Effects of assimilating phytoplankton carbon in marine ecosystem modelling in NEMO4.0.4-MEDUSA2.0-PDAF2.0

Yumeng Chen[1,2], Dale Partridge[2,3], and Lars Nerger[4]

[1]School of Mathematical, Physical and Computational Sciences, University of Reading, Reading RG6 6ET, UK
[2]National Centre for Earth Observation, University of Reading, Reading RG6 6ET, UK
[3]Plymouth Marine Laboratory, Plymouth, United Kingdom
[4]Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar-und Meeresforschung (AWI), 27570 Bremerhaven, Germany

Correspondence: Yumeng Chen (yumeng.chen@reading.ac.uk)

Abstract. The state of the marine ecosystem can be estimated by a combination of numerical models and satellite observations through data assimilation (DA) methods. Satellite data representing phytoplankton chlorophyll are typically used in operational marine ecosystem prediction. These data are derived from ocean colour ~~from~~ measured by optical satellite observations. Recently a ~~novel~~ new phytoplankton carbon product, from the ESA funded BICEP project made available from the UK CEDA Archive, has been derived through ~~an alternate~~ a novel processing of ocean colour. ~~With the~~ This novel carbon product ~~, the phytoplankton biomass is represented more directly than relying on the chlorophyll~~ captures aspects of phytoplankton biomass composition that chlorophyll alone does not represent. Here, we investigate the effects of assimilating the new carbon product on the modelling of the marine ecosystem. The investigation is carried out in a newly developed global ensemble DA system for the marine ecosystem using a coupled ocean-biogeochemistry model, NEMO-MEDUSA, and the Parallel Data Assimilation Framework. With the ensemble DA system, the evaluation can take the time-dependent uncertainty of the marine ecosystem and the reliability of the ensemble into account. We demonstrate that, compared with ~~solely assimilating~~ assimilating only the chlorophyll product, ~~with the~~ which may degrade the seasonality of phytoplankton carbon, assimilating the new carbon product ~~the DA~~ can provide different patterns of ~~adjustments in the phytoplankton concentration~~ adjustment and seasonal anomalies in phytoplankton concentrations, surface pCO_2, and oxygen. Our findings reveal that simultaneously assimilating both phytoplankton chlorophyll and carbon products in a complex marine ecosystem yields more ~~accurate and~~ balanced estimates of phytoplankton biomass than assimilating a single phytoplankton product.

1 Introduction

The marine ecosystem mediates the global carbon and oxygen budgets and supports human activities (Wanninkhof et al., 2013). In particular, as the base of the food web, phytoplankton sequestrate carbon and emit oxygen, accounting for ~~40% of global carbon uptake (Falkowski, 1994)~~50% of total global net primary production estimated on the order of 50 Gt C yr^{-1} (Falkowski, 1994; Buitenhuis et al., 2013; Johnson and Bif, 2021; Ryan-Keogh et al., 2023). Studies have shown the need for reliable forecasting and reanalysis data for the monitoring and management of the health of marine ecosystems and fisheries (Lehodey et al., 2008; Berx et al., 2011; Fennel et al., 2019).

Existing operational forecasting and reanalysis products adopt data assimilation (DA) methods to combine marine ecosystem models with observations. Observations are used by DA to optimise the model parameters (e.g., Hemmings et al., 2015; Schartau et al., 2017) and estimate the state of the marine ecosystem models (e.g., Ford et al., 2012; Gehlen et al., 2015; Ford and Barciela, 2017; Pradhan et al., 2020). Thanks to the good spatio-temporal coverage of the satellite observations, DA applications focus on assimilating total phytoplankton chlorophyll-a products derived from satellite ocean colour for their use as proxies for phytoplankton biomass in ocean ecosystem modelling (Gregg, 2008; Ford et al., 2012; Simon et al., 2015). Note that, for brevity, chlorophyll-a and chlorophyll are used interchangeably in this study.

One limitation of satellite ocean colour is its inability to differentiate the vertical structure of the euphotic zone near the ocean surface (Arteaga et al., 2022). This issue can be mitigated by the ~~new deployment of~~ deployment of new in-situ observations using autonomous instruments such as floats in the Biogeochemical-Argo program (Verdy and Mazloff, 2017; Ford, 2021) and gliders (Skákala et al., 2021). These in-situ observations not only observe the subsurface chlorophyll but also other key state variables of the marine ecosystem, such as the oxygen and nutrients that cannot be observed by satellite. ~~These features enable more accurate state estimates beyond phytoplankton (Ford, 2021). However, due to their recent deployment and the cost of instruments and limited spatial coverage~~However, besides its limited spatial and temporal coverage, operational assimilation of BGC-Argo floats can still face other technical challenges. For example, the chlorophyll estimates from BGC-Argo floats can be biased because they are estimated from measurements of fluorescence (Roesler et al., 2017). These biases are not addressed in studies assimilating BGC-Argo floats, e.g., in synthetic experiments performed by Ford (2021). This suggests the need for further research on bias correction for BGC-Argo floats in DA systems. In addition, research also suggested discrepancies between observations from satellites and BGC-Argo floats (Baudena et al., 2025). Such discrepancies pose additional challenges for joint assimilation of satellite and BGC-Argo data. Hence, satellite ocean colour is still a vital resource, especially on a global scale.

Assimilating ~~the~~ satellite phytoplankton chlorophyll products (referred to as chlorophyll hereafter) ~~products~~ can effectively control errors in modelled chlorophyll. However, even though the chlorophyll provides a proxy for biomass~~by adjusting primary production, this~~, assimilating chlorophyll does not necessarily correct other ~~constituents that make up~~ compositions of phytoplankton such as silicate, nitrogen and carbon, or factors that constrain chlorophyll such as the availability of light, nutrients, and species composition. For example, if the model simulations overestimate the phytoplankton biomass, reducing the chlorophyll by DA does not directly change phytoplankton carbon or nitrogen concentrations. Hence, operational marine biogeochemical DA requires balancing schemes~~to distribute the chlorophyll increments~~, which distribute DA increments of chlorophyll to other model variables for a consistent model state across variables, e.g., different phytoplankton compositions and nutrients, after the DA update (Hemmings et al., 2008), or rely on estimates of error covariances from an ensemble of model forecasts (Pradhan et al., 2020). Yet, as discussed by Skákala et al. (2024), phytoplankton community compositions are the most uncertain and least observable variables in the marine ecosystem even when a balancing scheme is applied. An additional challenge is that phytoplankton are typically represented in models by a number of functional types (PFTs), such as different sizes of chlorophyll and nitrogen. This means that either PFT products need to be derived from the ocean colour data prior to assimilation or the increments of total chlorophyll need to be split during assimilation. Recent ~~studies show improvements when~~

efforts were made to derive PFTs from satellite data with uncertainty validations (see e.g., Xi et al., 2021). These PFT products led to studies showing improvements from assimilating PFTs instead of total chlorophyll data~~are assimilated~~ (Ciavatta et al., 2018; Skákala et al., 2018; Pradhan et al., 2020), as inferring PFTs from chlorophyll errors using only balancing schemes is challenging. Alternatively, Pradhan et al. (2020) showed improvements in updating multiple PFTs through the use of ensemble error covariances when only total chlorophyll was assimilated. However, ~~it~~ this is not as desirable as directly assimilating PFT products. Nevertheless, due to the various definitions of PFTs in different models, universal PFT products are not yet available.

Besides chlorophyll data, ~~with further assumptions,~~ other phytoplankton products, such as particulate inorganic ocean carbon and particulate organic ocean carbon, can be derived from satellite ocean colour (Siegel et al., 2005; Yang et al., 2024) ~~.~~ independently of chlorophyll in their derivation. In particular, satellite ocean colour can be used to derive phytoplankton carbon (Behrenfeld et al., 2005; Graff et al., 2015; Bellacicco et al., 2020) using an empirical relationship between the backscattering of particles and phytoplankton carbon. These products are used to understand the ocean's biological pump and carbon cycles in addition to the chlorophyll products (Siegel et al., 2023; Bourdin et al., 2025). Recently, in the ESA Biological Pump and Carbon Export Processes (BICEP) project, a new phytoplankton carbon (referred to as carbon hereafter) product has been derived from the ocean colour observations (Sathyendranath et al., 2021b). The ~~carbon product could complement the chlorophyll product as an indicator of the phytoplankton biomass (Arteaga et al., 2016; Jakobsen and Markager, 2016)~~product is derived using an approach different from the backscattering of particles used in other carbon products. The new product is based on the chlorophyll-to-carbon ratio from a photoacclimation model where the ratio varies with the photosynthetically active radiation available in the ocean mixed layer. Compared to other carbon products with limited quality assessments (e.g., NASA Ocean Biology Process , the product's chlorophyll-to-carbon ratio is validated with field data (Sathyendranath et al., 2020).

In this study, we explore the effects of assimilating the newly derived carbon product ~~.~~ so that we can understand whether assimilating phytoplankton carbon can provide additional information compared with assimilating the chlorophyll product alone in the modelling system. Note that, here, as this study focuses on assessing the modelling response of assimilating carbon products, we do not seek the best carbon product for DA, and opt to use the validated off-the-shelf observational dataset. To our knowledge, whilst there are studies assimilating particulate organic carbon (POC, Kettle, 2009; Xiao and Friedrichs, 2014) , phytoplankton carbon assimilation is not well-studied in the literature. This new observation product ~~enables~~ allows us to investigate direct DA adjustments to carbon without the balancing scheme. Besides phytoplankton variables, we further investigate the impact of the assimilation on other model variables in a marine ecosystem model, such as zooplankton, pCO_2, and oxygen. These evaluations enable us to assess the potential benefits of assimilating the new ocean colour product and its impact in comparison with the widely assimilated total chlorophyll data. To perform such an evaluation, we adopt the coupled ocean-biogeochemistry model, NEMO-MEDUSA (Madec et al., 2023; Yool et al., 2013). Ford (2021) has applied NEMOVAR, a 3DVar system, to the global coupled NEMO-MEDUSA. In order to incorporate new model variables into the 3DVar system one would need an estimation of forecast errors. These are typically estimated by sophisticated background error covariance modelling accounting for the correlation scales and balancing conditions of the forecast errors. Without these efforts for the background error covariance, a 3DVar system cannot be used with the new product. To overcome this issue, instead of using the existing NEMOVAR setup, we develop an ensemble DA system using the Parallel Data Assimilation Framework (PDAF,

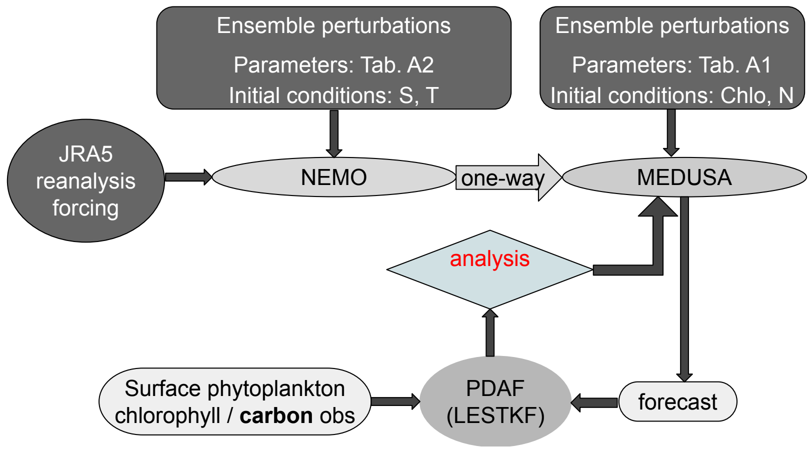

Figure 1. An illustration of the data assimilation system where an ensemble of forecasts is obtained from the one-way coupled NEMO-MEDUSA forced by the JRA-55 reanalysis whereas the satellite surface chlorophyll and carbon observations are assimilated. The ~~resulting~~ analysis (reconstructed state) is used to initialise the following ensemble forecast. Perturbed parameters are listed in Tab. A2 for MEDUSA and Tab. A1 for NEMO. Here, Parallel Data Assimilation Framework (PDAF) is the computational framework and local error subspace transform Kalman filter (LESTKF) is the DA algorithm provided by PDAF.

Nerger and Hiller, 2013). In this ensemble DA system, the forecast errors can be estimated by an ensemble of forecasts. Moreover, compared to NEMOVAR, where the background error covariance matrix does not change with time, the ensemble DA system can provide a time-dependent forecast error covariance matrix accounting for the temporal change of error dynamics.

In Sect. 2, we provide a description of the full ensemble DA system, comprised of the coupled ocean-biogeochemistry model~~and the observations~~, the assimilated observations, and independent in-situ observations for assessment. In Sect. 3, the processing of the state vector and observations used in this study are described. In Sect. 4, a suite of experiment setups used for ~~evaluating the effect~~ assessing the effects of assimilating the carbon product ~~is~~ are described. We evaluate the statistical skill of the analysis and the impact of the DA on the modelled marine ecosystem and pCO_2 and oxygen in Sect. 5. In Sect. 6, we discuss the benefits of assimilating additional satellite ocean colour products and the future developments of the ensemble DA system.

2 Data assimilation system

To investigate the impact of assimilating the newly derived carbon product on the modelling of the marine ecosystem, we develop a new global ensemble ocean-biogeochemistry DA system employing the Parallel Data Assimilation Framework (PDAF, Nerger and Hiller, 2013), as detailed in Fig. 1.

2.1 NEMO-MEDUSA-PDAF

The physical ocean model is the Nucleus for European Modelling of the Ocean (NEMO) version 4.0.4, which solves the primitive equations of the ocean (Madec et al., 2023). The model is configured on a 1° tripolar global extended ORCA grid with 75 vertical z*-coordinate levels. The model surface is forced by atmospheric wind, air temperature, dew point temperature, humidity, precipitation, sea level pressure, snowfall rate, and long- and short-wave radiation provided by the JRA-55 reanalysis product (Kobayashi et al., 2015).

The marine ~~biogeochemistry is modelled using~~ ecosystem model coupled to NEMO is the Model of Ecosystem Dynamics, nutrient Utilisation, Sequestration and Acidification (MEDUSA), an intermediate complexity model (Yool et al., 2013). ~~The coupling to~~ Although marine biogeochemical processes can have a strong impact on physics (Manizza et al., 2005; Skákala et al., 2022), the models in current DA systems are still primarily one-way coupled (Gehlen et al., 2015; Yumruktepe et al., 2022; Polton et al., 2023). Hence, even though two-way coupling can provide a better feedback between the ocean and biogeochemistry, following the common practice of using one-way coupled model avoids the need to investigate the impact of two-way coupling, which is outside the scope of this work. Hence, the coupling to NEMO is one-way; that is, the marine ecosystem model is forced by the physical ocean, but has no feedback from the ecosystem to the physics.

The model includes nutrients, phytoplankton, zooplankton and detritus ~~with nitrogen as the primary currency~~ to simulate marine nitrogen, silicon, iron, alkalinity and oxygen cycles with a benthic ecosystem for seafloor organic pools. In the model, the evolution of phytoplankton concentration is primarily represented in terms of nitrogen biomass. These features allow MEDUSA to simulate complex marine ecosystem processes relevant in climate models. Beyond these components, MEDUSA is capable of simulating the CO_2 uptake of marine ecosystems~~, an important component of the carbon cycle for climate prediction~~. The carbon cycle simulation is enabled by the $p\mathrm{CO_2}$, pH, alkalinity, and by carbonate species like H_2CO_3, HCO_3^- and CO_3^{2-}.

In MEDUSA, the ~~primary currency of nitrogen drives the model evolution. The phytoplankton biomass , including its primary production and losses,~~ phytoplankton biomass is represented by nitrogen. ~~Based on the nitrogen,~~ The phytoplankton chlorophyll is also explicitly modelled as a prognostic variable, ~~acting~~ influenced by the primary production of the phytoplankton biomass. Phytoplankton chlorophyll in turn act as the light limit of the primary production in MEDUSA~~and evolving due to losses of nitrogen by a space- and time-dependent scaling factor~~. Phytoplankton biomass in MEDUSA is divided into diatom and non-diatom PFTs~~based on their sizes. Similarly, zooplankton are divided into microscopic and mesoscopic size classes. Moreover~~. The diatom phytoplankton variable is responsible for silicon uptake and biogenic silica production and the non-diatom phytoplankton variable represents the non-silicifying phytoplankton community. Hence, diatom phytoplankton silicate is also explicitly modelled ~~, which limits the primary production of diatom phytoplankton.~~ besides nitrogen and chlorophyll. MEDUSA models diatom phytoplankton as "large phytoplankton" based on the fact that diatoms are a key component of large phytoplankton even though diatom phytoplankton span a large range of sizes (Yool et al., 2011). This is motivated by the fact that zooplankton graze phytoplankton according to their sizes as zooplankton are are divided into microscopic and mesoscopic size classes. For consistency with the ~~observed dataset~~ observations the nitrogen fields in this study are represented as carbon, with the conversion between carbon and nitrogen following the model assumption of a fixed ratio between C:N of 6.625:1. For clarity, we describe nitrogen in terms of its carbon equivalent (hereafter referred to simply as "carbon").

The data assimilation algorithm is implemented using PDAF, a flexible and efficient framework to enable ensemble data assimilation. The coupling between NEMO and PDAF is adapted based on an open-source NEMO-PDAF coupling available at https://github.com/PDAF/NEMO-PDAF. This ensemble DA system provides the flexibility for assimilating not only the ocean physical variables but also arbitrary marine biogeochemical variables. This system also comes with the flexibility to implement a new observation module, which permits the assimilation and comparison of different observation products in this study. The efficient NEMO-PDAF implementation allows for in-memory exchange of data between the model and the DA algorithms. As a result, the DA can be performed without interrupting the model simulation. This study adopts the local error subspace transform Kalman filter (LESTKF, Nerger et al., 2012), which is a specific ensemble DA algorithm provided by the computational framework of PDAF. Constructed with the assumption of a Gaussian distribution, the LESTKF approximates the forecast and analysis error distribution by an ensemble of model forecasts. The LESTKF can assimilate sparse and irregularly spaced observations without spatial interpolation of the observation product. In our configuration of the LESTKF, model variables at each grid point assimilate observations adjacent to that point within a radius of 200 km. The DA increment at grid points without observations is computed based on the spatial correlation between the grid point and the observation locations.

2.2 Observations

2.2.1 Assimilated observations

As our goal is to investigate the effects of assimilating the carbon product instead of constructing the best performing DA system, we follow common practices in operational systems (Skákala et al., 2025). In this study, two datasets are assimilated into NEMO-MEDUSA: a new monthly phytoplankton carbon product and a chlorophyll-a dataset for reference. This means that the DA system inherits some challenges in operational DA as we do not assimilate additional observations to improve vertical profiles.

The carbon product is created from the ESA Biological Pump and Carbon Export Processes (BICEP) Project (Sathyendranath et al., 2020). The carbon is derived from ocean colour using empirical models and additional datasets (Kulk et al., 2020). Note that the derived carbon product makes use of chlorophyll product suggesting a correlation between these two products. The dataset contains monthly global surface carbon from 1998 to 2020 at a spatial resolution of 9 km, with phytoplankton divided into PFTs of pico-, nano-, and micro-phytoplankton. This size division differs from the PFTs used in MEDUSA, whose phytoplankton are divided into diatom and non-diatom classes. Thus, instead of assimilating individual PFTs, the total surface carbon is assimilated. The analysis is distributed to the ~~size-dependent~~ model variables by post-processing, as will be discussed in detail in Sect. 3.

As a ~~benchmark~~reference, the commonly used satellite chlorophyll is also assimilated. The assimilated data set is the global chlorophyll-a data product gridded on a geographic projection, Version 5.0 from the ESA Ocean Colour Climate Change

Initiative (OC-CCI) project (http://www.esa-oceancolour-cci.org/). The product merges data from satellite sensors of SeaWiFS (Sea-viewing Wide Field-of-view Sensor), MODIS (Moderate Resolution Imaging Spectroradiometer), VIIRS (Visible Infrared Imaging Radiometer Suite), and OLCI (Ocean and Land Colour Instrument), matched to MERIS (Medium Resolution Imaging Spectrometer). The global daily product spans a time period of $1997-2020$, and has a spatial resolution of 4 km. Based on validation against in situ observations, this dataset provides an estimate of both observation error and biases. In this study, we assimilate both composite daily and monthly bias-corrected surface chlorophyll products.

Both observation products do not contain information under clouds. Daily products only contain pixels with available observations. The composite monthly product is computed as the monthly average of available daily data excluding missing data.

2.2.2 Independent assessment observations

To assess the DA system, we use in situ observations of chlorophyll and carbon data. The in situ chlorophyll data are obtained from the BGC-Argo program (Group, 2016; Roemmich et al., 2019). The BGC-Argo program extends the Argo program, which is an international program that deploys automatic instruments floating with ocean currents. The BGC-Argo floats provide vertical profiles for both physical variables such as temperature, salinity, and pressure, and biogeochemical variables. In this study, only quality-controlled measurements of chlorophyll with delayed-mode corrections from the upper 50 m were used, focusing the evaluation on the near-surface productive layer and the depth range most directly influenced by the DA. During the experiment period, the dataset has, on average, 448 observations per month.

The in situ observations of carbon are not routinely monitored, and hence are less common compared to chlorophyll. Here, we use the analytical measurements of phytoplankton carbon from the field campaign in the North Atlantic Aerosols and Marine Ecosystems Study (NAAMES). The project provides ship-based measurements for plankton stocks, rate processes, and community compositions. These in situ carbon observations are analytically determined for cells less than 64 μm using a BD Influx Flow Cytometer using methods detailed by Graff et al. (2015). Due to the limitations of the field campaign, the in situ dataset is limited to Northwest Atlantic and weeks of data in November 2015 and May 2016.

3 Experiment configurations

The DA component of the global ensemble marine ecosystem modelling system needs to handle both carbon and chlorophyll products, and need to have the ability to modify the model variables in the state vector. Additionally, the specification of forecast and observation uncertainties for those variables is crucial for the DA system to be effective.

3.1 State vector setup

Model variables that are directly modified by DA algorithms are represented mathematically as a state vector. Although any variables can be included technically, in our configuration, the state vector comprises only model variables for which corresponding observations are available. For example, if only chlorophyll observations are assimilated, the state vector consists solely of surface total chlorophyll, obtained by summing over all PFTs. This is implemented similarly for the carbon product. Depending on the experiment setup (Sect. 4), the state vector may contain carbon, chlorophyll, or both.

The LESTKF algorithm achieves an optimal state estimate when both forecast and observation errors follow a Gaussian distribution. Under the assumption that phytoplankton biomass follows a log-normal distribution ~~(Barnes et al., 2010)~~ (Campbell, 1995; Barnes et al., 2010), the state vector is transformed from phytoplankton concentration by applying the $\log_{10}$ function to each ensemble member following common practices of marine biogeochemical DA (Ford and Barciela, 2017; Ford et al., 2012). After each DA step, the state vector is transformed back to concentrations by a base-10 exponential function. The analysis *increment* is then computed as the difference between the analysis and forecast fields.

~~Here, we distribute the increment of total phytoplankton into~~ As the state vector contains only the total phytoplankton concentrations instead of model variables of PFTs, to exert a meaningful impact on DA cycles, the model variables have to be updated. There are multiple approaches to adjust model variables based on the total surface concentrations. One systematic approach used in ensemble DA is to update the model variables using an ensemble error covariance matrix (Pradhan et al., 2020; Higgs et al where model variables are included in the state vector along with total phytoplankton concentrations. Although, technically feasible in this system, this approach was not investigated in the context of MEDUSA, which requires further tuning of ensemble perturbations for ensemble cross error covariance. Investigating the update by ensemble covariance deviates from our goal of investigating the impact of the carbon product on the DA system. Hence, even though heuristic and potentially less optimal than relying on ensemble covariance, this work follows the practice in variational operational DA systems where increments in total phytoplankton are distributed to other model variables, including diatom and non-diatom ~~PFTs~~ PFT, by post-processing following Ford (2021). The increment of i-th PFT in the mixing layer, $\delta\mathbf{x}_i$, in MEDUSA is obtained by:

$$\delta\mathbf{x}_i = \delta\mathbf{x}_s \frac{\mathbf{x}_i^f}{\mathbf{x}_s^f}, \tag{1}$$

where $\delta\mathbf{x}_s$ is the increment of total surface chlorophyll or carbon in the state vector, and $\mathbf{x}_s^f$ is the corresponding forecast of the state vector, and $\mathbf{x}_i^f$ is the forecast of i-th ~~PFT~~model variable. When both carbon and chlorophyll are assimilated, post-processing is applied only to the corresponding PFTs. Due to the use of the heuristic post-processing, the impact of DA on model variables, such as PFTs, is more than a result of the observation product. It also reflects the imposed post-processing assumptions.

3.2 Observation setup

Similar to the modelled phytoplankton variables, observational data are transformed into a Gaussian distribution based on the following analytic equation:

$$\mathbf{y}^\mu = \log_{10}\mathbf{y} - \frac{1}{2}\ln(10)\sigma^2 \tag{2}$$

where $\mathbf{y}$ is the observed phytoplankton concentration, $\mathbf{y}^\mu$ is the mean of the transformed Gaussian distribution, and σ is the standard deviation of the Gaussian observation $\mathbf{y}^\mu$.

Equation (2) requires the standard deviation of the Gaussian distribution. For chlorophyll observations, ~~σ_i~~ σ is the same as the error provided by the observation product, ~~however~~which can be expressed as:

$$\mathbf{y}_o^\mu = \mathbf{y}_t^\mu + \sigma, \tag{3}$$

where $\mathbf{y}_t^\mu$ is the unknown true logarithm of chlorophyll concentration and $\mathbf{y}_o^\mu$ is the observed logarithm of chlorophyll concentration. However, the carbon product does not provide an error estimate. In this study, the carbon observation error is estimated as the chlorophyll observation error inflated by 10%. This means that we assume the carbon error, expressed for logarithmic carbon concentrations, to be 1.1σ. As a result, both observations produce a similar spatial pattern of observation error. This treatment is justified by several considerations. First, both products are derived from the same ocean colour measurements. Hence, we can assume that both products share similar sources of measurement error. Second, this choice of observation error avoids overconfidence in the carbon product that could lead to overly strong analysis increments. Third, phytoplankton biomass is assumed to follow a log-normal distribution. Due to this assumption, observation errors expressed as the corresponding standard deviation of a Gaussian distribution are multiplicative instead of additive to ~~observations. This~~ the unknown true observations:

$$\mathbf{y}_o = 10^{\mathbf{y}_o^\mu} = 10^{\mathbf{y}_t^\mu} 10^{\sigma} = \mathbf{y}_t 10^{\sigma}. \tag{4}$$

Equation (4) means that the errors are ~~a scaling factor of~~ scaling factors applied to the observation, regardless of the magnitude and unit of the observation itself. Moreover, this choice of the observation error is supported by the statistical diagnostic in observation-space (Desroziers et al., 2005), which is a common approach to tune observation errors based on the consistency of the linear statistical theory. The spatial and temporal averaged carbon observation error of 0.264 used in this study is similar to the observation error of 0.266 given by the diagnostic. This consistent global error statistic does not address changes in error in space and time, which will require further investigation in future studies.

This definition of the carbon observation errors still presents several difficulties. Firstly, to allow the carbon observation to utilise the chlorophyll observation error, both products are assumed to be co-located. However, this is not the case due to the different spatial resolutions of the two datasets. To avoid this issue, the observation errors of chlorophyll and carbon are taken after the observation thinning described below, after which both products are located on the model grid. Secondly, the spatial coverage of these products is different. In regions without chlorophyll observations, the observation errors of carbon cannot be estimated. This means that, in practice, carbon observations are discarded in regions without chlorophyll observations, leading to reduced number of observations. Lastly, the error correlation between the phytoplankton products that arises because they are derived from the same satellite ocean colour measurements is neglected ~~. Recognising~~ because the carbon and chlorophyll products are not estimated independently. Otherwise, there are no error correlations between them. The neglected error correlations are assumed to be accounted for by the 10% inflation in the carbon observation error. Despite these potential issues, these results still provide ~~qualitative validation~~ an assessment of the value of the carbon product.

After applying the transformation in Eq. (2), observation thinning is performed such that observations are coarse-grained to the same low resolution as the model. This prevents dense spatial observations from adversely impacting DA when only diagonal observation error covariances are used without fully accounting for spatial observation error correlations (Fowler et al., 2018) and representation error (Janjić et al., 2018). Following the treatment by Ford et al. (2012), the observation on a model grid is the median of observations within a rectangular box, where the length of the box is the distance between two grid points. This process can also serve as a quality control to eliminate outliers. To ensure that the carbon observation error has a similar spatial pattern as the chlorophyll observation, the second term in Eq. (2) is applied after the observation thinning process.

4 Experiment ~~Setup~~setup

To thoroughly understand the effect of assimilating phytoplankton carbon observations, a suite of experiments is performed over a two-year period for 2015 and ~~2016.~~ 2016 as shown in Tab. 1. The coupled NEMO-MEDUSA model is spun up without DA for a period of 15 years to create the initial state before applying the ensemble perturbation. The 15-year spinup is sufficient to equilibrate surface biogeochemical variables, but the deeper ocean variables may not be able to reach a steady state within the spinup period as they may require centuries or millennia to reach steady state. However, all of our experiments are performed under the same deep ocean conditions which should therefore have a limited impact on our comparisons between experiments. The initial ensemble is generated by perturbing the chlorophyll and carbon fields, ocean temperature and salinity, and selected physical and biogeochemical model parameters as detailed in Appendix A. ~~The following experiments are conducted:~~

- ~~**Freerun**: baseline ensemble experiment without DA~~
- ~~**Daily Chl**: daily chlorophyll assimilation that only updates the chlorophyll variable~~
- ~~**Monthly Chl**: monthly chlorophyll assimilation that only updates the chlorophyll variable~~
- ~~**Monthly Chl+**: monthly chlorophyll assimilation that updates both chlorophyll and carbon variables based on chlorophyll increments as given in Eq. (1)~~
- ~~**Monthly C**: monthly phytoplankton carbon assimilation that only updates the carbon variable~~
- ~~**Monthly C+**: monthly phytoplankton carbon assimilation that updates both carbon and chlorophyll variables based on carbon increments as given in Eq. (1)~~
- ~~**Monthly Chl & C**: monthly simultaneous phytoplankton chlorophyll and carbon assimilation that updates both chlorophyll and carbon variables~~

In these experiments, the "Daily Chl" experiment benefits from the high-frequency chlorophyll product. However, because only a monthly carbon product is available, monthly experiments are conducted for a fair comparison ~~between~~ among different phytoplankton products. The "Monthly Chl & C" experiment is distinctive from other experiments. In this experiment, the state vector contains both phytoplankton chlorophyll and carbon. The DA update makes use of the correlation between these variables such that each variable is updated by both observation products.

Table 1. Summary of the data assimilation experiments. “Directly updated” refers to variables modified by the data assimilation analysis, whereas post-processing refers to the additional increment-based adjustment defined in Eq. (1).

Experiment	Observation(s)	Model variable(s) directly updated	Post-processing	Frequency
Freerun	None	None	No	None
Daily Chl	Chlorophyll	Chlorophyll	No	Daily
Monthly Chl	Chlorophyll	Chlorophyll	No	Monthly
Monthly Chl+	Chlorophyll	Chlorophyll	Yes	Monthly
Monthly C	Carbon	Carbon	No	Monthly
Monthly C+	Carbon	Carbon	Yes	Monthly
Monthly Chl & C	Chlorophyll & carbon	Chlorophyll & carbon	No	Monthly

The daily assimilation experiment is implemented by assimilating the daily chlorophyll observation at the start of each day. In monthly assimilation experiments, the assimilation takes place at the start of the day in the middle of each month. In all experiments, to ensure a gradual adjustment of the model simulation, an incremental analysis update (IAU) is used with a window of 1 day corresponding to 31 time steps where $\frac{1}{31}$ of the total increment of the model variable is applied for each time step.

5 Results

The aim of this study is to investigate the impact of assimilating the phytoplankton carbon product on the modelled marine ecosystems. To understand the impact of the product globally, each experiment is evaluated against the composite monthly observations described in Sect. 2.2. We further assess the influences of each experiment on marine ecosystem variables such as PFTs, zooplankton and gases such as pCO_2 and oxygen without ~~available observations~~validating these variables against in situ datasets in the present study.

5.1 Statistical scores

~~The global statistical metrics can provide the effect of DA. These metrics evaluate the effectiveness of the DA and~~ Here, the effects of the DA on assimilated variables are evaluated using the bias between the model and observations, the root mean squared differences (RMSDs), continuous rank probability score (CRPS), and ensemble spread. These scores can expose potential disagreements between different observations~~. We also evaluate the uncertainty~~, the uncertainty, and reliability of the ensemble after DA.

5.1.1 Biases

The DA algorithm assumes that the forecastsand observations have no biases. However, ~~biases~~ this assumption is rarely met in real applications. Biases can exist both in observations and model forecasts ~~.~~ for both carbon and chlorophyll. In model forecasts, the bias could be a result of inaccurate model parameters, simplified treatment of model processes and biased initial conditions (Mamnun et al., 2025). In addition to model forecast, the application of fixed C:N ratio could affect the biases. This fixed ratio can be viewed as a global average in both time and space; biases could only occur when the distribution of the spatially and temporally varying C:N ratio is not symmetric that leads to a ratio that is systematically either too large or too small (Tanioka et al., 2022). Quantifying biases in models and observations is non-trivial considering the limited availability of ~~reliable observations (Fowler et al., 2023). As~~ so-called anchor observations with negligible biases (Eyre, 2016; Fowler et al., 2023). Because DA methods assume uncorrelated forecast and observation errors, as a proxy, the bias can be revealed by the misfit between observations and the model forecast over multiple time steps used in the DA~~.~~ ~~Without biases, the~~ (Saha et al., 2014; Zuo et al., 2019; Barton et al., 2021; Mignac et al., 2025). The expectation of the misfit is zero ~~.~~ in the absence of biases. Because the model parameters are not determined by the assimilated observations, this diagnostic can expose biases in the DA systems.

The misfits of logarithmic phytoplankton concentration represent the ratio between the model and observations, expressing relative differences that are independent of the absolute concentration values. Figure 2 shows the histogram of the misfits at the assimilation step during the 2-year experiment period. In Fig. 2a), ~~the~~ chlorophyll is less biased than carbon, but the modelled chlorophyll will, in general, be increased by the DA, as demonstrated by the positive mean misfits (vertical lines). There are differences in the misfit between experiments, with the “Daily Chl” experiment showing a much smaller distribution than the “Monthly Chl” experiments. The magnitude of the mean misfit in “Daily Chl” is similar to the bias given by Ford et al. (2012). Without the high-frequency adjustments provided by daily assimilation, the misfit is less likely to approach zero, and instead exhibits an almost bimodal distribution. Notably, using the balancing scheme to update the phytoplankton carbon in the “Monthly Chl+” experiment slightly improves the expectation of the misfits in chlorophyll~~, suggesting~~. This suggests that making direct changes in the ~~phytoplankton biomass~~ primary composition of phytoplankton biomass in MEDUSA leads to more sustained changes in the model. This is consistent with the study conducted by (Ford and Barciela, 2017).

In contrast, Fig. 2b) shows a negative mean misfit for carbon, which implies that the modelled carbon is constantly reduced by the DA in all experiments assimilating carbon. Unlike in the “Monthly Chl+” experiment, updating chlorophyll in the “Monthly C+” experiment does not lead to an improved distribution of the misfits compared to the “Monthly C” experiment. Interestingly, compared to other experiments, the simultaneous assimilation of chlorophyll and carbon does not bring the averaged misfits closer to zero. This may suggest that the DA ~~obtains~~ reaches a compromise between these observation products and the ratio of chlorophyll and carbon in observations differs from the model.

~~If the concentration of nutrients is sufficient, increased chlorophyll would lead to an increased phytoplankton production, which is in contradiction to an overall lower carbon concentration from observations. This, as also~~ As suggested by the histogram, ~~demonstrates that~~ the errors of the modelled carbon and chlorophyll have different signs, and the use of a balancing scheme may struggle to correctly handle corrections for both chlorophyll and carbon. Considering the carbon data are derived based on chlorophyll data and their ratio is validated by Sathyendranath et al. (2020), in addition to the balancing scheme, the inconsistency of different signs may also arise from model errors. This could be driven by inaccurate modelling of the primary production, the loss of phytoplankton, and/or the coupling of chlorophyll and carbon, where the tendencies of chlorophyll and

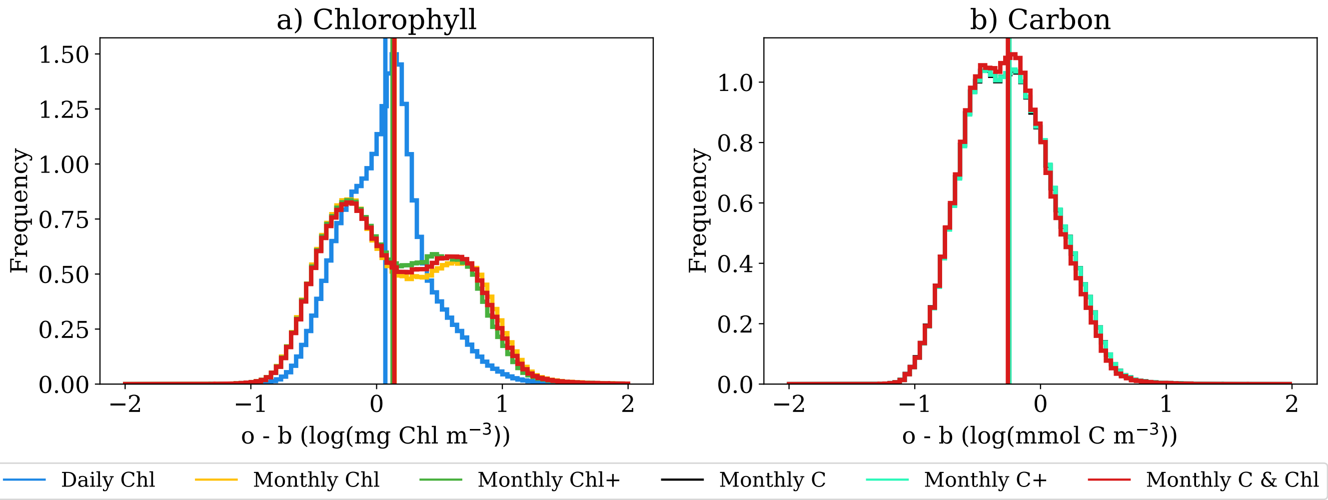

Figure 2. The histogram of the misfit between observations and forecast for log-transformed variables at the analysis step. ~~High~~ Large observation and forecast misfits with low probability in a) are neglected for display purposes. The mean of the misfits for each experiment is presented as vertical lines. This histogram ~~is~~ was constructed by global instantaneous data used by DA over the entire experiment period.

carbon are erroneous given modelled nutrient conditions. The assimilation of both products finds a compromise between the biased model and observation products.

5.1.2 Root mean square difference

~~The effectiveness~~ The effect of the DA is here demonstrated by the root mean square difference (RMSD) between the ensemble mean of monthly averaged model forecastsand monthly composite observations~~in~~. We first compare our model forecasts with independent in situ observations in Fig. 3 using the root mean squared difference of the logarithm of chlorophyll and carbon. The assessment is performed at observation locations using linear interpolation from model grid points. For convenience, the point observations are considered as monthly averaged values, which could cause additional errors in this evaluation. In Fig. 3a), the RMSD of the chlorophyll is calculated for the upper 50 m (around 18 model levels) of the ocean for model forecasts and observations. All experiments assimilating chlorophyll show reduced RMSD of chlorophyll compared to Freerun. Experiments assimilating carbon without chlorophyll show slightly increased RMSD of the chlorophyll. This suggests that assimilating carbon product does not necessarily improve chlorophyll in MEDUSA.

The comparison with in situ carbon observations is more complicated because the carbon observations include only cells smaller than 64 μm. MEDUSA has no explicit definition of the size of diatom and non-diatom phytoplankton. Hence, in Fig. 3b), we compare the observations with model variables of both total phytoplankton (cross signs) and non-diatom phytoplankton (dot signs), which is considered to represent "small" phytoplankton. Assimilating chlorophyll has little impact on the RMSD of non-diatom carbon in November 2015. The RMSD of non-diatom carbon is reduced by assimilating carbon in November 2015 but the RMSD is increased in May 2016. The increased RMSD by assimilating carbon could be a result of post-processing and

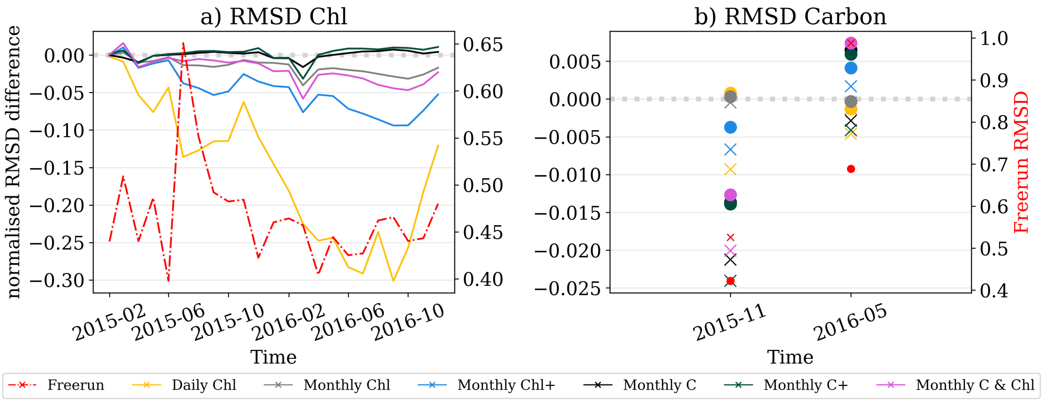

Figure 3. Differences in RMSD between DA experiments and Freerun normalised by the RMSD of Freerun between the logarithm of monthly model forecasts of and composites of chlorophyll from Argo and field campaign data from NAAMES on the left axis. The RMSD of Freerun is shown as a dashed red line on the right axis. The crosses represent comparisons with modelled total carbon while the dots represent comparisons with modelled non-diatom carbon.

mismatch between the observations and PFTs, especially considering the seasonal dependence of the results. When comparing with total modelled carbon, assimilating carbon better constrains the carbon than assimilating chlorophyll regardless of season. Nevertheless, the impact of DA on the RMSD of carbon is smaller in November 2015 than in May 2016. For both carbon and chlorophyll, Daily Chl shows better performance than monthly assimilations. Overall, simultaneous assimilation of chlorophyll and carbon provides a balanced RMSD reduction compared to assimilating a single type of phytoplankton composition.

The assessment against in situ data is still limited by the spatial and temporal coverage. Hence, we further assess the results with assimilated satellite data in Fig. 4a)-b). In Freerun, the RMSD of the logarithmic chlorophyll is higher than that of carbon. Because the RMSD of logarithmic concentration implicitly reflects the ratio between the ensemble mean and observations, the RMSD of the chlorophyll and the carbon can be compared directly. The lower RMSD of ~~the~~ carbon suggests that it is simulated more accurately than chlorophyll in Freerun, potentially due to its role as a representation of phytoplankton biomass in the MEDUSA formulation. The RMSD of chlorophyll is comparable to that reported in Pradhan et al. (2019, 2020), despite of a different biogeochemical model. Here, the RMSD of chlorophyll peaks during the boreal spring and autumn. However, Pradhan et al. (2019) showed a higher RMS error in the boreal autumn than in spring, whereas our results show the opposite pattern, with larger RMSD in autumn. The differences can be explained by various factors including the model formulations, forcing, and ensemble size, which cannot be easily attributed. For example, Pradhan et al. (2019) used the MITgcm-REcoM2 model on a variable resolution from 0.38° to 2° with 20 ensemble members forced by Coordinated Ocean-Ice Reference Experiment (CORE). Our study used NEMO-MEDUSA on a fixed ORCA1 grid with 30 ensemble members forced by JRA-55. Apart from this, REcoM2 uses a varying stoichiometry, while MEDUSA uses fixed ratios. Moreover, our experiments coincide with a strong El Niño event. These could all lead to different responses to the seasonality of the results.

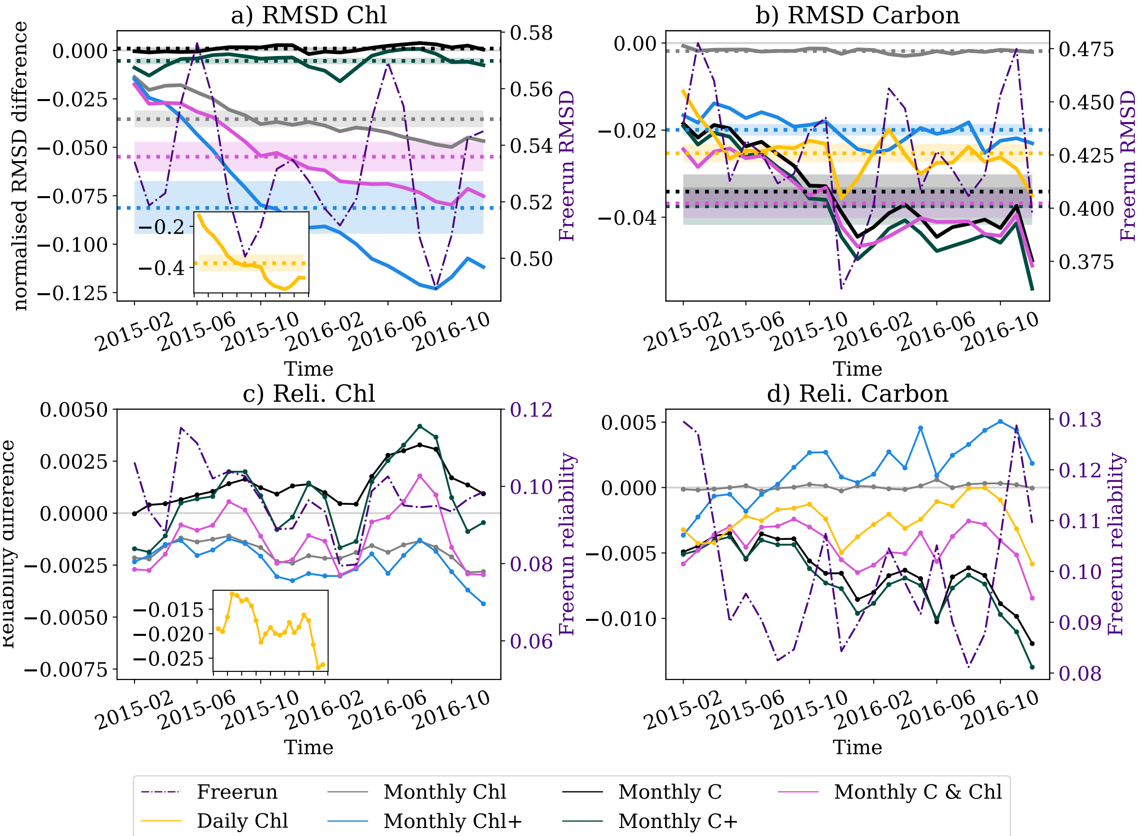

Figure 4. Upper row: time series of the RMSD of the ensemble mean of the logarithmic surface phytoplankton a) chlorophyll and b) carbon. The left axis shows the differences of RMSD between DA experiments and Freerun normalised by the RMSD of Freerun, and the right axis shows the RMSD of Freerun (red line). The light gray horizontal line is the zero line representing no global adjustments compared to Freerun. The inset in a) is the RMSD difference of chlorophyll in the "Daily Chl" experiment. The horizontal dashed lines with shaded area are the temporal average of the RMSD, and the shaded area represents the 95% confidence interval of the temporal average. Lower row: reliability score of the ensemble of surface phytoplankton a) chlorophyll and b) nitrogen compared against the assimilated observations derived from CRPS. The left axis is for the reliability differences between each DA experiment and Freerun, and the right axis is the reliability score of Freerun (red line). The inset in c) is the reliability differences of chlorophyll in the "Daily Chl" experiment.

For each DA experiment, the RMSD is normalised by the RMSD of the Freerun experiment representing the ratio of the RMSD changes. In all experiments, the assimilation reduces RMSD, reflected by the negative difference of the normalised RMSD between the DA experiment and Freerun. The RMSD differences show a decreasing trend over the experiment period as the model state gradually adjusts towards the observations.

For chlorophyll, the largest RMSD reduction is obtained in the "Daily Chl" experiment, as expected. The RMSD is reduced up to more than 40% at the end of the experiment period (see inset in Fig. 4a). With frequent adjustments to the model state, in the daily assimilation experiment, the carbon field is adjusted through modifying the light limitation of primary production in the model forecast, leading to a reduced RMSD of $2-4\%$ in carbon. Moreover, the daily chlorophyll assimilation outperforms monthly carbon assimilations with respect to the carbon field at the start of the assimilation.

The "Monthly Chl" experiment has a much smaller impact on chlorophyll compared to "Daily Chl", and as a result has little influence on phytoplankton carbon biomass. It only reduces the RMSD of the chlorophyll by up to around 5%. This result is consistent with the misfits statistics in Fig. 2. However, by updating the carbon, the RMSD of the "Monthly Chl+" is reduced by up to around 12%, which is higher than the "Monthly Chl" experiment, as it is more able to make lasting adjustments to the marine ecosystem. But it is still not as effective as in the "Daily Chl" experiment. The "Monthly Chl+" experiment also results in RMSD reductions in carbon whereas the "Monthly Chl" experiment has a very limited impact on the RMSD of carbon.

~~Upper row: time series of the RMSD of the ensemble mean of the logarithmic surface phytoplankton a) chlorophyll and b) carbon. The left axis shows the differences of RMSD between DA experiments and Freerun normalised by the RMSD of Freerun, and the right axis shows the RMSD of Freerun (red line). The light gray horizontal line is the zero line representing no global adjustments compared to Freerun. The inset in a) is the RMSD difference of chlorophyll in the "Daily Chl" experiment. Lower row: reliability score of the ensemble of surface phytoplankton a) chlorophyll and b) nitrogen compared against the assimilated observations derived from CRPS. The left axis is for the reliability differences between each DA experiment and Freerun, and the right axis is the reliability score of Freerun (red line). The inset in c) is the reliability differences of chlorophyll in the "Daily Chl" experiment.~~

In the "Monthly C" and "Monthly C+" experiments, in which carbon observations are assimilated, the RMSD reduction of the chlorophyll field is the lowest. In particular, without chlorophyll update, the "Monthly C" experiment results in almost no chlorophyll improvements, and can lead to deteriorated RMSD compared to the Freerun. For carbon the RMSD reduction ~~is~~ in the "Monthly C", "Monthly C+", and "Monthly Chl & C" experiments is similar, and is even larger across the year than the "Daily Chl" experiment. Lastly, even though showing less RMSD reduction in chlorophyll than the "Daily Chl" and "Monthly Chl+", assimilating both datasets in the "Monthly Chl & C" experiment shows the most balanced performance of all monthly experiments with respect to both variables. This shows the benefits of utilising both datasets even if we do not account for their correlations. Moreover, the improvements from "Monthly Chl+" show that updating both phytoplankton constituents is an effective and necessary approach when only chlorophyll is assimilated. The monthly carbon assimilation presents further reduced RMSD in carbon concentrations than only assimilating chlorophyll. The RMSD reduction suggests the need for direct assimilation of both phytoplankton constituents to achieve balanced adjustments and the potential for significant improvements with a more frequent and sophisticated phytoplankton carbon product.

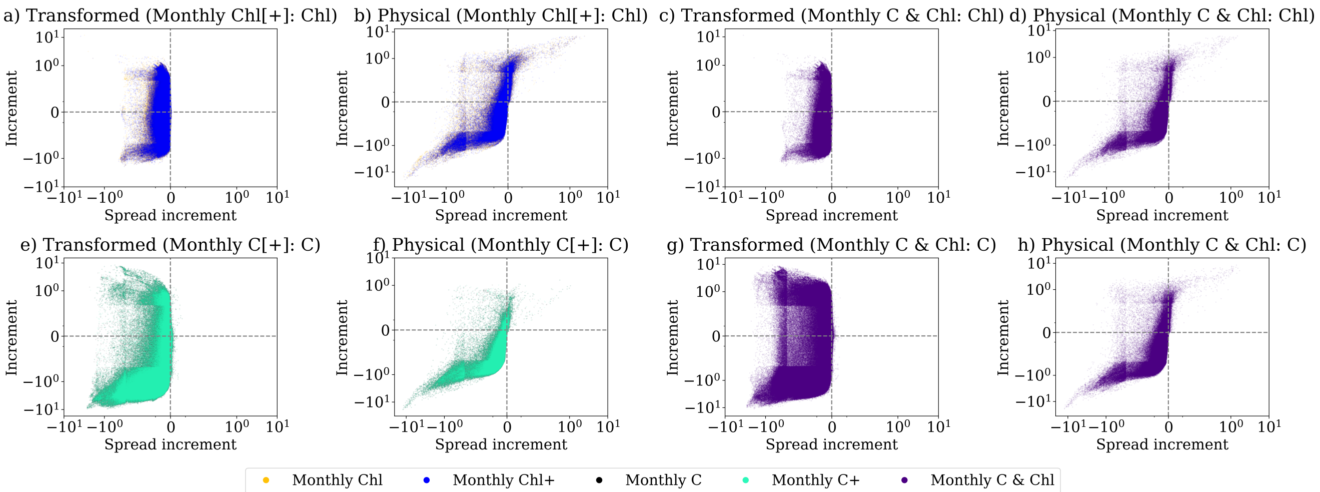

Figure 5. Scatter plot of the ensemble mean increment against the ensemble spread increment ~~in log-scale~~on a log scale. Here, the LESTKF is performed on transformed variables following the Gaussian distribution, and the physical values are the ensemble mean and standard deviation of the phytoplankton concentrations. ~~The squared brackets of the title represent experiments with~~ First row: Chlorophyll increment and ~~without post-processing update of other phytoplankton constituent~~spread increment; second row: carbon increment and spread increment. The “Monthly Chl[+]” ~~represent~~ represents both “Monthly Chl” and “Monthly Chl+” experiments, and “Monthly C[+]” ~~represent~~ represents both “Monthly C” and “Monthly C+” experiments; the variable is given after the colon.

5.1.3 Uncertainties of the ensemble system

In addition to the RMSD of the ensemble mean, a reliable ensemble DA system is expected to represent the probability of the occurrence of a given state, i.e. the uncertainty of the system. ~~The~~ In a perfectly reliable ensemble, by definition, such an ensemble will have an probability distribution function identical to the true probability distribution (Leutbecher and Palmer, 2008; Rodwell . Here, the true probability distribution is represented by the probability distribution of the assimilated observations as calculated by the reliability score of the continuous ranked probability score (CRPS). The CRPS is widely used to diagnose ~~an ensemble forecast~~ensemble forecasts (Hersbach, 2000), representing the differences between the cumulative probability of the distribution given by the model ensemble and observations. The CRPS can be decomposed into reliability and resolution scores. In a perfectly reliable ensemble, the reliability score is zero, and the ensemble represents the true uncertainty of the model simulations. The reliability score increases with less reliable ensemble. ~~In this study the reliability score is calculated against the assimilated observations based on the CRPS~~The reliability score of the CRPS is evaluated on a grid point by grid point basis. As the DA does not guarantee improved reliability by construction, this metric helps us evaluate the accuracy of quantified uncertainty from each experiment.

Fig. 4c)-d) focuses on the reliability score under the assumption that the climatology distributions are given by observations. The ~~reliability score of the Freerun experiment is much higher than that reported in Popov et al. (2024), which focused only~~

~~on the North Atlantic~~non-zero reliability score shows that the ensemble is not perfectly reliable. This highlights the challenge of achieving a reliable global ensemble system and underscores the need for further investigation into quantifying uncertainty in MEDUSA. The ensemble used in this study is in general under-dispersive. This could lead to smaller adjustments than an optimal system. The ensemble reliability can benefit from perturbations in atmosphere forcing and spatially and temporally dependent parameters. However, this will require further careful tuning and possible modifications to the model itself. Such a challenge is also ~~inspected~~ investigated by Anugerahanti et al. (2018) and Mamnun et al. (2022). For the chlorophyll field, daily chlorophyll assimilation continuously adjusts the chlorophyll ensemble, leading to the best reliability score. In the case of monthly DA, assimilating chlorophyll in general shows an improved reliability score compared to Freerun. In comparison, the "Monthly C" and "Monthly C+" experiments could lead to a less reliable chlorophyll ensemble regardless of the chlorophyll update.

For carbon, whilst the "Daily Chl" experiment improves the reliability of the carbon field, the "Monthly Chl" and "Monthly Chl+" experiments deteriorate the carbon reliability. All experiments with carbon assimilation lead to a more reliable carbon ensemble than any experiments without. The fact that assimilating a single observation product could potentially lead to a less reliable ensemble in the other variable demonstrates a benefit for assimilating multiple products in our ensemble system.

It is worth noting that, among all experiments, the simultaneous assimilation of ~~the~~ chlorophyll and carbon does not necessarily lead to the most reliable ensemble. The less reliable ensemble is likely an effect of the contradictory influence of these two types of observations in comparison to Freerun. Yet, the DA can provide a compromise between the two observation datasets.

~~The uncertainty~~ Besides reliability, the ensemble also provides uncertainty information of the model forecasts, which is typically estimated by the ensemble spread, i.e., the standard deviation of the ensemble. By construction of the LESTKF, the analysis ensemble should have a smaller total ensemble spread over the grid points compared to the forecast ensemble as the ensemble spread is a proxy of the error of the ensemble forecast in a reliable ensemble system (Leutbecher and Palmer, 2008). However, this does not mean that the analysis spread is smaller than the forecast spread for every grid point. Instead, the LESTKF reduces the total ensemble spread compared to the total forecast spread across the domain.

The log-transformation of the concentrations described in Sect. 3 leads to a Gaussian distribution, which is used in the LESTKF. Fig. 5 shows increments for these transformed variables as well as for the actual values. As expected, in Fig. 5a), ~~5~~c), ~~5~~e), ~~5~~g), the ensemble spread of the log-transformed variables is reduced by the LESTKF, i.e. the increment (differences between the analysis and forecast spread) is overall negative. However, when transformed back into physical space the DA leads to increased ensemble spread at some grid points. For a Gaussian distribution, the mean and standard deviation are independent of each other, whilst for a lognormal distribution this is not the case. Thus, even if the DA reduced the ensemble spread for the log-transformed variables, the corresponding ensemble spread of the actual variables can increase due to their log-normal distribution. This effect is shown in Fig. 5b), ~~5~~d), ~~5~~f), ~~5~~h), where a clear relationship between the change of the ensemble spread and the ensemble mean is visible. The increased ensemble spread corresponds closely to the increased ensemble mean value. This is a feature of the log-normal distribution, where the mean and spread of the distribution are related. Namely, when the DA increases the ensemble mean value, the ensemble spread of the actual variable can be increased even if the spread of the log-transformed variable is decreased. The ensemble spread can still be reduced if the observations are accurate enough to sufficiently suppress the spread of the log-transformed variable as is visible in Eq. 2. This emphasises the importance of accurate observations following a log-normal distribution, especially in high-concentration regions. When such high accuracy observations are not available, we postulate that post-processing may be necessary to prevent unphysical ensemble members.

In the "Monthly C" and "Monthly C+" experiments of Fig. 5e), the negative carbon bias from Fig. 2 is evident compared to the chlorophyll experiments in Fig. 5a). In contrast, the increment of the carbon in the "Monthly ~~C & Chl~~ Chl & C" experiment of Fig. 5g) shows more similarity with the "Monthly Chl[+]" of Fig. 5a) than the "Monthly C[+]" experiment of Fig. 5e). In addition to the interaction between phytoplankton chlorophyll and carbon, the similarity might be because the chlorophyll observations have lower uncertainty than carbon observations, so that the DA assimilates more information from the chlorophyll product. This suggests that, even if the phytoplankton products yield contradictory increments, the marine ecosystem prediction can benefit from the simultaneous assimilation of both observations for its ability to optimally combine different sources of observations.

5.2 Model adjustments

Besides the statistical metrics, to understand the spatial adjustments of the global system, we now compare the adjustments made to surface phytoplankton chlorophyll and carbon in different experiments.

5.2.1 Adjustments in total phytoplankton

As shown in Fig. 6, Freerun and the observations capture similar spatial patterns, with higher values in polar and subpolar regions of both hemispheres, coastal regions and the eastern equatorial Pacific than in other regions. Such spatial pattern is consistent with other studies, e.g., Yool et al. (2013); Ford (2021). The higher resolution of the observations captures more features, with the model showing higher chlorophyll concentration around the eastern equatorial Pacific region, Southern Ocean, the Benguela Current region, and some of the sub-polar regions of the northern hemisphere. Yet, the model does not capture the high chlorophyll concentration in open ocean areas such as the Atlantic and Indian oceans, and around some of the coastal regions of Antarctica. These all contribute to the overall positive misfits between observations and the model discussed in Sec. 5.1.1.

Similarly, in Fig. 6i) and 6m), the modelled carbon also has lower concentrationsin the open ocean and higher concentration in the eastern equatorial Pacific than the observations. The modelled carbon also shows spatial patterns similar to chlorophyll in the Benguela Current region and the equatorial Atlantic due to biased ocean upwelling (Yool et al., 2013). However, in high-concentration regions, the modelled carbon exhibits excessive values, spanning a wider latitudinal range than the modelled chlorophyll in several areas, such as the eastern equatorial and South Pacific Ocean. These patterns are distinct from the low carbon concentrations simulated in the open ocean basins when compared with satellite observations. The difference in spatial patterns demonstrate that a single observation product cannot fully represent the phytoplankton biomass, highlighting the need for carbon products. It is also worth noting that a strong El Niño occurred during our experiment period which had a strong impact on the eastern Equatorial Pacific. The differences between observations and Freerun may reflect the modelling capability

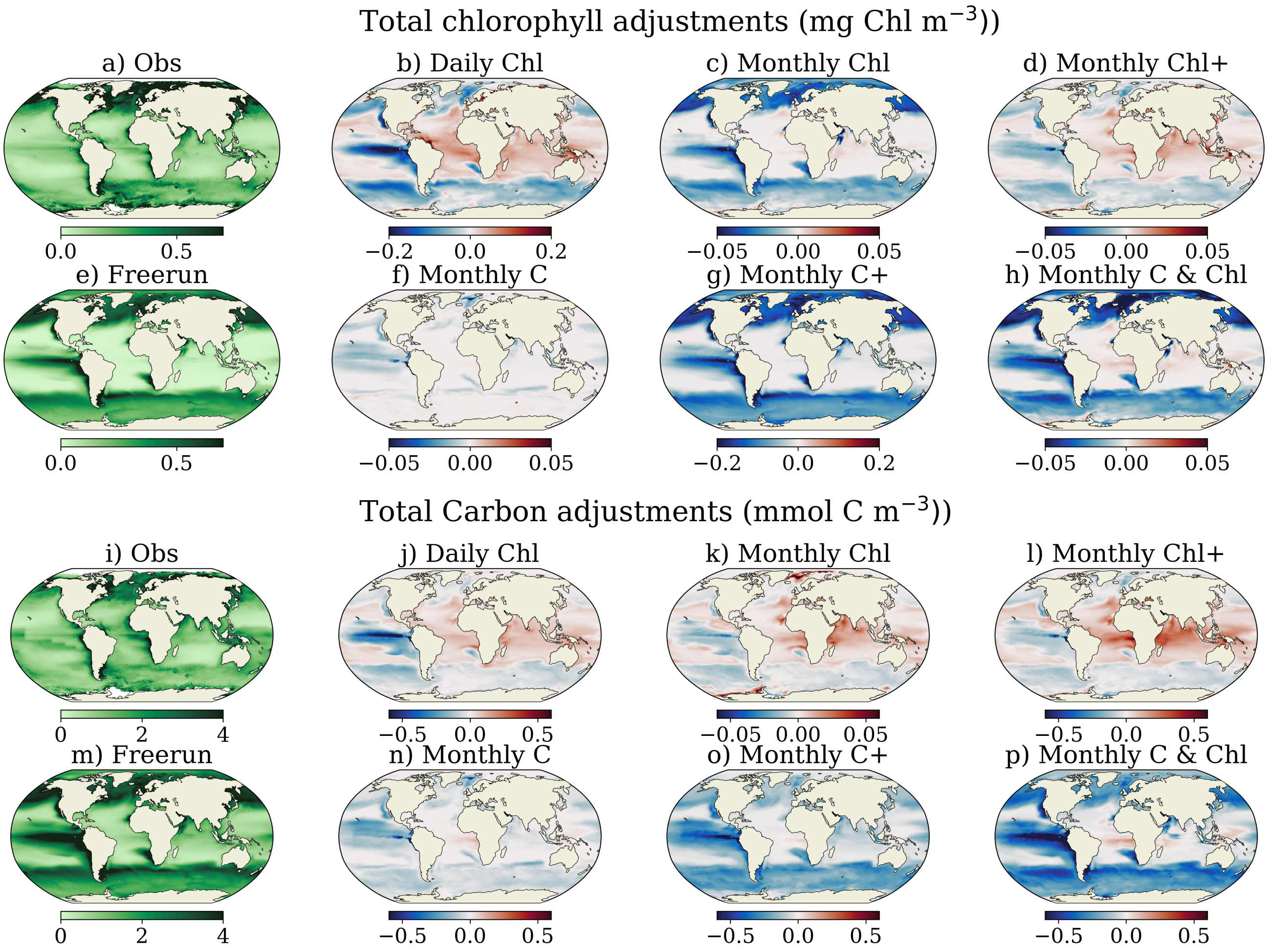

Figure 6. The phytoplankton chlorophyll concentration in a) observation and Freerun e) and the differences between the DA experiments and Freerun experiment in b) - d) and f) - h). The phytoplankton carbon concentration in i) observation and Freerun m) and the differences between the DA experiments and Freerun experiment in j) - l) and n) - p). The differences are computed based on the temporal mean of the ensemble mean of monthly model output over the experiment period. Note that some experiments use a different colour scale.

of the El Niño event as well. This could imply that results from Freerun and other experiments in the eastern Pacific region may not be generalised to years without El Niño events.

In the “Daily Chl” experiment, the frequent assimilation leads to one of the strongest adjustments compared to other experiments, as exhibited in Fig. 6b). These adjustments show a similar spatial pattern as the error correction in Pradhan et al. (2019). The frequent assimilation also results in a similar spatial pattern of adjustments in carbon (Fig. 6b) and j)). However, opposite signs of adjustment exist; e.g. in the Nordic Seas, the DA reduces chlorophyll but ~~increased~~ increases carbon. This increase in carbon is at odds with observations. This contradiction is related to a weak increase in chlorophyll during the spring bloom combined with a strong reduction in other seasons. This reduction has limited impact on the carbon, as the region is predominantly nutrient limited.

Similarly, the “Monthly Chl” and “Monthly C” experiments also only assimilate one type of observation without updating other phytoplankton constituents via post-processing. The less frequent adjustments in these experiments lead to smaller adjustments as shown by Fig. 6c), f), k) and n). In “Monthly Chl”, the weak model response of carbon to the monthly chlorophyll assimilation limits the adjustment of chlorophyll itself, as indicated by its colour scale. The changes in chlorophyll are not maintained over the one-month forecast, and this effect is amplified by the lack of a direct influence on carbon. Compared to “Monthly Chl+” in Fig. 6d, the stronger decreases and weaker increases of chlorophyll in “Monthly Chl” result from the MEDUSA formulation where a modified slope of the photosynthesis–irradiance curve, scaled by the chlorophyll-to-biomass ratio represented by carbon. The “Monthly C” experiment results in the weakest adjustment of chlorophyll among all experiments due to the infrequent adjustments and lack of direct DA update of chlorophyll.

When both ~~the~~ carbon and chlorophyll are updated, the DA adjustments become stronger. Modifying the chlorophyll based on the carbon adjustments leads to the strongest adjustments of the chlorophyll among all experiments, as exhibited by its colour scale in Fig. 6g). In regions with decreased chlorophyll and carbon, the DA adjustments of “Monthly C+”, “Monthly Chl+” and “Monthly ~~C & Chl~~ Chl & C” experiments share similar spatial patterns. There are regional differences between ~~the different experiments for~~ experiments in increases of phytoplankton. For example, the “Monthly Chl+” experiment shows strong increases in both the chlorophyll and carbon in the Black, Timor and Mediterranean Seas, low- and mid-latitude marine ~~zone, except for~~ zones, except in the eastern equatorial Pacific and ~~at~~ along the coasts of subpolar regions. These increases are absent in the “Monthly C+” experiment, with instead a strong reduction of the carbon and chlorophyll in the Black, Timor and Mediterranean Sea. These are regions where ~~there is the strongest~~ the differences between chlorophyll and carbon observations are strongest. Compared to the experiments that assimilate chlorophyll observations, this also suggests an unrealistic modification of the chlorophyll based on the carbon update.

Disagreements still exist in the simultaneous assimilation of both observations. However, in this case, these disagreements do not arise from changes in individual PFTs or from assumptions ~~of~~ in the post-processing approach. Instead, the chlorophyll and carbon corrections largely depend on their respective observational datasets.

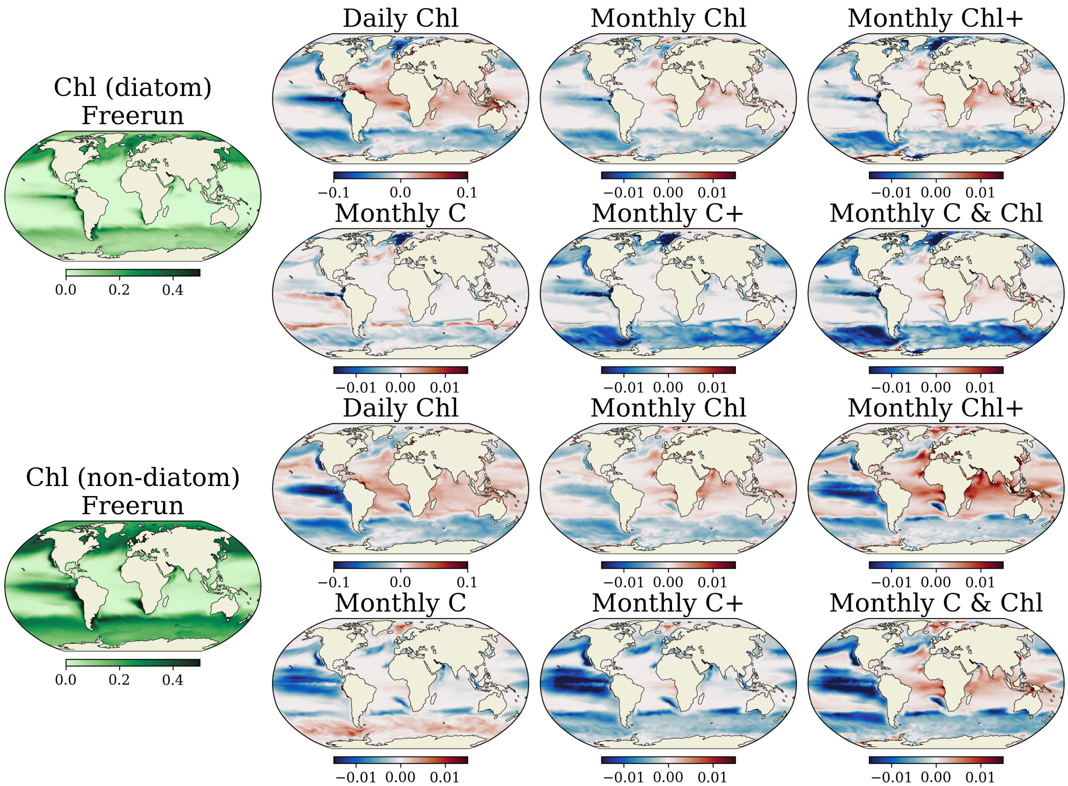

Figure 7. ~~The Freerun~~ Similar to Fig. 6 but for phytoplankton chlorophyll concentration (mg Chl m^{-3}) of diatom and non-diatom classes~~(first column) and the differences of the phytoplankton carbon between the DA experiments and the Freerun experiments~~. ~~The differences are computed based on the temporal mean of the ensemble mean of monthly model output over the experiment period.~~ Note that the “Daily Chl” experiment has a different colour scale from other experiments.

5.2.2 Adjustments in phytoplankton functional types

In MEDUSA, phytoplankton are represented as individual state variables of diatom and non-diatom functional types. These size classes are updated by post-processing in the DA system based on the ratio ~~of forecast between the PFT and~~ between the forecast PFT field and the total phytoplankton field as described in Sect. 3.1. Figs. 7 and 8 show the changes to each functional type for chlorophyll and carbon respectively. The diatom functional group tends to have a lower concentration than the non-diatom group for both chlorophyll and carbon content, with ~~similar spatial patterns consistent~~ spatial patterns similar to the total concentration (Fig. 6).

Although the spatial pattern of phytoplankton functional types in the Freerun experiment ~~is~~ are similar to that of the total phytoplankton, DA adjustments do not necessarily follow the same spatial pattern as their total concentration. These differences are evident in the cases where only one of the chlorophyll or carbon fields is updated.

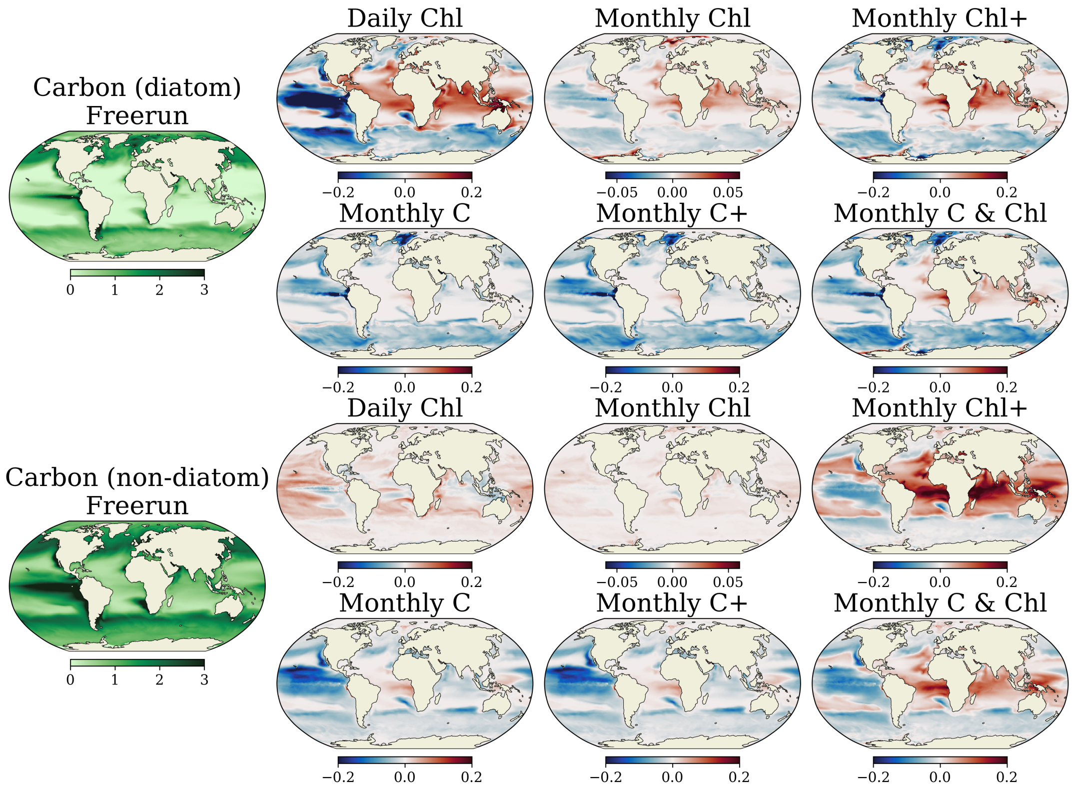

Figure 8. ~~The Freerun~~ Similar to Fig. 6 but for phytoplankton ~~carbon~~ nitrogen concentration (mmol C m^{-3}) of diatom and non-diatom classes~~(first column) and the differences of the phytoplankton carbon between the DA experiments and the Freerun experiments~~. ~~The differences is computed based on the temporal mean of the ensemble mean of monthly model output over the experiment period.~~ Note that the “Monthly Chl” experiment has a different colour scale than other experiments.

In the “Daily Chl” and “Monthly Chl” experiments, even though the spatial pattern of each chlorophyll size class follows the adjustments of the total chlorophyll, the carbon adjustments show a more complex spatial pattern. This is because in these experiments, the size classes of chlorophyll are updated based on the DA increments, but the carbon is adjusted through model responses to those increments. Diatom carbon responds similarly to chlorophyll, with strong decreases of phytoplankton in the eastern equatorial Pacific, whilst non-diatoms have a much weaker response. In the Southern Ocean and along the Alaska and California currents, the model shows increased non-diatom carbon due to increased primary production as a response to decreased chlorophyll. This is due to how MEDUSA handles the relationship, with decreased chlorophyll reducing the ratio between phytoplankton chlorophyll and carbon. Since the relationship between non-diatom growth rate and light limitation is non-linear, decreased chlorophyll leads to increases in primary production.

In the “Monthly C” experiment, we expect that the chlorophyll will largely follow the phytoplankton biomass changes. However, Figs. 7 and 8 show that both diatom and non-diatom chlorophyll can exhibit contrasting adjustments. These distinct features are evident in the Southern Ocean, the Benguela current and around the North Atlantic drift. These differences are a result of the model formulation of MEDUSA, where the chlorophyll tendency is related to the carbon tendency by a complex function of the primary production and various scaling factors.

The different signs between chlorophyll and carbon adjustments disappear when post-processing is applied in “Monthly Chl+” and “Monthly C+” experiments. In these experiments, the diatom and non-diatom components of chlorophyll and carbon are adjusted proportionally based on the forecast towards the assimilated observations. In the “Monthly C+” experiment, the model responses can become dominant compared to DA increments. There is increased non-diatom carbon in the Nordic Seas due to increased primary production, whilst diatom carbon reflects the decrease seen in the total carbon field. This type of model response occurs for all experiments that assimilate the carbon product.

As discussed in Sect. 5.1, different observation products can lead to contrasting phytoplankton increments. By assimilating both in the “Monthly ~~C & Chl~~ Chl & C” experiment, we see enhanced model adjustments in PFTs that reflect spatial patterns from both sets of observations compared to “Monthly Chl+” and “Monthly C+” experiments.

5.3 Seasonality

Phytoplankton have a strong seasonal variation~~s~~, as seen in Fig. 4. Evaluation of the seasonal adjustments from DA can reveal changes in key processes in marine ecosystem~~s~~ due to different assimilation strategies. It is worth discussing the impact of DA on the spatial pattern of phytoplankton in each season beyond the year-round climatology. As shown in Sect. 5.2.1, observations and each experiment have their own climatology. Here, we focus on the impact of DA on the seasonal anomalies without the climatological mean.

For the sake of simplicity, we only discuss the spatial adjustments in each season for the DA experiments that update both chlorophyll and carbon. The Freerun experiment generally captures the seasonal variations in Fig. 9 and 10, albeit with some bias. Consistent with Yool et al. (2013, 2021), the positive seasonal anomaly of both carbon and chlorophyll shifts to the northern hemisphere in boreal summer and shifts to the southern hemisphere in austral summer, following the seasonal variation of the radiation conditions. The Freerun also successfully captures regions where the seasonal anomalies of the chlorophyll and carbon are different. For example, in boreal winter, the model represents opposite anomalies ~~of~~ in carbon and chlorophyll in the mid-latitude Atlantic, the east Pacific, the Indian Ocean, and the south Pacific. Yet, there are regions where the Freerun experiment disagrees with the observations. For instance, in boreal spring a positive anomaly of phytoplankton extends toward higher latitudes in the North Atlantic that is not present in the observations. The model also fails to capture the correct sign of the seasonal anomaly in observations around the Nordic Seas and Arctic Ocean in boreal autumn for carbon and in boreal summer and autumn for chlorophyll. This suggests that, in addition to year-round biases discussed in Sect. 5.1, the Freerun also shows errors in seasonal biases.

Compared to the seasonal variation of the freerun experiment, each DA experiment shows different changes in its seasonal anomalies. In Fig. 9, although the “Monthly C+” and “Monthly ~~C & Chl~~ Chl & C” experiments show similar spatial patterns of seasonal chlorophyll adjustments in many regions, e.g., in mid-latitude and around ~~equators~~the Equator, the “Monthly ~~C & Chl~~ Chl & C” experiment also exhibits features characteristic of “Monthly Chl+” where only chlorophyll is assimilated.

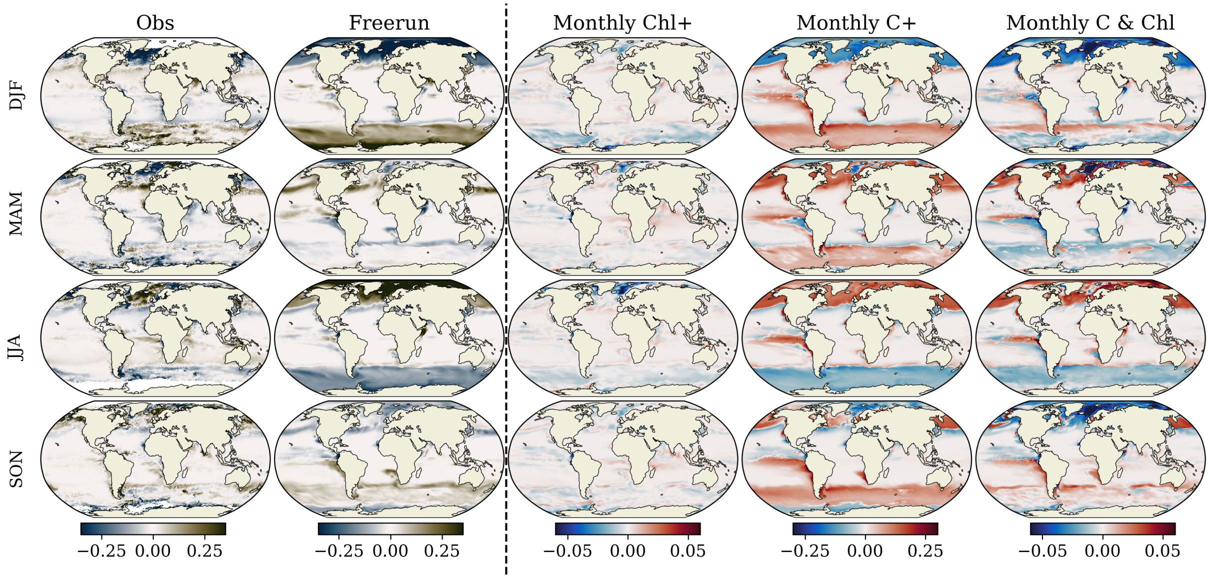

Figure 9. First two columns: seasonal anomalies of chlorophyll concentration (mg Chl m^{-3}) relative to the year-round climatology of observations and the Freerun experiment. Last three columns: differences in seasonal anomalies between the DA experiments and the Freerun experiment (positive values indicate increased anomalies, negative values indicate decreased anomalies). For each experiment, seasonal anomalies are computed relative to that experiment~~’~~’s own year-round climatology. Note: the ’Monthly C+’ experiment uses a different ~~color~~ colour scale than the others.

For example, the reduced seasonal anomaly in the Southern Ocean in boreal winter and spring aligns with “Monthly Chl+” instead of “Monthly C+”. This demonstrates that neither carbon nor chlorophyll alone can represent the seasonal cycle of ~~the~~ phytoplankton biomass. Furthermore, chlorophyll adjustments are larger in the “Monthly C+” experiment than in the “Monthly ~~C & Chl~~ Chl & C” experiment. The positive seasonal anomaly adjustments in “Monthly C+” indicate that assimilating carbon increases the positive and reduces the negative chlorophyll seasonal anomaly, suggesting a stronger seasonal variation than in the Freerun.

Similar to chlorophyll, the carbon anomaly is generally similar in the “Monthly C+” and “Monthly ~~C & Chl~~ Chl & C” experiments as shown in Fig. 10. In contrast, the “Monthly Chl+~~“~~” experiment not only shows the smallest adjustments (see colourbar in Fig. 10) but also fails to show a strengthened seasonal anomaly when the carbon product is assimilated. This suggests that assimilating chlorophyll alone with post-processing could deteriorate the seasonality of modelled global phytoplankton in MEDUSA. When carbon and chlorophyll products are assimilated simultaneously, the strengthened seasonal anomalies become stronger than those obtained by assimilating the carbon product alone.

Comparing Fig. 10 with Fig. 9, the seasonal anomalies are adjusted differently for the chlorophyll and carbon fields for the same experiment. For example, during the boreal spring, assimilating both carbon and chlorophyll reduces the carbon seasonal anomalies near the Arctic, but a strengthened seasonal anomaly is displayed in the same region. The different adjustments show

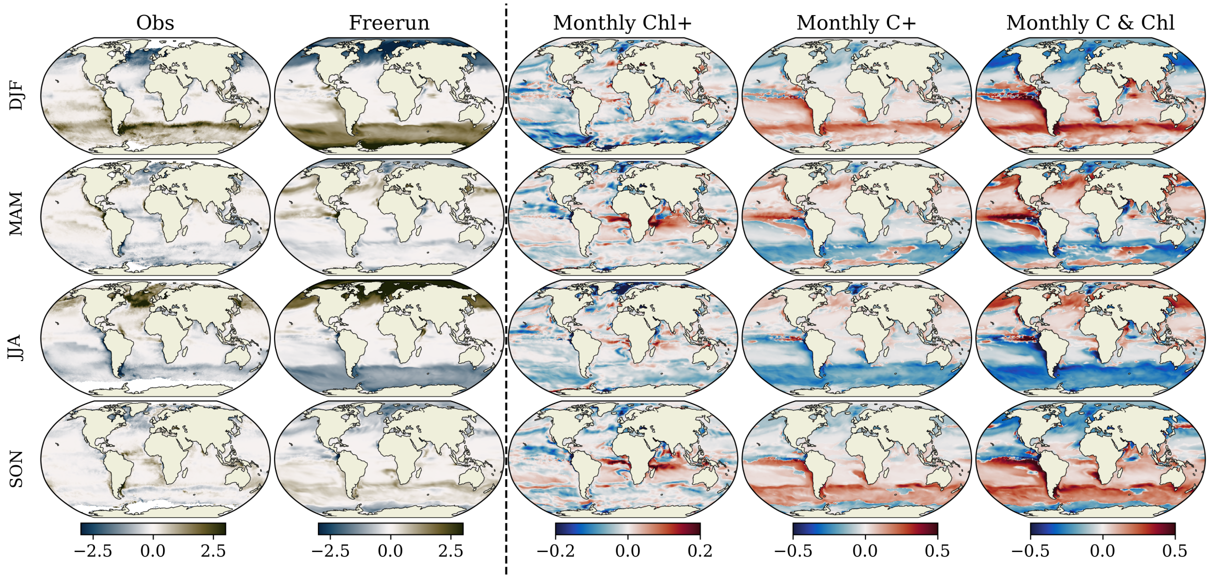

Figure 10. First two columns: seasonal anomalies of carbon concentration (mmol C m^{-3}) relative to the year-round climatology of observations and the Freerun experiment. Last three columns: differences in seasonal anomalies between the DA experiments and the Freerun experiment (positive values indicate increased anomalies, negative values indicate decreased anomalies). For each experiment, seasonal anomalies are computed relative to that experiment~~'~~’s own year-round climatology. Note: the 'Monthly Chl+' experiment uses a different ~~color~~ colour scale than the others.

that relying on multiple PFT observations improves the statistics of phytoplankton seasonal variations in ocean biogeochemical reanalyses.

5.4 Effects on zooplankton, inorganic carbon and oxygen cycles

Even though the ensemble DA system only adjusts phytoplankton, these changes of phytoplankton impact the entire marine ecosystem model, including zooplankton, nutrients and detritus. As with phytoplankton, MEDUSA represents zooplankton with two functional types: meso- and micro-zooplankton, where mesozooplankton are more abundant due to grazing on a wider variety of prey (Fig. 11). The spatial distribution of Freerun exhibits similar features as in other studies using MEDUSA, e.g. Yool et al. (2013); Petrik et al. (2022). In the DA experiments, the spatial pattern of zooplankton adjustments closely follow the phytoplankton adjustments. The spatial pattern of microzooplankton changes are largely identical to the non-diatom phytoplankton in Fig. 8 because they graze on detritus and non-diatom phytoplankton. Mesozooplankton graze on both microzooplankton and diatom phytoplankton, so the spatial pattern is more aligned with the total phytoplankton adjustments in Fig. 6. These changes are consistent across all DA experiments and follow the model formulation.

The time series of the Freerun experiment in Fig. 12 represents the globally averaged ocean surface pCO_2, which has a seasonal variation but a clear upward trend. The pCO_2 is within the range of pCO_2 given by other studies (e.g., Takahashi

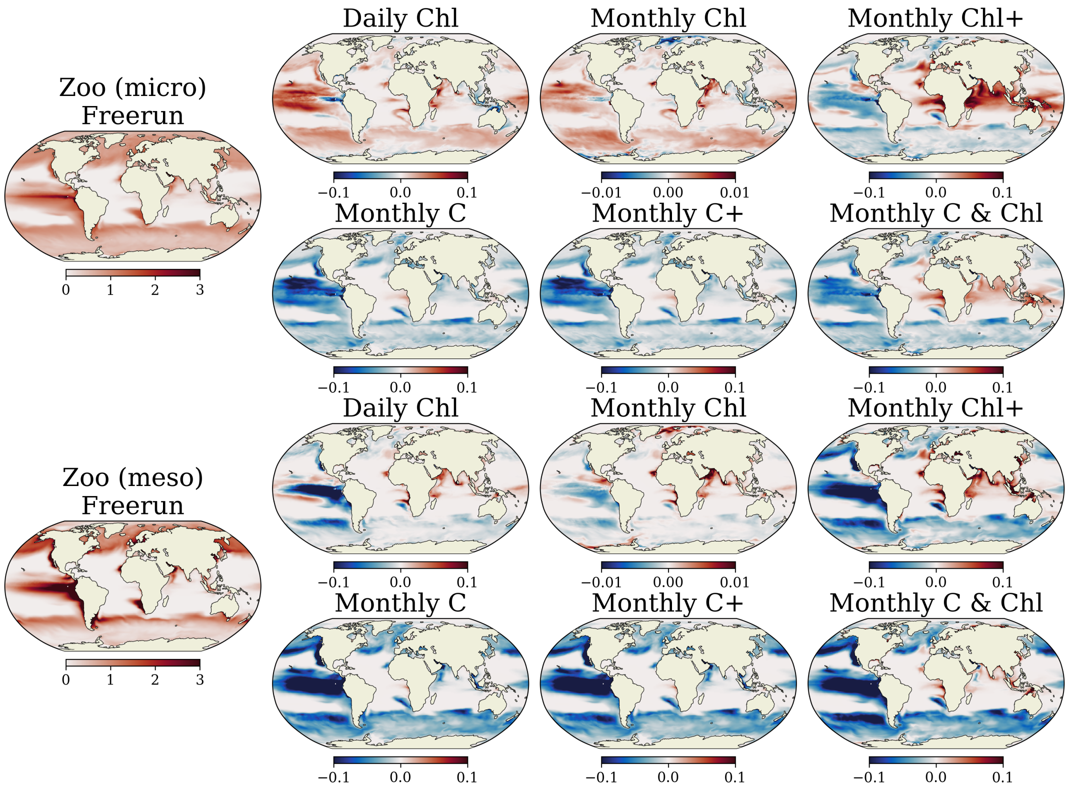

Figure 11. The micro- (ZMI) and meso- (ZME) zooplankton concentration (mmol C m^{-3}) of Freerun (first column) and the differences ~~of the~~ in phytoplankton nitrogen between the DA experiments and the Freerun experiments. The differences are computed based on the temporal mean of the ensemble mean of monthly model output over the experiment period.

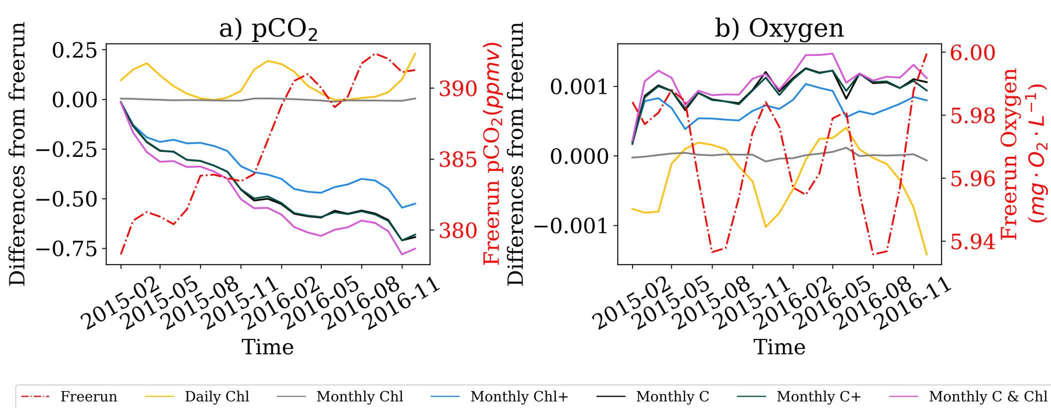

Figure 12. The time series of globally averaged ocean surface pCO_2 (left) and oxygen (right). Shown are Freerun (red line on right axis) and the difference from Freerun for each DA experiment.

et al., 2009; Carroll et al., 2020). The troughs of the seasonal variation correspond to the peaks in phytoplankton blooms. Changes to the globally averaged pCO_2 are minimal in the “Monthly Chl” experiment, but in the “Daily Chl” experiment, the adjustments in the marine ecosystem lead to increased pCO_2 growth. The increase is related to the strong reduction of phytoplankton around the eastern equatorial Pacific and the increased respiratory activities in zooplankton seen in Fig. 11. In other experiments, the pCO_2 concentration increases slower than the Freerun experiment over time. With updates of carbon based on chlorophyll observations in the “Monthly Chl+” experiment, the level of slowdown of pCO_2 increase is not as strong as experiments assimilating phytoplankton carbon. The “Monthly C” and “Monthly C+” experiments show similar adjustments in pCO_2, suggesting that the update of chlorophyll has little impact at monthly resolution. When both chlorophyll and carbon observations are assimilated, the pCO_2 increase is slowed down the most among all experiments. This suggests that the assimilation of carbon observations reduces the pCO_2 at the ocean surface by reducing global phytoplankton carbon.

Unlike pCO_2, the ocean surface oxygen concentration shows a strong seasonal variation without ~~obvious trend~~an obvious trend, which could be a result of the different equilibration timescales of O_2 from pCO_2. The “Monthly Chl” experiment has little impact on the oxygen level. However, the “Daily Chl” experiment shows reduced oxygen concentration while other experiments show increased oxygen at the ocean surface. The “Monthly Chl+” experiment shows a moderate increase, while the “Monthly Chl & C” experiment shows the highest level of increase in oxygen. The increased oxygen in these experiments is a result of less consumption of oxygen from the reduced zooplankton. This reduction in oxygen consumption exceeds the reduction in oxygen generation from photosynthesis due to reduced phytoplankton.

These results show that even though the “Monthly ~~C & Chl~~ Chl & C” experiment does not show the highest RMSD reduction in Sect. 5.1.2, the experiment has a stronger impact on the ocean pCO_2 and oxygen than assimilating only one type of observations. Besides, we expect a sustained impact with time due to the biases in the marine ecosystem model. This highlights the need to assimilate ~~at both a~~ observations at high temporal frequency, and to include both observational datasets to maximise the impact on the marine ecosystem.

6 Conclusions

In this study, we evaluated the effect of assimilating a novel satellite phytoplankton carbon product in comparison to the traditional phytoplankton chlorophyll dataset. To perform the evaluation, a global marine ecosystem ensemble DA system, PDAF-NEMO-MEDUSA, ~~is~~ was constructed using the DA software framework, PDAF.

In all DA experiments, both phytoplankton carbon and chlorophyll become ~~globally more~~ more globally aligned with observations of phytoplankton carbon and chlorophyll. These improvements arise mainly from increments with the same sign ~~of increments~~ when assimilating either observation type in a majority of the global ocean. Despite these alignments, the overall misfits between the model and observations differ in sign for carbon and chlorophyll. This is because these alignments do not hold ~~for~~ in regions with contrasting increments when different observation products are assimilated. ~~These show~~ This shows the disparities between ~~the~~ phytoplankton chlorophyll and carbon. Assimilating both observational datasets simultaneously can account for the contrasting signs.

When only one product is assimilated, adjustments to other phytoplankton components are primarily driven by the model or by post-processing based on increments of the assimilated variables. The majority of these adjustments lead~~s~~to a similar pattern of increments ~~to~~ as the dataset that is being assimilated. However, these adjustments to variables other than the assimilated product could also lead to increased differences with observations and can deteriorate the seasonal variation in some regions. For example, in Monthly Chl+, the seasonality of phytoplankton carbon could deteriorate. This finding could suggest deteriorated seasonality for other DA systems assimilating chlorophyll alone, but, considering the seasonal discrepancy of RMSD with Pradhan et al. (2019), this could also be model and period dependent, which requires further investigations. When no post-processing is used, the model cannot accurately adjust itself to changes, particularly at monthly resolution. In the case of assimilating the widely used chlorophyll observation, only adjusting the light limitation of primary production does not ensure a changed phytoplankton biomass because of the intricate interactions between nutrients and zooplankton. The update from post-processing shows that inferring increments of other variables is not always reliable as well, especially when increments are apportioned to PFTs that do not change proportionally to each other.

Any assimilation experiment has an impact beyond phytoplankton, with the assimilation of a single product potentially having a deteriorating response in other parts of the ecosystem such as the zooplankton, pCO_2 and oxygen. The largest response was seen when assimilating both products.

Our results demonstrate that the simultaneous assimilation of carbon products and chlorophyll can yield more balanced adjustments in phytoplankton biomass. Nevertheless, these adjustments may vary depending on the formulations of individual marine ecosystem models. In MEDUSA, carbon and nitrogen are assumed to have a fixed stoichiometric relationship. By contrast, more complex models such as ERSEM (Butenschön et al., 2016) or quota-based models like REcoM2 (Schourup-Kristensen et al., 2014) may exhibit distinct responses to perturbations in carbon or chlorophyll. ~~We anticipate that in such~~ For example, Ciavatta et al. (2025) compared responses to assimilating chlorophyll in models of different complexities. It is of interest to make such comparisons for carbon products. We hypothesise that in more complex models, where the representation of carbon and chlorophyll dynamics is more sophisticated, the DA is likely to produce more robust ecosystem responses.

In this study, we also investigated the performance of the ensemble DA system. The more frequent daily chlorophyll assimilation improved the reliability of the ensemble compared to monthly assimilation experiments, where the low temporal resolution impacted the ability of the model to tend towards the observed state. Due to the assumption of a log-normal distribution for phytoplankton biomass, the DA system required a statistical transformation to perform the DA using a Gaussian distribution. However, the DA does not necessarily decrease the ensemble uncertainty due to the dependence of the variance on the mean value of the log-normal distribution, even though the resulting analysis is better aligned with the observation than the forecast. From a technical point of view, for future work, the ensemble DA system can be improved through better ensemble generation with more reliable perturbations for ocean forcing and biogeochemical parameters. ~~The~~ A reliable ensemble could further lead to multivariate DA increments that could perform better than the post-processing scheme used here as demonstrated by Pradhan et al. (2019).

There are potential further benefits of assimilating even more information about the phytoplankton. From a DA perspective, assimilating the phytoplankton carbon product still requires further investigation, including better quantification of the observation error ~~, and assessing~~ and assessment of the correlation between the carbon and chlorophyll products. It is of interest to understand the impact of daily assimilation of phytoplankton carbon observations similar to operational biogeochemical DA where chlorophyll is assimilated daily. This would require a daily product in line with the current chlorophyll product. Moreover, independent observations will be needed ~~for quantitatively validating~~ to quantitatively validate the DA system ~~.~~ in addition to phytoplankton chlorophyll and carbon. For example, BGC-Argo floats, the Surface Ocean CO_2 Atlas (SOCAT) or World Ocean Atlas datasets could be used for nutrients, oxygen, and carbonate variables. Furthermore, this study follows the common practice of operational DA systems by only assimilating satellite ocean colour (Skákala et al., 2025). This means that updates to the subsurface structure are not as accurate as assimilating profiles of marine ecosystem variables such as BGC-Argo data. Hence, it is also of interest to investigate a DA system that assimilates both vertical profiles along with ocean colour data.

Code and data availability. The phytoplankton carbon observations from ESA Biological Pump and Carbon Export Processes (BICEP) Project are open access, and are available from UK CEDA Archive (Sathyendranath et al., 2021b). The phytoplankton chlorophyll observations from ESA Ocean Colour Climate Change Initiative: Global chlorophyll-a data products gridded on a geographic projection, Version 5.0 are available from UK CEDA Archive (Sathyendranath et al., 2021a). The source code of NEMO-MEDUSA-PDAF, data analysis and experiment setup are available from the Zenodo repository at https://doi.org/10.5281/zenodo.15209556 (Chen et al., 2025).

Appendix A: Ensemble perturbations

In an ensemble data assimilation system, the forecast error distribution is estimated from the forecast ensemble. The forecast errors come from the initial condition and the epistemic uncertainties arising from physical and biogeochemical parameters. Limited by the computational cost, 30 ensemble members are used in this configuration. As the ensemble is used to estimate the error covariance of the forecast, the limited ensemble size unavoidably leads to sampling errors that could lead to reduced ensemble spread and filter divergence. To overcome these issues, an inflation of 5%, implemented by a forgetting factor of 0.95 (Nerger et al., 2012), is applied to inflate the ensemble during the analysis step.

In this study, the initial conditions of surface chlorophyll and nitrogen as well as 3D ocean temperature and salinity are perturbed. The perturbation of these fields are sampled from a Gaussian distribution. The covariance matrix of the Gaussian distribution is approximated from a model trajectory from Jan - March of 2000 - 2005. In this case, each time step is considered as a sample following the Gaussian distribution. This approach assumes that the uncertainty of model initial condition can be captured by model variability over this period. The ensemble is generated by the 2nd-order exact sampling Pham (2001) provided by PDAF from sampled error covariances with zero mean. The ensemble is run for one month without any DA allowing for improved physical consistency among model variables by model adjustments.

To take the physical model error into account, STochasticOcean physics PACKage (STOPACK, Storto and Andriopoulos, 2021), is used to perturb the physical parametrisations. The stochastic physics package provides both stochastically perturbed parameters (SPP) and a stochastic kinetic energy backscatter (SKEB) scheme. The SPP scheme perturbs a suite of dynamical

Table A1. Perturbed parameters and standard deviation used in SPP scheme in STOPACK

Parameter	Standard deviation
SST and SSS relaxation coefficient	0.5
Solar penetration scheme	0.01
Horizontal diffusivity for velocity and temperature	0.1
Langmuir cell coefficient	0.1
Kolmogoroff dissipation coefficient	0.1
Air-sea drag coefficient	0.01
Tracer damping	0.1

Table A2. Perturbed parameters in MEDUSA. The last column is the constraints of parameters after the sampling from a uniform distribution.

Parameter	Description	Default value	Constraints
$p_{\mu\mathrm{Pn}}, p_{\mu\mathrm{D}}$	microzooplankton grazing preferences for non-diatom (Pn) and detritus (D)	0.75, 0.25	$p_{\mu\mathrm{D}} = 1 - p_{\mu\mathrm{Pn}}$
$p_{m\mathrm{Pn}}$, $p_{m\mathrm{Pd}}$, $p_{m\mathrm{Z}\mu}, p_{m\mathrm{D}}$	mesozooplankton grazing preferences for diatom (Pd), non-diatom (Pn), microzooplankton (Zμ) and detritus (D)	0.15, 0.35, 0.35, 0.15	$p_{m\mathrm{Pd}} = p_{m\mathrm{Z}\mu}$; $p_{m\mathrm{Pn}} = 0.5 - p_{m\mathrm{Pd}}$; $p_{m\mathrm{D}} = p_{m\mathrm{Pn}}$
ϕ	zooplankton grazing inefficiency	0.2	
β_N	zooplankton nitrogen assimilation efficiency	0.77	
β_C	zooplankton carbon assimilation efficiency	0.64	
k_C	zooplankton net C growth efficiency	0.8	
$\mu_{1,\mathrm{Pn}}, \mu_{1,\mathrm{Pd}}$	phytoplankton loss rates d^{-1}	0.02, 0.02	
$\mu_{1,\mathrm{Z}\mu}, \mu_{1,\mathrm{Z}m}$	zooplankton loss rates d^{-1}	0.02, 0.02	

and physical parameters in NEMO while the SKEB scheme transfers the eddy kinetic energy from unresolved scales to resolved scales which mimics an inversed energy cascade. In the SPP scheme, a selection of parameters is perturbed where the perturbations follow a log-normal distribution as given in Tab. A1. The SKEB scheme only perturbs the eddy kinetic energy with a first-order autoregressive model with a standard deviation of 1 and a time decorrelation scale of 1.

Further, a total of 11 biogeochemical parameters have been selected to be perturbed, as shown in Tab. A2. The parameters are first sampled from a uniform distribution $\mathcal{U}(0.8p, 1.2p)$, where p is the default parameter value. Constraints are then imposed shown in the last column of Tab. A2.

Author contributions. YC conducted the experiments, performed the data analysis, and wrote the paper. DP designed the experiments, contributed to the data analysis and paper writing. LN contributed to the code and paper writing.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors acknowledge the UK ~~National~~ Natural Environment Research Council's support for the National Centre for Earth Observation (Contract Number: PR140015). The authors are also grateful for helpful discussions with David Ford, Matthew Martin and Andrea Storto for their support on the ocean stochastic perturbations.

References

Anugerahanti, P., Roy, S., and Haines, K.: A perturbed biogeochemistry model ensemble evaluated against in situ and satellite observations, Biogeosciences, 15, 6685–6711, https://doi.org/10.5194/bg-15-6685-2018, 2018.

Arteaga, L., Pahlow, M., and Oschlies, A.: Modeled Chl:C ratio and derived estimates of phytoplankton carbon biomass and its contribution to total particulate organic carbon in the global surface ocean, Global Biogeochemical Cycles, 30, 1791–1810, https://doi.org/https://doi.org/10.1002/2016GB005458, 2016.

Arteaga, L. A., Behrenfeld, M. J., Boss, E., and Westberry, T. K.: Vertical Structure in Phytoplankton Growth and Productivity Inferred From Biogeochemical-Argo Floats and the Carbon-Based Productivity Model, Global Biogeochemical Cycles, 36, e2022GB007 389, https://doi.org/https://doi.org/10.1029/2022GB007389, e2022GB007389 2022GB007389, 2022.

Barnes, C., Irigoien, X., De Oliveira, J. A. A., Maxwell, D., and Jennings, S.: Predicting marine phytoplankton community size structure from empirical relationships with remotely sensed variables, Journal of Plankton Research, 33, 13–24, https://doi.org/10.1093/plankt/fbq088, 2010.

Barton, N., Metzger, E. J., Reynolds, C. A., Ruston, B., Rowley, C., Smedstad, O. M., Ridout, J. A., Wallcraft, A., Frolov, S., Hogan, P., Janiga, M. A., Shriver, J. F., McLay, J., Thoppil, P., Huang, A., Crawford, W., Whitcomb, T., Bishop, C. H., Zamudio, L., and Phelps, M.: The Navy's Earth System Prediction Capability: A New Global Coupled Atmosphere-Ocean-Sea Ice Prediction System Designed for Daily to Subseasonal Forecasting, Earth and Space Science, 8, e2020EA001 199, https://doi.org/https://doi.org/10.1029/2020EA001199, 2021.

Baudena, A., Riom, W., Taillandier, V., Mayot, N., Mignot, A., and D'Ortenzio, F.: Comparing satellite and BGC-Argo chlorophyll estimation: A phenological study, Remote Sensing of Environment, 326, 114 743, https://doi.org/https://doi.org/10.1016/j.rse.2025.114743, 2025.

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M.: Carbon-based ocean productivity and phytoplankton physiology from space, Global Biogeochemical Cycles, 19, https://doi.org/https://doi.org/10.1029/2004GB002299, 2005.

Bellacicco, M., Pitarch, J., Organelli, E., Martinez-Vicente, V., Volpe, G., and Marullo, S.: Improving the Retrieval of Carbon-Based Phytoplankton Biomass from Satellite Ocean Colour Observations, Remote Sensing, 12, https://doi.org/10.3390/rs12213640, 2020.

Berx, B., Dickey-collas, M., Skogen, M. D., hervé De Roeck, Y., Klein, H., Barciela, R., Forster, R. M., Dombrowsky, E., Huret, M., Payne, M., Sagarminaga, Y., and Schrum, C.: Does Operational Oceanography Address the Needs of Fisheries and Applied Environmental Scientists?, Oceanography, 24, 166–171, http://www.jstor.org/stable/24861249, 2011.

Bourdin, G., Karp-Boss, L., Lombard, F., Gorsky, G., and Boss, E.: Dynamics of island mass effect – Part II: Phytoplankton physiological responses, EGUsphere, 2025, 1–56, https://doi.org/10.5194/egusphere-2025-4261, 2025.

Buitenhuis, E. T., Hashioka, T., and Quéré, C. L.: Combined constraints on global ocean primary production using observations and models, Global Biogeochemical Cycles, 27, 847–858, https://doi.org/https://doi.org/10.1002/gbc.20074, 2013.

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., Bruggeman, J., Cazenave, P., Ciavatta, S., Kay, S., Lessin, G., van Leeuwen, S., van der Molen, J., de Mora, L., Polimene, L., Sailley, S., Stephens, N., and Torres, R.: ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels, Geoscientific Model Development, 9, 1293–1339, https://doi.org/10.5194/gmd-9-1293-2016, 2016.

Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, Journal of Geophysical Research: Oceans, 100, 13 237–13 254, https://doi.org/https://doi.org/10.1029/95JC00458, 1995.

Carroll, D., Menemenlis, D., Adkins, J. F., Bowman, K. W., Brix, H., Dutkiewicz, S., Fenty, I., Gierach, M. M., Hill, C., Jahn, O., Landschützer, P., Lauderdale, J. M., Liu, J., Manizza, M., Naviaux, J. D., Rödenbeck, C., Schimel, D. S., Van der Stocken, T., and Zhang, H.: 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, e2019MS001 888, https://doi.org/https://doi.org/10.1029/2019MS001888, e2019MS001888 2019MS001888, 2020.

Chen, Y., Nerger, L., and Partridge, D.: Assimilation of carbon data into NEMO-MEDUSA, https://doi.org/10.5281/zenodo.15209556, 2025.

Ciavatta, S., Brewin, R. J. W., Skákala, J., Polimene, L., de Mora, L., Artioli, Y., and Allen, J. I.: Assimilation of Ocean-Color Plankton Functional Types to Improve Marine Ecosystem Simulations, Journal of Geophysical Research: Oceans, 123, 834–854, https://doi.org/https://doi.org/10.1002/2017JC013490, 2018.

Ciavatta, S., Lazzari, P., Álvarez, E., Bertino, L., Bolding, K., Bruggeman, J., Capet, A., Cossarini, G., Daryabor, F., Nerger, L., Popov, M., Skákala, J., Spada, S., Teruzzi, A., Wakamatsu, T., Yumruktepe, V., and Brasseur, P.: Control of simulated ocean ecosystem indicators by biogeochemical observations, Progress in Oceanography, 231, 103 384, https://doi.org/https://doi.org/10.1016/j.pocean.2024.103384, 2025.

Desroziers, G., Berre, L., Chapnik, B., and Poli, P.: Diagnosis of observation, background and analysis-error statistics in observation space, Quarterly Journal of the Royal Meteorological Society, 131, 3385–3396, https://doi.org/https://doi.org/10.1256/qj.05.108, 2005.

Eyre, J. R.: Observation bias correction schemes in data assimilation systems: a theoretical study of some of their properties, Quarterly Journal of the Royal Meteorological Society, 142, 2284–2291, https://doi.org/https://doi.org/10.1002/qj.2819, 2016.

Falkowski, P. G.: The role of phytoplankton photosynthesis in global biogeochemical cycles, Photosynthesis Research, 39, 235–258, https://doi.org/10.1007/BF00014586, 1994.

Fennel, K., Gehlen, M., Brasseur, P., Brown, C. W., Ciavatta, S., Cossarini, G., Crise, A., Edwards, C. A., Ford, D., Friedrichs, M. A. M., Gregoire, M., Jones, E., Kim, H.-C., Lamouroux, J., Murtugudde, R., Perruche, C., , t. G. O. M. E. A., and Team, P. T.: Advancing Marine Biogeochemical and Ecosystem Reanalyses and Forecasts as Tools for Monitoring and Managing Ecosystem Health, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00089, 2019.

Ford, D.: Assimilating synthetic Biogeochemical-Argo and ocean colour observations into a global ocean model to inform observing system design, Biogeosciences, 18, 509–534, https://doi.org/10.5194/bg-18-509-2021, 2021.

Ford, D. and Barciela, R.: Global marine biogeochemical reanalyses assimilating two different sets of merged ocean colour products, Remote Sensing of Environment, 203, 40–54, https://doi.org/https://doi.org/10.1016/j.rse.2017.03.040, earth Observation of Essential Climate Variables, 2017.

Ford, D. A., Edwards, K. P., Lea, D., Barciela, R. M., Martin, M. J., and Demaria, J.: Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model, Ocean Science, 8, 751–771, https://doi.org/10.5194/os-8-751-2012, 2012.

Fowler, A. M., Dance, S. L., and Waller, J. A.: On the interaction of observation and prior error correlations in data assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 48–62, https://doi.org/https://doi.org/10.1002/qj.3183, 2018.

Fowler, A. M., Skákala, J., and Ford, D.: Validating and improving the uncertainty assumptions for the assimilation of ocean-colour-derived chlorophyll a into a marine biogeochemistry model of the Northwest European Shelf Seas, Quarterly Journal of the Royal Meteorological Society, 149, 300–324, https://doi.org/https://doi.org/10.1002/qj.4408, 2023.

Gehlen, M., Barciela, R., Bertino, L., Brasseur, P., Butenschön, M., Chai, F., Crise, A., Drillet, Y., Ford, D., Lavoie, D., Lehodey, P., Perruche, C., Samuelsen, A., and Simon, E.: Building the capacity for forecasting marine biogeochemistry and ecosystems: recent advances and future developments, Journal of Operational Oceanography, 8, s168–s187, https://doi.org/10.1080/1755876X.2015.1022350, 2015.

Graff, J. R., Westberry, T. K., Milligan, A. J., Brown, M. B., Dall'Olmo, G., van Dongen-Vogels, V., Reifel, K. M., and Behrenfeld, M. J.: Analytical phytoplankton carbon measurements spanning diverse ecosystems, Deep Sea Research Part I: Oceanographic Research Papers, 102, 16–25, https://doi.org/https://doi.org/10.1016/j.dsr.2015.04.006, 2015.

Gregg, W. W.: Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model, Journal of Marine Systems, 69, 205–225, https://doi.org/https://doi.org/10.1016/j.jmarsys.2006.02.015, physical-Biological Interactions in the Upper Ocean, 2008.

Group, B.-A. P.: The scientific rationale, design and implementation plan for a Biogeochemical-Argo float array, Report, https://doi.org/10.13155/46601, 2016.

Hemmings, J. C., Barciela, R. M., and Bell, M. J.: Ocean color data assimilation with material conservation for improving model estimates of air-sea CO2 flux, Journal of Marine Research, 2008.

Hemmings, J. C. P., Challenor, P. G., and Yool, A.: Mechanistic site-based emulation of a global ocean biogeochemical model (MEDUSA 1.0) for parametric analysis and calibration: an application of the Marine Model Optimization Testbed (MarMOT 1.1), Geoscientific Model Development, 8, 697–731, https://doi.org/10.5194/gmd-8-697-2015, 2015.

Hersbach, H.: Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems, Weather and Forecasting, 15, 559 – 570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2, 2000.

Higgs, I., Bannister, R., Skákala, J., Carrassi, A., and Ciavatta, S.: Hybrid machine learning data assimilation for marine biogeochemistry, Biogeosciences, 23, 315–344, https://doi.org/10.5194/bg-23-315-2026, 2026.

Jakobsen, H. H. and Markager, S.: Carbon-to-chlorophyll ratio for phytoplankton in temperate coastal waters: Seasonal patterns and relationship to nutrients, Limnology and Oceanography, 61, 1853–1868, https://doi.org/https://doi.org/10.1002/lno.10338, 2016.

Janjić, T., Bormann, N., Bocquet, M., Carton, J. A., Cohn, S. E., Dance, S. L., Losa, S. N., Nichols, N. K., Potthast, R., Waller, J. A., and Weston, P.: On the representation error in data assimilation, Quarterly Journal of the Royal Meteorological Society, 144, 1257–1278, https://doi.org/https://doi.org/10.1002/qj.3130, 2018.

Johnson, K. S. and Bif, M. B.: Constraint on net primary productivity of the global ocean by Argo oxygen measurements, Nature Geoscience, 14, 769–774, https://doi.org/10.1038/s41561-021-00807-z, 2021.

Kettle, H.: Using satellite-derived backscattering coefficients in addition to chlorophyll data to constrain a simple marine biogeochemical model, Biogeosciences, 6, 1591–1601, https://doi.org/10.5194/bg-6-1591-2009, 2009.

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, C., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, Journal of the Meteorological Society of Japan. Ser. II, 93, 5–48, https://doi.org/10.2151/jmsj.2015-001, 2015.

Kulk, G., Platt, T., Dingle, J., Jackson, T., Jönsson, B. F., Bouman, H. A., Babin, M., Brewin, R. J. W., Doblin, M., Estrada, M., Figueiras, F. G., Furuya, K., González-Benítez, N., Gudfinnsson, H. G., Gudmundsson, K., Huang, B., Isada, T., Kovač, v., Lutz, V. A., Marañón, E., Raman, M., Richardson, K., Rozema, P. D., Poll, W. H. v. d., Segura, V., Tilstone, G. H., Uitz, J., Dongen-Vogels, V. v., Yoshikawa, T., and Sathyendranath, S.: Primary Production, an Index of Climate Change in the Ocean: Satellite-Based Estimates over Two Decades, Remote Sensing, 12, https://doi.org/10.3390/rs12050826, 2020.

Lehodey, P., Senina, I., and Murtugudde, R.: A spatial ecosystem and populations dynamics model (SEAPODYM) – Modeling of tuna and tuna-like populations, Progress in Oceanography, 78, 304–318, https://doi.org/https://doi.org/10.1016/j.pocean.2008.06.004, 2008.

Leutbecher, M. and Palmer, T.: Ensemble forecasting, Journal of Computational Physics, 227, 3515–3539, https://doi.org/https://doi.org/10.1016/j.jcp.2007.02.014, predicting weather, climate and extreme events, 2008.

Madec, G., Bell, M., Blaker, A., Bricaud, C., Bruciaferri, D., Castrillo, M., Calvert, D., Chanut, J., Clementi, E., Coward, A., Epicoco, I., Éthé, C., Ganderton, J., Harle, J., Hutchinson, K., Iovino, D., Lea, D., Lovato, T., Martin, M., Martin, N., Mele, F., Martins, D., Masson, S., Mathiot, P., Mele, F., Mocavero, S., Müller, S., Nurser, A. G., Paronuzzi, S., Peltier, M., Person, R., Rousset, C., Rynders, S., Samson, G., Téchené, S., Vancoppenolle, M., and Wilson, C.: NEMO Ocean Engine Reference Manual, https://doi.org/10.5281/zenodo.8167700, 2023.

Mamnun, N., Völker, C., Vrekoussis, M., and Nerger, L.: Uncertainties in ocean biogeochemical simulations: Application of ensemble data assimilation to a one-dimensional model, Frontiers in Marine Science, Volume 9 - 2022, https://doi.org/10.3389/fmars.2022.984236, 2022.

Mamnun, N., Völker, C., Vrekoussis, M., and Nerger, L.: Spatially Varying Biogeochemical Parameter Estimation in a Global Ocean Model, Journal of Geophysical Research: Oceans, 130, e2025JC022 752, https://doi.org/https://doi.org/10.1029/2025JC022752, e2025JC022752 2025JC022752, 2025.

Manizza, M., Le Quéré, C., Watson, A. J., and Buitenhuis, E. T.: Bio-optical feedbacks among phytoplankton, upper ocean physics and sea-ice in a global model, Geophysical Research Letters, 32, https://doi.org/https://doi.org/10.1029/2004GL020778, 2005.

Mignac, D., Waters, J., Lea, D. J., Martin, M. J., While, J., Weaver, A. T., Vidard, A., Guiavarc'h, C., Storkey, D., Ford, D., Blockley, E. W., Baker, J., Haines, K., Price, M. R., Bell, M. J., and Renshaw, R.: Improvements to the Met Office's global ocean–sea ice forecasting system including model and data assimilation changes, Geoscientific Model Development, 18, 3405–3425, https://doi.org/10.5194/gmd-18-3405-2025, 2025.

NASA Ocean Biology Processing Group: Aqua MODIS Level-4 Global Mapped Carbon Data, Version 2022.0, https://doi.org/10.5067/AQUA/MODIS/L4M/CARBON/2022.0, dataset; accessed 2026-06-20, 2025.

Nerger, L. and Hiller, W.: Software for ensemble-based data assimilation systems—Implementation strategies and scalability, Computers & Geosciences, 55, 110–118, https://doi.org/https://doi.org/10.1016/j.cageo.2012.03.026, ensemble Kalman filter for data assimilation, 2013.

Nerger, L., Janjić, T., Schröter, J., and Hiller, W.: A Unification of Ensemble Square Root Kalman Filters, Monthly Weather Review, 140, 2335 – 2345, https://doi.org/https://doi.org/10.1175/MWR-D-11-00102.1, 2012.

Petrik, C. M., Luo, J. Y., Heneghan, R. F., Everett, J. D., Harrison, C. S., and Richardson, A. J.: Assessment and Constraint of Mesozooplankton in CMIP6 Earth System Models, Global Biogeochemical Cycles, 36, e2022GB007 367, https://doi.org/https://doi.org/10.1029/2022GB007367, 2022.

Pham, D. T.: Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems, Monthly Weather Review, 129, 1194 – 1207, https://doi.org/https://doi.org/10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2, 2001.

Polton, J., Harle, J., Holt, J., Katavouta, A., Partridge, D., Jardine, J., Wakelin, S., Rulent, J., Wise, A., Hutchinson, K., Byrne, D., Bruciaferri, D., O'Dea, E., De Dominicis, M., Mathiot, P., Coward, A., Yool, A., Palmiéri, J., Lessin, G., Mayorga-Adame, C. G., Le Guennec, V., Arnold, A., and Rousset, C.: Reproducible and relocatable regional ocean modelling: fundamentals and practices, Geoscientific Model Development, 16, 1481–1510, https://doi.org/10.5194/gmd-16-1481-2023, 2023.

Popov, M., Brankart, J.-M., Capet, A., Cosme, E., and Brasseur, P.: Ensemble analysis and forecast of ecosystem indicators in the North Atlantic using ocean colour observations and prior statistics from a stochastic NEMO–PISCES simulator, Ocean Science, 20, 155–180, https://doi.org/10.5194/os-20-155-2024, 2024.

Pradhan, H. K., Völker, C., Losa, S. N., Bracher, A., and Nerger, L.: Assimilation of Global Total Chlorophyll OC-CCI Data and Its Impact on Individual Phytoplankton Fields, Journal of Geophysical Research: Oceans, 124, 470–490, https://doi.org/https://doi.org/10.1029/2018JC014329, 2019.

Pradhan, H. K., Völker, C., Losa, S. N., Bracher, A., and Nerger, L.: Global Assimilation of Ocean-Color Data of Phytoplankton Functional Types: Impact of Different Data Sets, Journal of Geophysical Research: Oceans, 125, e2019JC015 586, https://doi.org/https://doi.org/10.1029/2019JC015586, e2019JC015586 2019JC015586, 2020.

Rodwell, M. J., Lang, S. T. K., Ingleby, N. B., Bormann, N., Hólm, E., Rabier, F., Richardson, D. S., and Yamaguchi, M.: Reliability in ensemble data assimilation, Quarterly Journal of the Royal Meteorological Society, 142, 443–454, https://doi.org/https://doi.org/10.1002/qj.2663, 2016.

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., Oke, P., Owens, W. B., Pouliquen, S., Purkey, S., Scanderbeg, M., Suga, T., Wijffels, S., Zilberman, N., Bakker, D., Baringer, M., Belbeoch, M., Bittig, H. C., Boss, E., Calil, P., Carse, F., Carval, T., Chai, F., Conchubhair, D. O., d'Ortenzio, F., Dall'Olmo, G., Desbruyeres, D., Fennel, K., Fer, I., Ferrari, R., Forget, G., Freeland, H., Fujiki, T., Gehlen, M., Greenan, B., Hallberg, R., Hibiya, T., Hosoda, S., Jayne, S., Jochum, M., Johnson, G. C., Kang, K., Kolodziejczyk, N., Körtzinger, A., Traon, P.-Y. L., Lenn, Y.-D., Maze, G., Mork, K. A., Morris, T., Nagai, T., Nash, J., Garabato, A. N., Olsen, A., Pattabhi, R. R., Prakash, S., Riser, S., Schmechtig, C., Schmid, C., Shroyer, E., Sterl, A., Sutton, P., Talley, L., Tanhua, T., Thierry, V., Thomalla, S., Toole, J., Troisi, A., Trull, T. W., Turton, J., Velez-Belchi, P. J., Walczowski, W., Wang, H., Wanninkhof, R., Waterhouse, A. F., Waterman, S., Watson, A., Wilson, C., Wong, A. P. S., Xu, J., and Yasuda, I.: On the Future of Argo: A Global, Full-Depth, Multi-Disciplinary Array, Frontiers in Marine Science, Volume 6 - 2019, https://doi.org/10.3389/fmars.2019.00439, 2019.

Roesler, C., Uitz, J., Claustre, H., Boss, E., Xing, X., Organelli, E., Briggs, N., Bricaud, A., Schmechtig, C., Poteau, A., D'Ortenzio, F., Ras, J., Drapeau, S., Haëntjens, N., and Barbieux, M.: Recommendations for obtaining unbiased chlorophyll estimates from in situ chlorophyll fluorometers: A global analysis of WET Labs ECO sensors, Limnology and Oceanography: Methods, 15, 572–585, https://doi.org/https://doi.org/10.1002/lom3.10185, 2017.

Ryan-Keogh, T. J., Thomalla, S. J., Chang, N., and Moalusi, T.: A new global oceanic multi-model net primary productivity data product, Earth System Science Data, 15, 4829–4848, https://doi.org/10.5194/essd-15-4829-2023, 2023.

Saha, S., Moorthi, S., Wu, X., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y.-T., ya Chuang, H., Iredell, M., Ek, M., Meng, J., Yang, R., Mendez, M. P., van den Dool, H., Zhang, Q., Wang, W., Chen, M., and Becker, E.: The NCEP Climate Forecast System Version 2, Journal of Climate, 27, 2185 – 2208, https://doi.org/10.1175/JCLI-D-12-00823.1, 2014.

Sathyendranath, S., Platt, T., Žarko Kovač, Dingle, J., Jackson, T., Brewin, R. J. W., Franks, P., nón, E. M., Kulk, G., and Bouman, H. A.: Reconciling models of primary production and photoacclimation, Appl. Opt., 59, C100–C114, https://doi.org/10.1364/AO.386252, 2020.

Sathyendranath, S., Jackson, T., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., Clements, O., Cipollini, P., Danne, O., Dingle, J., Donlon, C., Grant, M., Groom, S., Krasemann, H., Lavender, S., Mazeran, C., Mélin, F., Müller, D., Steinmetz, F., Valente, A., Zühlke, M., Feldman, G., Franz, B., Frouin, R., Werdell, J., and Platt, T.: ESA Ocean Colour Climate Change Initiative (Ocean_Colour_cci): Version 5.0 Data. NERC EDS Centre for Environmental Data Analysis, NERC EDS Centre for Environmental Data Analysis, https://doi.org/https://dx.doi.org/10.5285/1dbe7a109c0244aaad713e078fd3059a, 2021a.

Sathyendranath, S., Platt, T., Kovač, v., Dingle, J., Jackson, T., Brewin, R., Franks, P., Kulk, G., and Bouman, H.: BICEP / NCEO: Monthly global Phytoplankton Carbon, between 1998-2020 at 9 km resolution (derived from the Ocean Colour Climate Change Initiative v5.0 dataset), NERC EDS Centre for Environmental Data Analysis, https://doi.org/https://dx.doi.org/10.5285/6a6ccbb8ef2645308a60dc47e9b8b5fb, 2021b.

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., Ward, B. A., Slawig, T., and Oschlies, A.: Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling, Biogeosciences, 14, 1647–1701, https://doi.org/10.5194/bg-14-1647-2017, 2017.

Schourup-Kristensen, V., Sidorenko, D., Wolf-Gladrow, D. A., and Völker, C.: A skill assessment of the biogeochemical model REcoM2 coupled to the Finite Element Sea Ice–Ocean Model (FESOM 1.3), Geoscientific Model Development, 7, 2769–2802, https://doi.org/10.5194/gmd-7-2769-2014, 2014.

Siegel, D. A., Maritorena, S., Nelson, N. B., and Behrenfeld, M. J.: Independence and interdependencies among global ocean color properties: Reassessing the bio-optical assumption, Journal of Geophysical Research: Oceans, 110, https://doi.org/https://doi.org/10.1029/2004JC002527, 2005.

Siegel, D. A., DeVries, T., Cetinić, I., and Bisson, K. M.: Quantifying the Ocean's Biological Pump and Its Carbon Cycle Impacts on Global Scales, Annual Review of Marine Science, 15, 329–356, https://doi.org/https://doi.org/10.1146/annurev-marine-040722-115226, 2023.

Simon, E., Samuelsen, A., Bertino, L., and Mouysset, S.: 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/https://doi.org/10.1016/j.jmarsys.2015.07.004, 2015.

Skákala, J., Ford, D., Brewin, R. J., McEwan, R., Kay, S., Taylor, B., de Mora, L., and Ciavatta, S.: The Assimilation of Phytoplankton Functional Types for Operational Forecasting in the Northwest European Shelf, Journal of Geophysical Research: Oceans, 123, 5230–5247, https://doi.org/https://doi.org/10.1029/2018JC014153, 2018.

Skákala, J., Ford, D., Bruggeman, J., Hull, T., Kaiser, J., King, R. R., Loveday, B., Palmer, M. R., Smyth, T., Williams, C. A. J., and Ciavatta, S.: Towards a Multi-Platform Assimilative System for North Sea Biogeochemistry, Journal of Geophysical Research: Oceans, 126, e2020JC016 649, https://doi.org/https://doi.org/10.1029/2020JC016649, e2020JC016649 2020JC016649, 2021.

Skákala, J., Bruggeman, J., Ford, D., Wakelin, S., Akpınar, A., Hull, T., Kaiser, J., Loveday, B. R., O'Dea, E., Williams, C. A., and Ciavatta, S.: The impact of ocean biogeochemistry on physics and its consequences for modelling shelf seas, Ocean Modelling, 172, 101 976, https://doi.org/https://doi.org/10.1016/j.ocemod.2022.101976, 2022.

Skákala, J., Ford, D., Fowler, A., Lea, D., Martin, M. J., and Ciavatta, S.: How uncertain and observable are marine ecosystem indicators in shelf seas?, Progress in Oceanography, 224, 103 249, https://doi.org/https://doi.org/10.1016/j.pocean.2024.103249, 2024.

Skákala, J., Ford, D., Haines, K., Lawless, A., Martin, M. J., Browne, P., Chrust, M., Ciavatta, S., Fowler, A., Lea, D., Palmer, M., Rochner, A., Waters, J., Zuo, H., Banerjee, D. S., Bell, M., Carneiro, D. M., Chen, Y., Kay, S., Partridge, D., Price, M., Renshaw, R., Shapiro, G., and While, J.: Marine data assimilation in the UK: the past, the present, and the vision for the future, Ocean Science, 21, 1709–1734, https://doi.org/10.5194/os-21-1709-2025, 2025.

Storto, A. and Andriopoulos, P.: A new stochastic ocean physics package and its application to hybrid-covariance data assimilation, Quarterly Journal of the Royal Meteorological Society, 147, 1691–1725, https://doi.org/https://doi.org/10.1002/qj.3990, 2021.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B., Friederich, G., Chavez, F., Sabine, C., Watson, A., Bakker, D. C., Schuster, U., Metzl, N., Yoshikawa-Inoue, H., Ishii, M., Midorikawa, T., Nojiri, Y., Körtzinger, A., Steinhoff, T., Hoppema, M., Olafsson, J., Arnarson, T. S., Tilbrook, B., Johannessen, T., Olsen, A., Bellerby, R., Wong, C., Delille, B., Bates, N., and de Baar, H. J.: Climatological mean and decadal change in surface ocean pCO2, and net sea–air CO2 flux over the global oceans, Deep Sea Research Part II: Topical Studies in Oceanography, 56, 554–577, https://doi.org/https://doi.org/10.1016/j.dsr2.2008.12.009, surface Ocean CO2 Variability and Vulnerabilities, 2009.

Tanioka, T., Garcia, C. A., Larkin, A. A., Garcia, N. S., Fagan, A. J., and Martiny, A. C.: Global patterns and predictors of C:N:P in marine ecosystems, Communications Earth & Environment, 3, 271, https://doi.org/10.1038/s43247-022-00603-6, 2022.

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

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., Gruber, N., Doney, S. C., McKinley, G. A., Lenton, A., Le Quéré, C., Heinze, C., Schwinger, J., Graven, H., and Khatiwala, S.: Global ocean carbon uptake: magnitude, variability and trends, Biogeosciences, 10, 1983–2000, https://doi.org/10.5194/bg-10-1983-2013, 2013.

Xi, H., Losa, S. N., Mangin, A., Garnesson, P., Bretagnon, M., Demaria, J., Soppa, M. A., Hembise Fanton d'Andon, O., and Bracher, A.: Global Chlorophyll a Concentrations of Phytoplankton Functional Types With Detailed Uncertainty Assessment Using Multisensor Ocean Color and Sea Surface Temperature Satellite Products, Journal of Geophysical Research: Oceans, 126, e2020JC017 127, https://doi.org/https://doi.org/10.1029/2020JC017127, e2020JC017127 2020JC017127, 2021.

Xiao, Y. and Friedrichs, M. A. M.: Using biogeochemical data assimilation to assess the relative skill of multiple ecosystem models in the Mid-Atlantic Bight: effects of increasing the complexity of the planktonic food web, Biogeosciences, 11, 3015–3030, https://doi.org/10.5194/bg-11-3015-2014, 2014.

Yang, G., Bellacicco, M., Organelli, E., and Xing, X.: Global Variability of Phytoplankton Carbon and Non-Algal Particles From Ocean Color Data Based on a Photoacclimation Model, Journal of Geophysical Research: Oceans, 129, e2023JC019 922, https://doi.org/https://doi.org/10.1029/2023JC019922, e2023JC019922 2023JC019922, 2024.

Yool, A., Popova, E. E., and Anderson, T. R.: Medusa-1.0: a new intermediate complexity plankton ecosystem model for the global domain, Geoscientific Model Development, 4, 381–417, https://doi.org/10.5194/gmd-4-381-2011, 2011.

Yool, A., Popova, E. E., and Anderson, T. R.: MEDUSA-2.0: an intermediate complexity biogeochemical model of the marine carbon cycle for climate change and ocean acidification studies, Geoscientific Model Development, 6, 1767–1811, https://doi.org/10.5194/gmd-6-1767-2013, 2013.

Yool, A., Palmiéri, J., Jones, C. G., de Mora, L., Kuhlbrodt, T., Popova, E. E., Nurser, A. J. G., Hirschi, J., Blaker, A. T., Coward, A. C., Blockley, E. W., and Sellar, A. A.: Evaluating the physical and biogeochemical state of the global ocean component of UKESM1 in CMIP6 historical simulations, Geoscientific Model Development, 14, 3437–3472, https://doi.org/10.5194/gmd-14-3437-2021, 2021.

Yumruktepe, V. c., Samuelsen, A., and Daewel, U.: ECOSMO II(CHL): a marine biogeochemical model for the North Atlantic and the Arctic, Geoscientific Model Development, 15, 3901–3921, https://doi.org/10.5194/gmd-15-3901-2022, 2022.

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M.: The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment, Ocean Science, 15, 779–808, https://doi.org/10.5194/os-15-779-2019, 2019.