. <https://doi.org/10.1098/rsta.2016.0328>
- Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, 8(8), 2465–2513. <https://doi.org/10.5194/gmd-8-2465-2015>
- Behrenfeld, M. J., & Falkowski, P. G. (1997a). A consumer's guide to phytoplankton primary productivity models. *Limnology & Oceanography*, 42(7), 1479–1491. <https://doi.org/10.4319/lo.1997.42.7.1479>
- Behrenfeld, M. J., & Falkowski, P. G. (1997b). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnology & Oceanography*, 42(1), 1–20. <https://doi.org/10.4319/lo.1997.42.1.0001>
- Behrenfeld, M. J., & Milligan, A. J. (2013). Photophysiological expressions of iron stress in phytoplankton. *Annual Review of Marine Science*, 5(1), 217–246. <https://doi.org/10.1146/annurev-marine-121211-172356>
- Bopp, L., Resplandy, L., Untersee, A., Le Mezo, P., & Kageyama, M. (2017). Ocean (de)oxygenation from the last glacial maximum to the twenty-first century: Insights from Earth system models. *Philosophical Transactions*, 375(2102), 20160323. <https://doi.org/10.1098/rsta.2016.0323>
- Bouman, H., Platt, T., Sathyendranath, S., & Stuart, V. (2005). Dependence of light-saturated photosynthesis on temperature and community structure. *Deep Sea Research Part I: Oceanographic Research Papers*, 52(7), 1284–1299. <https://doi.org/10.1016/j.dsr.2005.01.008>
- Bouman, H. A., Platt, T., Doblin, M., Figueiras, F. G., Gudmundsson, K., Gudfinnsson, H. G., et al. (2018). Photosynthesis-irradiance parameters of marine phytoplankton: Synthesis of a global data set. *Earth System Science Data*, 10(1), 251–266. <https://doi.org/10.5194/essd-10-251-2018>
- Boyer, T. P., Garcia, H. E., Locarnini, R. A., Zweng, M. M., Mishonov, A. V., Reagan, J. R., et al. (2018). World ocean atlas 2018 [Dataset]. NOAA National Centers for Environmental Information. Retrieved from <https://www.ncei.noaa.gov/products/world-ocean-atlas>
- Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *Journal of Geophysical Research*, 100(C7), 13237–13254. <https://doi.org/10.1029/95jc00458>
- Carroll, D., Menemenlis, D., Adkins, J. F., Bowman, K. W., Brix, H., Dutkiewicz, S., et al. (2020). The ECCO-darwin data-assimilative global ocean biogeochemistry model: Estimates of seasonal to multidecadal surface ocean pCO₂ and air-sea CO₂ flux. *Journal of Advances in Modeling Earth Systems*, 12(10), e2019MS001888. <https://doi.org/10.1029/2019ms001888>
- Carvalho, F., Fitzsimmons, J. N., Couto, N., Waite, N., Gorbunov, M., Kohut, J., et al. (2020). Testing the Canyon Hypothesis: Evaluating light and nutrient controls of phytoplankton growth in penguin foraging hotspots along the West Antarctic Peninsula. *Limnology and Oceanography*, 65(3), 455–470. <https://doi.org/10.1002/lno.11313>
- Chai, F., Dugdale, R. C., Peng, T. H., Wilkerson, F. P., & Barber, R. T. (2002). One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part I: Model development and silicon and nitrogen cycle. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(13), 2713–2745. [https://doi.org/10.1016/S0967-0645\(02\)00055-3](https://doi.org/10.1016/S0967-0645(02)00055-3)
- Cheung, W. W. L., Lam, V. W. Y., Sarmiento, J. L., Kearney, K., Watson, R. E. G., Zeller, D., & Pauly, D. (2010). Large-scale redistribution of maximum fisheries catch potential in the global ocean under climate change. *Global Change Biology*, 16(1), 24–35. <https://doi.org/10.1111/j.1365-2486.2009.01995.x>
- Ciavatta, S., Kay, S., Saux-Picart, S., Butenschon, M., & Allen, J. I. (2016). Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *Journal of Geophysical Research-Oceans*, 121(3), 1824–1845. <https://doi.org/10.1002/2015jc011496>
- Crisp, D., Dolman, H., Tanhua, T., McKinley, G. A., Hauck, J., Bastos, A., et al. (2022). How well do we understand the land-ocean-atmosphere carbon cycle? *Reviews of Geophysics*, 60(2), e2021RG000736. <https://doi.org/10.1029/2021rg000736>
- Cullen, J. J. (1982). The deep chlorophyll maximum: Comparing vertical profiles of chlorophyll a. *Canadian Journal of Fisheries and Aquatic Sciences*, 39(5), 791–803. <https://doi.org/10.1139/f82-108>
- Daase, M., Falk-Petersen, S., Varpe, O., Darnis, G., Søreide, J. E., Wold, A., et al. (2013). Timing of reproductive events in the marine copepod *Calanus glacialis*: A pan-Arctic perspective. *Canadian Journal of Fisheries and Aquatic Sciences*, 70(6), 871–884. <https://doi.org/10.1139/cjfas-2012-0401>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Denman, K. L. (2003). Modelling planktonic ecosystems: Parameterizing complexity. *Progress in Oceanography*, 57(3–4), 429–452. [https://doi.org/10.1016/S0079-6611\(03\)00109-5](https://doi.org/10.1016/S0079-6611(03)00109-5)
- Dong, H., Zhou, M., Hu, Z., Zhang, Z., Zhong, Y., Basedow, S. L., & Smith, W. O. (2021). Transport barriers and the retention of *Calanus finmarchicus* on the Lofoten shelf in early spring. *Journal of Geophysical Research: Oceans*, 126(8), e2021JC017408. <https://doi.org/10.1029/2021jc017408>
- Doron, M., Brasseur, P., Brankart, J. M., Losa, S. N., & Melet, A. (2013). Stochastic estimation of biogeochemical parameters from global colour ocean colour satellite data in a North Atlantic 3D ocean coupled physical-biogeochemical model. *Journal of Marine Systems*, 117, 81–95. <https://doi.org/10.1016/j.jmarsys.2013.02.007>
- Droop, M. R. (1983). 25 years of algal growth-kinetics - A personal view. *Botanica Marina*, 26(3), 99–112. <https://doi.org/10.1515/botm.1983.26.3.99>
- Dugdale, R. C., Barber, R. T., Chai, F., Peng, T. H., & Wilkerson, F. P. (2002). One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part II: Sensitivity analysis and comparison with JGOFS EqPac data. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(13), 2747–2768. [https://doi.org/10.1016/S0967-0645\(02\)00056-5](https://doi.org/10.1016/S0967-0645(02)00056-5)
- Evelyn, J. L., & Michael, C. M. (1998). Microzooplankton herbivory and phytoplankton growth in the northwestern Sargasso Sea. *Aquatic Microbial Ecology*, 16(2), 173–188. <https://doi.org/10.3354/ame016173>

- Fennel, K., Mattern, J. P., Doney, S. C., Bopp, L., Moore, A. M., Wang, B., & Yu, L. (2022). Ocean biogeochemical modelling. *Nature Reviews Methods Primers*, 2(1), 76. <https://doi.org/10.1038/s43586-022-00154-2>
- Fiechter, J., Broquet, G., Moore, A. M., & Arango, H. G. (2011). A data assimilative, coupled physical-biological model for the coastal Gulf of Alaska. *Dynamics of Atmospheres and Oceans*, 52(1–2), 95–118. <https://doi.org/10.1016/j.dynatmoce.2011.01.002>
- Follows, M. J., Dutkiewicz, S., Grant, S., & Chisholm, S. W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science*, 315(5820), 1843–1846. <https://doi.org/10.1126/science.1138544>
- Ford, D., & Barciela, R. (2017). Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products. *Remote Sensing of Environment*, 203, 40–54. <https://doi.org/10.1016/j.rse.2017.03.040>
- Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, 8(10), 3071–3104. <https://doi.org/10.5194/gmd-8-3071-2015>
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., et al. (2022). Global carbon budget 2022. *Earth System Science Data*, 14(11), 4811–4900. <https://doi.org/10.5194/essd-14-4811-2022>
- 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(C8). <https://doi.org/10.1029/2006jc003852>
- García, H. E., Weathers, K., Paver, C. R., Smolyar, I., Boyer, T. P., Locarnini, R. A., et al. (2019). World ocean atlas 2018, volume 4: Dissolved inorganic nutrients (phosphate, nitrate and nitrite + nitrate, silicate). *NOAA Atlas NESDIS 84*.
- Gattuso, J.-P., Magnan, A. K., Bopp, L., Cheung, W. W. L., Duarte, C. M., Hinkel, J., et al. (2018). Ocean solutions to address climate change and its effects on marine ecosystems. *Frontiers in Marine Science*, 5, 337. <https://doi.org/10.3389/fmars.2018.00337>
- Gehlen, M., Gangstø, R., Schneider, B., Bopp, L., Aumont, O., & Etche, C. (2007). The fate of pelagic CaCO₃ production in a high CO₂ ocean: A model study. *Biogeosciences*, 4(4), 505–519. <https://doi.org/10.5194/bg-4-505-2007>
- Geider, R. J., & La Roche, J. (1994). The role of iron in phytoplankton photosynthesis, and the potential for iron-limitation of primary productivity in the sea. *Photosynthesis Research*, 39(3), 275–301. <https://doi.org/10.1007/BF00014588>
- Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1996). A dynamic model of photoadaptation in phytoplankton. *Limnology & Oceanography*, 41(1), 1–15. <https://doi.org/10.4319/lo.1996.41.1.0001>
- Geider, R. J., MacIntyre, H. L., & Kana, T. M. (1998). A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnology & Oceanography*, 43(4), 679–694. <https://doi.org/10.4319/lo.1998.43.4.0679>
- Gharamti, M. E., Samuelsen, A., Bertino, L., Simon, E., Korosov, A., & Daewel, U. (2017). Online tuning of ocean biogeochemical model parameters using ensemble estimation techniques: Application to a one-dimensional model in the North Atlantic. *Journal of Marine Systems*, 168, 1–16. <https://doi.org/10.1016/j.jmarsys.2016.12.003>
- Goodliff, M., Bruening, T., Schwichtenberg, F., Li, X., Lindenthal, A., Lorkowski, I., & Nerger, L. (2019). Temperature assimilation into a coastal ocean-biogeochemical model: Assessment of weakly and strongly coupled data assimilation. *Ocean Dynamics*, 69(10), 1217–1237. <https://doi.org/10.1007/s10236-019-01299-7>
- Gregg, W. W. (2008). Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model. *Journal of Marine Systems*, 69(3–4), 205–225. <https://doi.org/10.1016/j.jmarsys.2006.02.015>
- Gruber, N., Boyd, P. W., Frolicher, T. L., & Vogt, M. (2021). Biogeochemical extremes and compound events in the ocean. *Nature*, 600(7889), 395–407. <https://doi.org/10.1038/s41586-021-03981-7>
- Gutknecht, E., Reffray, G., Mignot, A., Dabrowski, T., & Sotillo, M. G. (2019). Modelling the marine ecosystem of iberia-biscay-ireland (IBI) European waters for CMEMS operational applications. *Ocean Science*, 15(6), 1489–1516. <https://doi.org/10.5194/os-15-1489-2019>
- Harrison, W. G., & Platt, T. (1980). Variations in assimilation number of coastal marine phyto-phytoplankton: Effects of environmental co-variables. *Journal of Plankton Research*, 2(4), 249–260. <https://doi.org/10.1093/plankt/2.4.249>
- Harrison, W. G., & Platt, T. (1986). Photosynthesis-irradiance relationships in polar and temperate phytoplankton populations. *Polar Biology*, 5(3), 153–164. <https://doi.org/10.1007/BF00441695>
- Hauck, J., Volker, C., Wang, T., Hoppema, M., Losch, M., & Wolf-Gladrow, D. A. (2013). Seasonally different carbon flux changes in the Southern Ocean in response to the southern annular mode. *Global Biogeochemical Cycles*, 27(4), 1236–1245. <https://doi.org/10.1002/2013GB004600>
- Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., et al. (2020). Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Frontiers in Marine Science*, 7, 571720. <https://doi.org/10.3389/fmars.2020.571720>
- Hauri, C., Pagès, R., McDonnell, A. M. P., Stuecker, M. F., Danielson, S. L., Hedstrom, K., et al. (2021). Modulation of ocean acidification by decadal climate variability in the Gulf of Alaska. *Communications Earth & Environment*, 2(1), 191. <https://doi.org/10.1038/s43247-021-00254-z>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hohn, S. (2009). Coupling and decoupling of biogeochemical cycles in marine ecosystems (Thesis). *Fachbereich 2 Biologie*.
- Ilyina, T., Li, H., Spring, A., Müller, W. A., Bopp, L., Chikamoto, M. O., et al. (2021). Predictable variations of the carbon sinks and atmospheric CO₂ growth in a multi-model framework. *Geophysical Research Letters*, 48(6), e2020GL090695. <https://doi.org/10.1029/2020gl090695>
- Ilyina, T., Zeebe, R. E., Maier-Reimer, E., & Heinze, C. (2009). Early detection of ocean acidification effects on marine calcification. *Global Biogeochemical Cycles*, 23(1). <https://doi.org/10.1029/2008GB003278>
- Johnson, R., Strutton, P. G., Wright, S. W., McMinn, A., & Meiners, K. M. (2013). Three improved satellite chlorophyll algorithms for the southern ocean. *Journal of Geophysical Research-Oceans*, 118(7), 3694–3703. <https://doi.org/10.1002/jgrc.20270>
- Joli, N., Concia, L., Mocaer, K., Guterman, J., Laude, J., Guerin, S., et al. (2024). Hypometabolism to survive the long polar night and subsequent successful return to light in the diatom *fragilariopsis cylindrus*. *New Phytologist*, 241(5), 2193–2208. <https://doi.org/10.1111/nph.19387>
- Jones, E. M., Baird, M. E., Mongin, M., Parslow, J., Skerratt, J., Lovell, J., et al. (2016). Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the Great Barrier Reef. *Biogeosciences*, 13(23), 6441–6469. <https://doi.org/10.5194/bg-13-6441-2016>
- Karakaş, O., Völker, C., Iversen, M., Hagen, W., & Hauck, J. (2022). The role of zooplankton grazing and nutrient recycling for global ocean biogeochemistry and phytoplankton phenology. *Journal of Geophysical Research: Biogeosciences*, 127(10), e2022JG006798. <https://doi.org/10.1029/2022jg006798>
- Kim, H. H., Luo, Y. W., Ducklow, H. W., Schofield, O. M., Steinberg, D. K., & Doney, S. C. (2021). WAP-ID-VAR v1.0: Development and evaluation of a one-dimensional variational data assimilation model for the marine ecosystem along the West Antarctic Peninsula. *Geoscientific Model Development*, 14(8), 4939–4975. <https://doi.org/10.5194/gmd-14-4939-2021>

- Krumhardt, K. M., Lovenduski, N. S., Long, M. C., Levy, M., Lindsay, K., Moore, J. K., & Nissen, C. (2019). Coccolithophore growth and calcification in an acidified Ocean: Insights from community Earth system model simulations. *Journal of Advances in Modeling Earth Systems*, 11(5), 1418–1437. <https://doi.org/10.1029/2018MS001483>
- Lam, V. W. Y., Cheung, W. W. L., Reygondeau, G., & Sumaila, U. R. (2016). Projected change in global fisheries revenues under climate change. *Scientific Reports*, 6(1), 32607. <https://doi.org/10.1038/srep32607>
- Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences*, 12(23), 6955–6984. <https://doi.org/10.5194/bg-12-6955-2015>
- Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X. H., et al. (2016). A new global interior ocean mapped climatology: The 1 degrees x 1 degrees GLODAP version 2. *Earth System Science Data*, 8(2), 325–340. <https://doi.org/10.5194/essd-8-325-2016>
- Leles, S. G., Polimene, L., Bruggeman, J., Blackford, J., Ciavatta, S., Mitra, A., & Flynn, K. J. (2018). Modelling mixotrophic functional diversity and implications for ecosystem function. *Journal of Plankton Research*, 40(6), 627–642. <https://doi.org/10.1093/plankt/fby044>
- Longhurst, A. R. (2007). Chapter 7 - Provinces: The secondary compartments. In A. R. Longhurst (Ed.), *Ecological geography of the sea* (2nd ed., pp. 103–114). Academic Press. <https://doi.org/10.1016/B978-0-12455521-1/50008-5>
- Löptien, U., & Dietze, H. (2019). Reciprocal bias compensation and ensuing uncertainties in model-based climate projections: Pelagic biogeochemistry versus ocean mixing. *Biogeosciences*, 16(9), 1865–1881. <https://doi.org/10.5194/bg-16-1865-2019>
- Losa, S. N., Kivman, G. A., & Ryabchenko, V. A. (2004). Weak constraint parameter estimation for a simple ocean ecosystem model: What can we learn about the model and data? *Journal of Marine Systems*, 45(1), 1–20. <https://doi.org/10.1016/j.jmarsys.2003.08.005>
- Losch, M., Menemenlis, D., Campin, J. M., Heimbach, P., & Hill, C. (2010). On the formulation of sea-ice models. Part 1: Effects of different solver implementations and parameterizations. *Ocean Modelling*, 33(1–2), 129–144. <https://doi.org/10.1016/j.ocemod.2009.12.008>
- Loukos, H., Monfray, P., Bopp, L., & Lehodey, P. (2003). Potential changes in skipjack tuna (*Katsuwonus pelamis*) habitat from a global warming scenario: Modelling approach and preliminary results. *Fisheries Oceanography*, 12(4–5), 474–482. <https://doi.org/10.1046/j.1365-2419.2003.00241.x>
- MacIntyre, H. L., Kana, T. M., Anning, T., & Geider, R. J. (2002). Photoacclimation of photosynthesis irradiance response curves and photo-synthetic pigments in microalgae and cyanobacteria. *Journal of Phycology*, 38(1), 17–38. <https://doi.org/10.1046/j.1529-8817.2002.00094.x>
- Mamnun, N. (2025a). Global ocean model outputs: Data assimilative and forward modeling for biogeochemical parameter estimation [Dataset]. *Zenodo*. <https://doi.org/10.5281/zenodo.15209270>
- Mamnun, N. (2025b). Jupyter notebooks to generate figures for parameter estimation in global ocean biogeochemistry (v1.1). *Zenodo*. <https://doi.org/10.5281/zenodo.15209843>
- Mamnun, N., Völker, C., Krumscheid, S., Vrekoussis, M., & Nerger, L. (2023). Global sensitivity analysis of a one-dimensional ocean biogeochemical model. *Socio Environmental Systems Modelling*, 5, 18613. <https://doi.org/10.18174/sesmo.18613>
- Mamnun, N., Völker, C., Vrekoussis, M., & Nerger, L. (2022). Uncertainties in ocean biogeochemical simulations: Application of ensemble data assimilation to a one-dimensional model. *Frontiers in Marine Science*, 9, 980388. <https://doi.org/10.3389/fmars.2022.984236>
- Marañón, E., & Holligan, P. M. (1999). Photosynthetic parameters of phytoplankton from 50°N to 50°S in the Atlantic Ocean. *Marine Ecology Progress Series*, 176, 191–203. <https://doi.org/10.3354/meps176191>
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997a). A finite-volume, incompressible Navier stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research*, 102(C3), 5753–5766. <https://doi.org/10.1029/96jc02775>
- Marshall, J., Hill, C., Perelman, L., & Adcroft, A. (1997b). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research*, 102(C3), 5733–5752. <https://doi.org/10.1029/96jc02776>
- Mattern, J. P., Fennel, K., & Dowd, M. (2013). Sensitivity and uncertainty analysis of model hypoxia estimates for the Texas-Louisiana shelf. *Journal of Geophysical Research - Oceans*, 118(3), 1316–1332. <https://doi.org/10.1002/jgrc.20130>
- Mattern, J. P., Fennel, K., & Dowd, M. (2014). Periodic time-dependent parameters improving forecasting abilities of biological ocean models. *Geophysical Research Letters*, 41(19), 6848–6854. <https://doi.org/10.1002/2014gl0106178>
- Mattern, J. P., Song, H., Edwards, C. A., Moore, A. M., & Fiechter, J. (2017). Data assimilation of physical and chlorophyll-a observations in the California Current system using two biogeochemical models. *Ocean Modelling*, 109, 55–71. <https://doi.org/10.1016/j.ocemod.2016.12.002>
- McDonald, C. P., Bennington, V., Urban, N. R., & McKinley, G. A. (2012). 1-D test-bed calibration of a 3-D Lake Superior biogeochemical model. *Ecological Modelling*, 225, 115–126. <https://doi.org/10.1016/j.ecolmod.2011.11.021>
- National Academies of Sciences, Engineering and Medicine. (2022). *A research strategy for ocean-based carbon dioxide removal and sequestration*. The National Academies Press. <https://doi.org/10.17226/26278>
- Nerger, L. (2023). The parallel data assimilation framework (PDAF). *Zenodo*. <https://doi.org/10.5281/zenodo.7861829>
- Nerger, L., & Gregg, W. W. (2007). Assimilation of SeaWiFS data into a global ocean-biogeochemical model using a local SEIK filter. *Journal of Marine Systems*, 68(1–2), 237–254. <https://doi.org/10.1016/j.jmarsys.2006.11.009>
- Nerger, L., & Gregg, W. W. (2008). Improving assimilation of SeaWiFS data by the application of bias correction with a local SEIK filter. *Journal of Marine Systems*, 73(1–2), 87–102. <https://doi.org/10.1016/j.jmarsys.2007.09.007>
- Nerger, L., & Hiller, W. (2013). Software for ensemble-based data assimilation systems – Implementation strategies and scalability. *Computers & Geosciences*, 55, 110–118. <https://doi.org/10.1016/j.cageo.2012.03.026>
- Nerger, L., Janjic, T., Schroter, J., & Hiller, W. (2012a). A regulated localization scheme for ensemble-based Kalman filters. *Quarterly Journal of the Royal Meteorological Society*, 138(664), 802–812. <https://doi.org/10.1002/qj.945>
- Nerger, L., Janjic, T., Schroter, J., & Hiller, W. (2012b). A unification of ensemble square root kalman filters. *Monthly Weather Review*, 140(7), 2335–2345. <https://doi.org/10.1175/Mwr-D-11-00102.1>
- Omta, A. W., Talmy, D., Sher, D., Finkel, Z. V., Irwin, A. J., & Follows, M. J. (2017). Extracting phytoplankton physiological traits from batch and chemostat culture data. *Limnology and Oceanography: Methods*, 15(5), 453–466. <https://doi.org/10.1002/lom3.10172>
- Orr, J. C., Najjar, R. G., Aumont, O., Bopp, L., Bullister, J. L., Danabasoglu, G., et al. (2017). Biogeochemical protocols and diagnostics for the CMIP6 ocean model intercomparison project (OMIP). *Geoscientific Model Development*, 10(6), 2169–2199. <https://doi.org/10.5194/gmd-10-2169-2017>
- Oschlies, A., & Schartau, M. (2005). Basin-scale performance of a locally optimized marine ecosystem model. *Journal of Marine Research*, 63(2), 335–358. <https://doi.org/10.1357/0022240053693680>
- Platt, T., Caverhill, C., & Sathyendranath, S. (1991). Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic. *Journal of Geophysical Research*, 96(C8), 15147–15159. <https://doi.org/10.1029/91JC01118>
- Pradhan, H. K., Völker, C., Losa, S. N., Bracher, A., & Nerger, L. (2019). Assimilation of global total chlorophyll OC-CCI data and its impact on individual phytoplankton fields. *Journal of Geophysical Research: Oceans*, 124(1), 470–490. <https://doi.org/10.1029/2018jc014329>

- Pradhan, H. K., Völker, C., Losa, S. N., Bracher, A., & Nerger, L. (2020). Global assimilation of ocean-color data of phytoplankton functional types: Impact of different data sets. *Journal of Geophysical Research: Oceans*, 125(2), e2019JC015586. <https://doi.org/10.1029/2019JC015586>
- Prieur, C., Viry, L., Blayo, E., & Brankart, J. M. (2019). A global sensitivity analysis approach for marine biogeochemical modeling. *Ocean Modelling*, 139, 101402. <https://doi.org/10.1016/j.ocemod.2019.101402>
- Puxty, R. J., Millard, A. D., Evans, D. J., & Scanlan, D. J. (2016). Viruses inhibit CO₂ fixation in the Most abundant phototrophs on Earth. *Current Biology*, 26(12), 1585–1589. <https://doi.org/10.1016/j.cub.2016.04.036>
- Rohr, T., Richardson, A. J., Lenton, A., Chamberlain, M. A., & Shadwick, E. H. (2023). Zooplankton grazing is the largest source of uncertainty for marine carbon cycling in CMIP6 models. *Communications Earth & Environment*, 4(1), 212. <https://doi.org/10.1038/s43247-023-00871-w>
- Roy, S., Broomhead, D. S., Platt, T., Sathyendranath, S., & Ciavatta, S. (2012). Sequential variations of phytoplankton growth and mortality in an NPZ model: A remote-sensing-based assessment. *Journal of Marine Systems*, 92(1), 16–29. <https://doi.org/10.1016/j.jmarsys.2011.10.001>
- Saito, H., Suzuki, K., Hinuma, A., Ota, T., Fukami, K., Kiyosawa, H., et al. (2005). Responses of microzooplankton to in situ iron fertilization in the western subarctic Pacific (SEEDS). *Progress in Oceanography*, 64(2), 223–236. <https://doi.org/10.1016/j.pocean.2005.02.010>
- Sasai, Y., Smith, S. L., Siswanto, E., Sasaki, H., & Nonaka, M. (2022). Physiological flexibility of phytoplankton impacts modelled chlorophyll and primary production across the North Pacific Ocean. *Biogeosciences*, 19(20), 4865–4882. <https://doi.org/10.5194/bg-19-4865-2022>
- 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(19), 4285. <https://doi.org/10.3390/s19194285>
- Sathyendranath, S., Platt, T., Horne, E. P. W., Harrison, W. G., Ulloa, O., Outerbridge, R., & Hoepffner, N. (1991). Estimation of new production in the ocean by compound remote sensing. *Nature*, 353(6340), 129–133. <https://doi.org/10.1038/353129a0>
- 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., & Oschlies, A. (2003). Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the North Atlantic: Part I - Method and parameter estimates. *Journal of Marine Research*, 61(6), 765–793. <https://doi.org/10.1357/002224003322981147>
- Schartau, M., Wallhead, P., Hemmings, J., Loptien, 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>
- Schourup-Kristensen, V., Sidorenko, D., Wolf-Gladrow, D. A., & Völker, C. (2014). A skill assessment of the biogeochemical model recom2 coupled to the finite element sea ice-ocean model (FESOM 1.3). *Geoscientific Model Development*, 7(6), 2769–2802. <https://doi.org/10.5194/gmd-7-2769-2014>
- Simon, E., Samuelson, A., Bertino, L., & Mouysset, S. (2015). Experiences in multiyear combined state-parameter estimation with an ecosystem model of the North Atlantic and Arctic Oceans using the ensemble Kalman Filter. *Journal of Marine Systems*, 152, 1–17. <https://doi.org/10.1016/j.jmarsys.2015.07.004>
- Singh, T., Counillon, F., Tjiputra, J., Wang, Y., & Gharamti, M. E. (2022). Estimation of ocean biogeochemical parameters in an Earth system model using the dual one step ahead smoother: A twin experiment. *Frontiers in Marine Science*, 9, 775394. <https://doi.org/10.3389/fmars.2022.775394>
- Smith, W. O., Jr., & Donaldson, K. (2015). Photosynthesis-irradiance responses in the Ross Sea, Antarctica: A meta-analysis. *Biogeosciences*, 12(11), 3567–3577. <https://doi.org/10.5194/bg-12-3567-2015>
- Søreide, J. E., Leu, E. V. A., Berge, J., Graeve, M., & Falk-Petersen, S. (2010). Timing of blooms, algal food quality and Calanus glacialis reproduction and growth in a changing Arctic. *Global Change Biology*. <https://doi.org/10.1111/j.1365-2486.2010.02175.x>
- 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), 1733–1768. [https://doi.org/10.1016/S0967-0645\(00\)00159-4](https://doi.org/10.1016/S0967-0645(00)00159-4)
- St-Laurent, P., Friedrichs, M. A. M., Najjar, R. G., Martins, D. K., Herrmann, M., Miller, S. K., & Wilkin, J. (2017). Impacts of atmospheric nitrogen deposition on surface waters of the Western north Atlantic mitigated by multiple feedbacks. *Journal of Geophysical Research: Oceans*, 122(11), 8406–8426. <https://doi.org/10.1002/2017jc013072>
- Taylor, A. H., Geider, R. J., & Gilbert, F. J. H. (1997). Seasonal and latitudinal dependencies of phytoplankton carbon-to-chlorophyll a ratios: Results of a modelling study. *Marine Ecology Progress Series*, 152(1/3), 51–66. <http://www.jstor.org/stable/24857808>
- Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, 106(D7), 7183–7192. <https://doi.org/10.1029/2000jd900719>
- Thacker, W. C. (1989). The role of the Hessian matrix in fitting models to measurements. *Journal of Geophysical Research*, 94(C5), 6177–6196. <https://doi.org/10.1029/JC094iC05p06177>
- Tjiputra, J. F., Polzin, D., & Winguth, A. M. E. (2007). Assimilation of seasonal chlorophyll and nutrient data into an adjoint three-dimensional ocean carbon cycle model: Sensitivity analysis and ecosystem parameter optimization. *Global Biogeochemical Cycles*, 21(1). <https://doi.org/10.1029/2006gb002745>
- Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., et al. (2018). JRA-55 based surface dataset for driving ocean-sea-ice models (JRA55-do). *Ocean Modelling*, 130, 79–139. <https://doi.org/10.1016/j.ocemod.2018.07.002>
- Turner, M. G., & Gardner, R. H. (2015). Introduction to models. In *Landscape ecology in theory and practice: Pattern and process* (pp. 63–95). Springer New York. https://doi.org/10.1007/978-1-4939-2794-4_3
- Vetra-Carvalho, S., Van Leeuwen, P. J., Nerger, L., Barth, A., Altai, M. U., Brasseur, P., et al. (2018). State-of-the-art stochastic data assimilation methods for high-dimensional Non-Gaussian problems. *Tellus Series a-Dynamic Meteorology and Oceanography*, 70(1), 1445364. <https://doi.org/10.1080/16000870.2018.1445364>
- Ward, B. A., Dutkiewicz, S., Jahn, O., & Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnology & Oceanography*, 57(6), 1877–1891. <https://doi.org/10.4319/lo.2012.57.6.1877>
- Ward, B. A., Friedrichs, M. A. M., Anderson, T. R., & Oschlies, A. (2010). Parameter optimisation 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>
- Westberry, T., Behrenfeld, M. J., Siegel, D. A., & Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochemical Cycles*, 22(2). <https://doi.org/10.1029/2007gb003078>
- Xu, M., Liu, Y., Zhao, Z., Fu, K., & Lv, X. (2022). Advancing parameter estimation with characteristic finite difference method (CFDM) for a marine ecosystem model by assimilating satellite observations: Spatial distributions. *Frontiers in Marine Science*, 9, 997537. <https://doi.org/10.3389/fmars.2022.997537>
- Zaiss, J., Boyd, P. W., Doney, S. C., Havenhand, J. N., & Levine, N. M. (2021). Impact of lagrangian sea surface temperature variability on Southern Ocean phytoplankton community growth rates. *Global Biogeochemical Cycles*, 35(8), e2020GB006880. <https://doi.org/10.1029/2020gb006880>