Using BGC-Argo floats for the assessment of marine biogeochemical models: a case study with CMEMS global forecast system

Alexandre Mignot[1], Hervé Claustre[2,3], Gianpiero Cossarini[4], Fabrizio D'Ortenzio[2,3], Elodie Gutknecht[1], Julien Lamouroux[1], Paolo Lazzari[4], Coralie Perruche[1], Stefano Salon[4], Raphaelle Sauzède[3], Vincent Taillandier[2,3], Anna Teruzzi[4]

[1]Mercator Océan International, Toulouse, France
[2]Laboratoire d'Océanographie de Villefranche-sur-Mer, Villefranche-sur-Mer, CNRS and Sorbonne Université, 06230 Villefranche-sur-Mer, France
[3]Institut de la Mer de Villefranche, CNRS and Sorbonne Université, 06230 Villefranche-sur-Mer, France
[4]National Institute of Oceanography and Applied Geophysics - OGS, Trieste, Italy

Numerical models of ocean biogeochemistry are becoming a major tool to detect and predict the impact of climate change on marine resources and monitor ocean health. Classically, the validation of such models relies on comparison with surface quantities from satellite (such as chlorophyll-*a* concentrations), climatologies, or sparse *in situ* data (such as cruises observations, and permanent fixed oceanic stations). However, these datasets are not fully suitable to assess how models represent many climate-relevant biogeochemical processes. These limitations now begin to be overcome with the availability of a large number of vertical profiles of light, pH, oxygen, nitrate, chlorophyll-*a* concentrations and particulate backscattering acquired by the Biogeochemical-Argo (BGC-Argo) floats network. Additionally, other key biogeochemical variables such as dissolved inorganic carbon and alkalinity, not measured by floats, can be predicted by machine learning-based methods applied to float oxygen concentrations. Here, we demonstrate the use of the global array of BGC-Argo floats for the assessment of biogeochemical models through a concise evaluation of the Copernicus Marine Environment Marine Service (CMEMS) global forecasting system. We first detail the handling of the BGC-Argo data set for model assessment purposes, then we present 22 assessment metrics to quantify the consistency of BGC model simulations with respect to BGC-Argo data. The metrics evaluate either the model state accuracy or the skill of the model in capturing emergent properties, such as the

Deep Chlorophyll Maximums (DCMs) or Oxygen Minimum Zones (OMZs). These metrics are associated with the air-sea CO_2 flux, the biological carbon pump, and the oceanic pH and oxygen levels. We also suggest four diagnostic plots for displaying such metrics.

1. Introduction

Since pre-industrial times, the ocean had taken up ~36 % of the CO_2 emitted by the combustion of fossil fuel (Friedlingstein et al., 2019) leading to dramatic change in the ocean's biogeochemical (BGC) cycles, such as ocean acidification (Iida et al., 2020). Moreover, deoxygenation (Breitburg et al., 2018) and change in the biological carbon pump are now manifesting on a global scale (Capuzzo et al., 2018; Osman et al., 2019; Roxy et al., 2016). Together with plastic pollution (Eriksen et al., 2014) and an increase in fisheries pressure (Crowder et al., 2008), major changes are therefore occurring in marine ecosystems at the global scale. In order to monitor these ongoing changes, derive climate projections and develop better mitigation strategies, realistic numerical simulations of the oceans' BGC state are required.

Numerical models of ocean biogeochemistry represent a prime tool to address these issues because they produce three dimensional estimates of a large number of chemical and biological variables that are dynamically consistent with the ocean circulation (Fennel et al., 2019). They can assess past and current states of the biogeochemical ocean, produce short-term to seasonal forecasts as well as climate projections. However, these models are far from being flawless, mostly because there are still huge knowledge gaps in the understanding of key biogeochemical processes and, as a result, the mathematical functions that describe BGC fluxes and ecosystems dynamics are too simplistic (Schartau et al., 2017). For instance, most models do not include a radiative component for the penetration of solar radiation in the ocean. It has been nevertheless shown that coupling such a component with a BGC model improves the representation of the dynamics of phytoplankton in the lower euphotic zone (Dutkiewicz et al., 2015). Additionally, the parameterisation of the mathematical functions generally results from laboratory experiments on few *a priori* expected representative species and may not be suitable for extrapolation to ocean simulations that need to represent the large range of organisms present in oceanic ecosystems (Schartau et al., 2017; Ward et al., 2010). Furthermore, the assimilation of physical data in coupled physical-BGC models that improves the physical ocean state can paradoxically degrade the simulation of the BGC state of the ocean (Fennel et al., 2019; Park et al., 2018; Gasparin et al., 2021). A rigorous validation of BGC models is thus essential to test their predictive skills, their ability to reproduce BGC processes and estimate confidence intervals on model predictions (Doney et al., 2009; Stow et al., 2009).

However, the validation of BGC models is presently limited by the availability of data. It relies principally on comparison with surface quantities from satellite (such as chlorophyll-*a* concentrations), cruises observations, and few permanent oceanic stations (e.g., Doney et al., 2009; Dutkiewicz et al., 2015; Lazzari et al., 2012, 2016; Lynch et al., 2009; Séférian et al., 2013; Stow et al., 2009). All these datasets neither have a sufficient vertical or temporal resolution, nor a synoptic view, nor can provide all variables necessary to evaluate how models represent climate-relevant processes such as the air-sea CO_2 fluxes, the biological carbon pump, ocean acidification or deoxygenation.

In 2016, the Biogeochemical-Argo (BGC-Argo) program was launched with the goal to operate a global array of 1000 BGC-Argo floats equipped with oxygen (O_2), chlorophyll *a* (Chl*a*) and nitrate (NO_3) concentrations, particulate backscattering (b_{bp}), pH and downwelling irradiance sensors (Biogeochemical-Argo Planning Group, 2016; Claustre et al., 2020). Although the planned number of 1000 floats has not been reached yet, the BGC-Argo program has already provided a large number of quality-controlled vertical profiles of O_2, Chl*a*, NO_3, b_{bp}, and pH (Fig. 1). With respect to O_2, Chl*a*, NO_3, and b_{bp}, the North Atlantic and the Southern Ocean are reasonably well sampled whereas pH is so far essentially sampled in the Southern Ocean. At regional scale, the Mediterranean Sea is also fairly well sampled by BGC-Argo floats (Salon et al., 2019; Terzić et al., 2019). However, there are still, large under-sampled areas, like the subtropical gyres or the sub-polar North Pacific. Nevertheless, the number of quality-controlled observations collected by the BGC-Argo fleet is already greater than any other data set (Claustre et al., 2020). The BGC-Argo data also have a satisfactory level of accuracy and stability over time (Johnson et al., 2017; Mignot et al., 2019). Thanks to machine learning based methods (Bittig et al., 2018; Sauzède et al., 2017), floats equipped with O_2 sensors can be additionally used to derive vertical profiles of NO_3, phosphate (PO_4), silicate (Si), alkalinity (Alk), dissolved inorganic carbon (DIC), pH and pCO_2. All these specificities overcome the limitations of previous data sets, in terms of vertical and temporal resolution, from now and open new perspectives for the validation of BGC models (Gutknecht et al., 2019; Salon et al., 2019; Terzić et al., 2019).

The BGC-Argo data set represent a significant improvement for the assessment of models comparing to large databases such as the World Ocean Atlas (WOA) or the World Ocean Database (WOD). One of the issues of large databases such as WOD is the interoperability of the data that compose it, which, ultimately, affects their overall accuracy (Snowden et al., 2019). Using the BGC-Argo dataset separately is a way to ensure consistent accuracy as it is an interoperable homogenous data set with strict data QC procedures. The BGC-Argo floats also provide observations at high vertical and temporal resolutions and for long periods of time providing nearly continuous time series of the vertical distribution of a number of biogeochemical variables. This is not possible with discrete vertical samplings provided by cruise cast *in situ* measurements or from climatological values derived from the WOA.

We aim to demonstrate the use of the BGC-Argo global array for the assessment of BGC models at the global scale. To that end, we performed a concise evaluation of Copernicus Marine Environment Marine Service (CMEMS) global BGC forecasting system using the global fleet of BGC-Argo floats. We expect that the methodology employed here (from the data handling to the use of assessment metrics) would be useful and informative for other research teams interested in model evaluation with BGC-Argo floats. In this study, the BGC-Argo dataset is used in conjunction with the model evaluation framework developed by Hipsey et al. (2020). In particular, they propose three levels of assessment metrics to evaluate the skill of a model simulation: state variables validation (e.g., Chl*a*, NO_3, O_2, etc…), mass fluxes and process rates validation (e.g., primary production or division rates), and emergent properties validation (e.g., deep chlorophyll maximum, or oxygen minimum zones). In this study we present 22 metrics for the assessment of a model simulation with BGC-Argo data. Most of them evaluate the model state accuracy through the comparison of simulated state variables with BGC-Argo observations in the mixed layer or at fixed depth. In addition, some of the metrics assess the skill of the model in capturing emergent properties. These metrics are associated with the air-sea CO_2 flux, the biological carbon pump, the oceanic pH, and oxygen levels and Oxygen Minimum Zones (OMZs). Further, our validation framework could, in principle, include the second level of assessment metrics (i.e., flux and process). Indeed, recent works demonstrated the feasibility of calculation at basin scale, from BGC-Argo observations, of mass fluxes and process rates, such as primary production, phytoplankton division and accumulation rates (Yang et al., 2021; Mignot et al., 2018), net community production (Plant et al., 2016), and carbon export (Dall'Olmo and Mork, 2014). However, it would be arduous to achieve such estimations on the global BGC-Argo dataset as it requires *ad hoc* calibration that cannot be easily defined. Consequently, the evaluation of simulated process rates with BGC-Argo data is not addressed in this study.

The paper is organised as follow: section 2 presents the data sets used in the study. In section 3, we define the metrics necessary to compare the model to floats' observations. In section 4, we show examples of diagnostic plots for displaying the metrics. In section 5, we discuss metrics relative to optical properties in the water column. Finally, section 6 summarizes and concludes the study.

2. Data

a. BGC-Argo floats observations

The float data were downloaded from the Argo Coriolis Global Data Assembly Centre in France (ftp://ftp.ifremer.fr/argo). The CTD and trajectory data were quality controlled using the standard Argo protocol (Wong et al., 2015). The raw BGC signals were transformed to biogeochemical variables (i.e., O_2, Chl*a*, NO_3, b_{bp}, and pH) and quality-controlled according to international BGC-Argo protocols (Johnson et al., 2018b, a; Schmechtig et al., 2015, 2018; Thierry et al., 2018; Thierry and Bittig, 2018).

In the Argo data-system, the data are available in three data modes: "Real-Time", "Adjusted" and "Delayed" (Bittig et al., 2019). In the "Real-time" mode, the raw data are converted into state variable and an automatic quality-control is applied to "flag" gross outliers. In the "Adjusted" mode, the "Real-time" data receive a calibration adjustment in an automated manner. In the "Delayed" mode, the "Adjusted" data are adjusted and validated by a scientific expert. While the "Real-Time" and "Adjusted" data are considered acceptable for operational application (data assimilation), the "Delayed" mode" is designed for scientific exploitation and represent the highest quality of data with the ultimate goal, when time-series with sufficient duration will have been acquired, to possibly extract climate-related trends. However, for some variables, only a limited fraction of data is accessible in "Delayed-Mode".

Consequently, for each variable, we selected the highest data modes, where at least 80 % of the data are available (see Table 1). Note that this criterion does not apply to O_2, where only delayed mode data were selected in order to generate the pseudo-observations from CANYON-B neural network (see after). We removed data with missing location or time information and flagged as “Bad data” (flag =4). Depending on the parameter and the associated data mode, we also excluded data flagged as “potentially bad data” (flag=3) (see Table 1).

Particulate Organic Carbon (POC) concentrations were derived from b_{bp} observations. First, three consecutive low-pass filters were applied on the vertical profiles of b_{bp} to remove spikes (Briggs et al., 2011): a 2-points running median followed by a 5-points running minimum and 5-points running maximum. Then, the filtered b_{bp} profiles were converted into POC using a POC vs b_{bp} relationship developed for the global ocean (https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-MOB-QUID-015-010.pdf) based on a global database of *in situ* POC and satellite b_{bp} (Evers-King et al., 2017). This relationship, POC= 38687.27* $b_{bp}{}^{0.95}$, developed for global applications, has been shown to outperform regional relationships, applied at global scales. Negative values resulting from this transformation were set to 0.

Finally, we complemented the existing BGC-Argo dataset with pseudo-observations of NO_3, PO_4 , Si, and DIC concentrations as well as pH and pCO_2 using the CANYON-B neural network (Bittig et al., 2018). CANYON-B estimates vertical profiles of nutrients as well as the carbonate system variables from concomitant measurements of floats pressure, temperature, salinity and O_2 qualified in “Delayed“ mode together with the associated geolocalization and date of sampling. The CANYON-B estimates of NO_3 and pH were merged with measured values on the rationale that CANYON-B estimates have RMS errors (NO_3 = 0.7 μmol kg^{-1} , pH = 0.013) (Bittig et al., 2018) that are of the same order of magnitude as those of the BGC-Argo observations errors (NO_3 = 0.5 μmol kg^{-1}, pH = 0.07) (Mignot et al., 2019; Johnson et al., 2017) .

Finally, we verified that the RMS errors of BGC-Argo data (both measured and from CANYON-B estimates) are lower than the RMS difference between the model and BGC-Argo data, so that the comparison of simulated properties with the BGC-Argo data leads to a meaningful evaluation of the model performance. We believe it is reasonable to draw conclusions on the model uncertainty from BGC-Argo data as long as the BGC-Argo errors are much lower than the model-observations RMS difference.

b. CMEMS global BGC Model

The global model simulation used in this study (see Appendix A.1) originates from the Global Ocean hydrodynamic-biogeochemical model, implemented and operated by the Global Monitoring and Forecasting Center of the EU, the Copernicus Marine Environment Monitoring Service (CMEMS). It is based on the coupled NEMO–PISCES model and it is constrained by the assimilation of satellite Chl*a* concentrations. The BGC model is forced offline by daily fields of ocean, sea ice and atmosphere. The ocean and sea ice forcing come from Mercator Ocean global high-resolution ocean model (Lellouche et al., 2018) that assimilates along-track altimeter data, satellite Sea Surface Temperature and Sea-Ice Concentration, and *in situ* temperature and salinity vertical profiles. The BGC model has a 1/4° horizontal resolution, 50 vertical levels (with 22 levels in the upper 100 m, the vertical resolution is 1 m near the surface and decreases to 450 m resolution near the bottom). It produces daily outputs of Chl*a*, NO_3, PO_4, Si, O_2, pH, DIC and Alk, and weekly outputs of POC (resampled offline from weekly to daily frequency through linear interpolation) from 2009 to 2017. Note that the method of linear resampling, while artificially increasing the number of data, could potentially bias the statistical results, especially in regions with poor data coverage. Following the approach of Gali et al. (2021), the POC simulated by the model corresponds to the sum of the two sizes classes of phytoplankton, the small detrital particles and microzooplankton modelled by PISCES. This particular combination of phytoplanktonic and non-phytoplanktonic organisms has been shown match the small POC observed by the floats (Galí et al., 2021). Partial pressures of CO_2 values are calculated offline from the modelled DIC, Alk, temperature and salinity data using the seacarb program for R (https://CRAN.R-project.org/package=seacarb). The Black Sea was not considered in the present analysis because the model solutions are of very poor qualities. Finally, the daily model outputs were collocated in time and the closest to the BGC-Argo floats positions, and they were interpolated to the sampling depth of the float observations. The characteristics of the model are further detailed in the appendix.

3. Metrics

In this section, we present 22 metrics for the assessment of a model simulation with BGC-Argo data. The metrics are associated with the air-sea CO_2 flux, the biological carbon pump, oceanic pH, oxygen levels and Oxygen minimum zones (OMZs). The metrics are described below and summarized in Table 2.

a. Air-sea CO_2 flux

The air-sea CO_2 flux is generally calculated following a bulk formulation (Wanninkhof, 2014), $F_{CO2} = k\alpha(pCO_{2atm} - spCO_2)$, where F_{CO2} is the air-sea CO_2 flux, α is the CO_2 solubility in seawater, k is a gas transfer coefficient that depends on wind speed, $spCO_2$ is the partial pressure of CO_2 at the ocean's surface, and pCO_{2atm} is the partial pressure of CO_2 in the atmosphere. Among the uncertainties affecting the different components of the model CO_2 flux, BGC-Argo data can contribute to estimate that on $spCO_2$. Thus, the validation of pCO_2 plays a critical role to assess the skill of a BGC model in representing correctly the air-sea CO_2 flux.

Here, $spCO_2$ is defined as the average of pCO_2 profile between the surface and the mixed layer depth (MLD). Following De Boyer et al. (2004), the MLD is computed as the depth at which the change in potential density from its value at 10 m exceeded 0.03 kg m^{-3}. We verified that the MLD is correctly represented in the model -- the global bias between the model and the BGC-Argo observations is 0.3 m.

b. Oceanic pH

Ocean acidification is the decrease in oceanic pH due to the absorption of anthropogenic CO_2. The acidification of the ocean is expected to impact primarily the surface oceanic waters as well as the 200-400 m layer (Kwiatkowski et al., 2020). Assessing how models correctly represent oceanic pH at the surface and in the 200-400 m layer is therefore critical if we aim to derive accurate climate projections on acidification. The surface ocean pH (spH) is defined as the average of pH profile between the surface and the base of the mixed layer and the pH in the 200-400 m layer ($pH_{200\text{-}400}$) as the average of pH profile in this layer.

c. Biological carbon pump

The biological carbon pump is the transformation of nutrients and dissolved inorganic carbon into organic carbon in the upper part of the ocean through phytoplankton photosynthesis and the subsequent transfer of this organic material into the deep ocean. The functioning of this pump relies on key pools of nutrients and carbon as well as several processes that control mass fluxes between the pools.

The first level of assessment of a biological carbon pump simulated by a model consists in evaluating the different pools (or state variables) of the pump (Hipsey et al., 2020). In particular, the comparison of simulated surface nutrients (NO_3, PO_4, and Si), DIC, Chla and POC with BGC-Argo observations gives an indirect evaluation of the model capability to capture key processes of the biological carbon pump in the ocean upper layer, such as primary production, respiration, and grazing. A second level assessment would be to directly compare these key processes with measured mass fluxes, but this assessment level is not addressed in this study. The surface nutrients, DIC, Chla and POC (hereinafter denoted sNO_3, sPO_4, sSi, sDIC, sChl and sPOC) are calculated as the average concentrations in the mixed layer.

Similarly, the assessment of the mesopelagic nutrients, DIC and POC concentration (hereinafter indicated with the subscript $_{meso}$) provides an indirect evaluation of the key mesopelagic layer processes, such as export production, respiration, etc. The mesopelagic concentrations are calculated as the depth-averaged concentrations between the base of the mixed layer down to 1000 m.

In stratified systems, a Chl*a* maximum (hereinafter denoted Deep Chlorophyll Maximum, DCM) is formed at the base of the euphotic layer (Barbieux et al., 2019; Cullen, 2015; Letelier et al., 2004; Mignot et al., 2014, 2011). It has been suggested that the DCM plays an important role in the synthesis of organic carbon by phytoplankton (Macías et al., 2014). DCMs are therefore important features to be assessed in BGC models with respect to processes involved in the biological carbon pump processes such as the primary production, however the DCM layer generally escapes detection by remote sensing. Furthermore, DCM is also an emergent feature that develops in response to complex physical and biogeochemical interactions (Cullen, 2015). Thus, its evaluation provides critical information regarding the accuracy of the model in capturing complex patterns of key ecosystem processes. The depth and magnitude of DCM (H_{dcm} and Chl_{dcm}) are helpful metrics for the assessment of DCM dynamics. The depth of the DCM is calculated as the depth where the maximum of Chl*a* occurs in the profile with the criterion that H_{dcm} should be deeper than the MLD. The magnitude of the DCM is computed at the value at H_{dcm}.

The vertical supply of NO_3 to the surface layers is a critical process of the biological carbon pump as NO_3 is often depleted in the surface layers and is a limiting factor for phytoplankton growth in most oceanic regions. This NO_3 vertical supply depends, among other factors, on the vertical gradient of NO_3 (the nitracline), and, in particular, on its depth (the nitracline depth) (Cermeno et al., 2008; Omand and Mahadevan, 2015). Therefore, the comparison of the simulated nitracline depth with BGC-Argo observations allows for an indirect assessment of the model quality in reproducing vertical fluxes of NO_3. Following previous studies (Cermeno et al., 2008; Lavigne et al., 2013; Richardson and Bendtsen, 2019), the depth of the nitracline corresponds to the first depth where NO_3 is detected. The detection threshold is set to 1 µmol kg^{-1}, which corresponds to an upper estimate of BGC-Argo NO_3 data accuracy (Johnson et al., 2017; Mignot et al., 2019).

d. Oxygen levels and oxygen minimum zones

Oxygens levels in the global and coastal waters have declined over the whole water column over the past decades (Schmidtko et al., 2017) and OMZs are expanding (Stramma et al., 2008). Assessing how models correctly represent ocean oxygen levels as well as the OMZs is therefore critical to monitor their change over time. Similarly to DCMs, the assessment of OMZs is also informative on how the model simulates emergent dynamics as OMZs originate from complex physical and biogeochemical interactions (Paulmier and Ruiz-Pino, 2009). We evaluate oxygen levels in 3 layers, at the surface, at 300 m and at 1000 m. The surface O_2 (sO_2), important for the air-sea O_2 flux, is defined as the average of O_2 profile in the mixed layer. The oxygen at 300 m ($O_{2\ 300}$), a depth where large areas of the global ocean have very low O_2 (Breitburg et al., 2018), is defined as the average of O_2 profile between 250 and 300

m. The deep oxygen content, ($O_{2\ 1000}$), is defined as the average of O_2 profile between 950 and 1000 m. Finally, to characterize the OMZs, we evaluate the depth (H_{O2min}) and concentration (O_{2min}) of O_2 minimums. O_2 level lower than 80 µmol kg^{-1} are used to characterize OMZs (Schmidtko et al., 2017).

4. Diagnostic plots to display the BGC-Argo based metrics

Based upon the existing literature (e.g., Aumont et al., 2015; Cossarini et al., 2019; Doney et al., 2009; Dutkiewicz et al., 2015; Gutknecht et al., 2019; Salon et al., 2019; Séférian et al., 2013; Terzić et al., 2019), we propose 4 graphical representations that can be used to display the novel validation metrics and to assess the skill of a model in reproducing a particular process or variable: Taylor diagrams, scatterplots, spatial maps, and time series.

a. Taylor diagram

Taylor diagrams are useful to display simultaneously information on model-data skill for a suite of metrics (Taylor, 2001). These diagrams combine the Pearson correlation coefficient (r), root-mean-square difference (RMSD) and the model standard deviation (SD). In order to represent all metrics with different units into a single diagram, we use a normalized Taylor diagram (RMSD and the model SD are divided by the SD of the observations). In the diagram, the Pearson correlation coefficient between the model and the observations is related to the azimuthal angle. The normalized SDs are proportional to the radial distances from the origin. The observational reference is indicated along the x-axis and corresponds to the normalized SD and r =1. Finally, the normalized RMSD is proportional to the distance from the observational difference.

b. Scatter/Density plots

In validation exercises, scatter plots are useful to identify relationships between the predicted and observed values. It is common to add a least squares regression line to quantify the strength of the linear relationship between the observed and predicted values. In those cases, when a large amount of data points has to be plotted (like in our study), the points overlap to a degree where it can be difficult to distinguish the relationship between the variables. To overcome this, scatter plots are displayed as density plots, where each axis is divided in several bins while the colour within each bin indicates the number of points.

c. Spatial maps

Spatial maps draw attention to the spatial distribution of a given metric. The maps are handy to determine if the model is skilled in reproducing global patterns, spatial gradients, and basins inter-difference. It is also helpful to display the BIAS and RMSD between predicted and observed values on a spatial map to quickly determine regions where the model uncertainty is the highest. Depending on the context, the comparison between the model and the observation can be performed either on a climatological level, or for a specific period (year, month, etc ..). In our case, the scarcity of observations imposes us to display all data (from 2009 to 2017; the period of analysis of the model simulation) in a climatological way if we want to highlight large scale patterns. To do so, the metrics from 2009 to 2017 are averaged in 4°x4° bins, bins excluding those with less than 4 points. The 4° distance is an upper estimate of the autocorrelation length scales for O_2, nutrients, and pCO_2 (comprised between 300 and 400 km) between 20° and 40° of latitude in both hemispheres (Biogeochemical-Argo Planning Group, 2016). We also computed the BIAS and RMSD within each bin. Standard deviation can also be displayed on spatial maps as an indicator of the model skill in properly reproducing variability scales. For clarity, it is not shown in this study.

d. Seasonal time-series

Taylor diagrams, scatter plots and spatial maps are powerful diagnostics plots to evaluate the global skills of a model but understanding the causes of difference remains somewhat limited with these diagrams. Rather, the comparative analysis of seasonal time-series of multiple metrics and their inter-relationships is a powerful tool to highlight and to understand BGC processes. This is especially true for the biological carbon pump that has a strong seasonal variability due to the seasonal variation in sunlight, surface heating and surface wind (Williams and Follows, 2011). As a matter of fact, the analysis of seasonal dynamics in nutrients as well as in phyto- and zoo- plankton has a rich history for the development of

BGC models (Evans and Parslow, 1985; Riley, 1946). In addition to the time series of metrics, we also display normalized skill scores such as percent BIAS and RMSD as a function of season in order to combine quantitative metrics with visual comparison.

5. Results: Application to CMEMS global model

Examples of the diagnostic plots described in section 4 in combination with the metrics defined in Section 3 are shown. The objective of this section is to illustrate the opportunities offered by the BGC-Argo data for evaluating global BGC model solutions, rather than to provide a full evaluation of the CMEMS global model. Consequently, for each diagnostic plot, we only present one detailed example. The density plots and spatial maps for all metrics are displayed in the Appendix section (Fig. A1-A44).

a. Taylor diagram

The CMEMS global model skill is summarized in the normalized Taylor diagram (Fig. 2) and Table 3. The oxygen levels metrics (sO_2, $O_{2\ 300}$, $O_{2\ 1000}$), $pH_{200-400}$, the average nutrients and DIC concentrations in the mixed layer and in the mesopelagic layer are particularly well represented in the model. The correlation coefficients are greater than 0.95, the predicted SDs are close the observed SDs and the normalized RMSDs are lower than 0.4. The OMZs as well as the depths of DCM and nitracline are reasonably well represented in the model, with r > 0.9 (OMZs) and r > 0.8 (for H_{nit} and H_{dcm}) and normalized RMSDs <0.6. The variability in the predicted O_{2min} is however larger than the observed ones. Finally, the POC concentrations, the Chl*a* in the mixed layer and at the DCM as well as $spCO_2$ and spH are the worst predicted metrics. The normalised RMSDs are greater than 0.7-0.8, and r is between 0.3 and 0.7.

The fact that surface nutrients are well represented in the model suggests that the model captures the combination of process rates that drive nutrients dynamics. Some of these process rates drive both the nutrients, Chl*a* and POC dynamics, but there are also rates that are specific to each state variable. This probably explains why Chl*a* and POC are not performing while the surface nutrients are well simulated. However, it must be recognised that without a direct assessment of the individual rates, we cannot verify this hypothesis.

The representation of all metrics into a single Taylor diagram allows to rapidly evaluate the strengths and the weaknesses of a model simulation. For instance, the CMEMS global model is skilled in reproducing oxygen levels and the cycling of nutrients and DIC, but the representation of Chl*a*, POC, $spCO_2$ and spH needs to be improved.

b. Scatter/Density plots

The density plots for all metrics are displayed in the Appendix section (Fig. A1-A22). Here, we detail only the density plot for O_{2min} to illustrate the potential of such representations.

Figure 3 shows the comparison between the observed and predicted O_{2min} values. The regression line, the slope, and the intercept as well the coefficient of determination (R^2) are indicated. Overall, the model and the float O_{2min} are in good agreement with a slope close to 1 and R^2 close to 0.8. There is however a positive offset of ~11 µmol kg^{-1} across all O_{2min} values suggesting that the modelled OMZs are on average too much oxygenated by a constant value. It is worth noting that the scatter around the regression line is larger for O_{2min} > 50 µmol kg^{-1}, which corresponds to the Atlantic OMZ around Cap Verde (Fig. A43). This suggests that the uncertainty in this OMZ is particularly high, as confirmed in Fig. A43.

c. Spatial maps

The spatial maps for all metrics are displayed in the Appendix section (Fig. A23-A44), while we detail hereafter the spatial distribution of sChl.

Figure 4 shows the spatial distribution of sChl estimated from the BGC-Argo floats (Fig. 4a), the model (Fig. 4b), the BIAS (Fig. 4c) and the RMSD (Fig. 4d). As already noticed in Fig. 1, the density of sChl observations is satisfactory for high latitude regions (latitudes > 50° N and S) whereas it is poor in subtropical gyres and the Equatorial band. Nevertheless, large scale patterns in sChl are still distinguishable in Fig. 1a, especially the juxtaposition of the high-latitudes-high- sChl regions with the low-latitudes-low- sChl regions. The model (Fig. 4b) exhibits large-scale, coherent patterns. However, the model tends to be lower than the BGC-Argo observations in the high-latitudes region and higher in the subtropical gyres (Fig. 4c). The RMS difference between the predicted and the observed values seems to be quite uniform, suggesting the uncertainty in model sChl is fairly constant in all oceanic basins.

d. Seasonal time-series

Two examples of BGC-Argo float seasonal time-series compared to the same time-series simulated by the model along the float trajectory are presented in Figs. 5 and 6. The two figures present a case study in the North Atlantic during the "spring bloom" and a case study in the South Pacific subtropical gyre.

Figure 5 compares the seasonal time series of MLD, sChl, sNO_3, sSi and sPO_4 derived from the BGC-Argo floats observations (blue) and from the model simulation (yellow). To avoid relying only on visual inspection, the percent BIAS and percent RMSD are also represented for each metrics and for each season.

The seasonal cycles of MLD, sChl and nutrients are typical of the North Atlantic bloom dynamics (Dale et al., 1999; Mignot et al., 2018). The temporal dynamics of sChl and nutrients are well approximated by the model with the timings of minima, maxima and the onset of the bloom being correctly represented. The winter- sChl -minimum and winter-nutrients-maxima are also properly estimated by the model (Figs. 5g and h). However, the summer- sChl -maximum is underestimated and the summer- sNO_3 -minimum and summer- sPO_4 -minimum are overestimated (Fig. 5g). This is coherent with the negative BIASs observed in the spatial map of sChl in the North Atlantic (Fig. 4) and the positive BIAS in the spatial map of sNO_3 and sPO_4 in the North Atlantic (Figs. A27 and A28).

Figure 6 shows similar time series than Fig. 5 but for an oligotrophic environment in the South Pacific subtropical gyre. The time series of H_{DCM} and Chl_{DCM} are also shown as this gyre is characterized by a seasonal and permanent DCM (Mignot et al., 2011). The model correctly represents the seasonal cycle of sChl, H_{DCM} and Chl_{DCM}, which are characteristic of this region. The average percent RMSD for these three metrics is 17 %, 12 % and 16 % respectively. The more stable time series of sSi and sPO_4 are also well simulated by the model; the average percent RMSD being 19 % and 11 % respectively. Finally, sNO_3 are constantly underestimated by the model by an average negative BIAS of roughly 0.25 µmol kg^{-1}.

6. Perspectives: metrics relative to ocean optical properties

BGC-Argo floats equipped with sensors measuring the downward planar irradiance are essential observations to evaluate the performance of recently-developed BGC models that resolve the spectral and directional properties of the underwater light field. For several years, the number of BGC models coupled with a multispectral light module has been steadily increasing (Baird et al., 2016; Dutkiewicz et al., 2015; Gregg and Rousseaux, 2016; Lazzari et al., 2020; Skákala et al., 2020). Such models require dedicated observations and metrics to evaluate their skill in representing the ocean's optical properties of the ocean. Diffuse attenuation coefficient for downwelling irradiance (K_d) is one of the most common properties to characterise the optical state of the ocean (Sosik, 2008). Values of K_d can be derived at three different wavelengths (380, 412, 490 nm) from the BGC-Ago floats observations. This metric also provides information about the constituents of seawater (Organelli 2017) (phytoplankton for K_d at 490 nm and coloured dissolved organic carbon for K_d at 380 nm and 412 nm) and is complementary to Chl*a* measurements for the assessment of the modelled phytoplankton dynamics.

BGC-Argo floats equipped with optical sensors are available on the global ocean, but the global model used in this study does not resolve the spectral and directional properties of the underwater light field. Therefore, to show the potentiality of such comparison, we use a model of the Mediterranean Sea equipped with a multispectral light module (Lazzari et al., 2020) (Appendix A.2). The spatial distribution of K_d at 490 nm in the first optical depth estimated from the BGC-Argo floats and from the Mediterranean Sea model are shown in Fig. 7. The BGC-Argo estimated K_d at 490 nm exhibits a basin-scale pattern, with high values in the North-Western Mediterranean Sea and lower values in the Eastern Mediterranean Sea, consistent with the spatial distribution of surface Chl*a* in the Mediterranean Sea (Bosc et al., 2004). The model is able to reproduce the large-scale pattern of K_d at 490 nm, but it tends to underestimate K_d at 490 nm in the North-Western Mediterranean Sea; area where the RMSD is also the highest. The annual cycle of phytoplankton being largely influenced by a spring bloom in this region (Bosc et al., 2004; D'Ortenzio et al., 2014), we can speculate that the underestimation of K_d at 490 nm highlights a possible misrepresentation of the spring bloom in the model that yields to lower phytoplankton and Chl*a* concentrations. The comparison exercise performed in the Mediterranean Sea shows the added value of BGC-Argo optical data for the assessment of biogeochemical model dynamics at the global scale.

7. Conclusion

Biogeochemical ocean models are powerful tools to monitor changes in marine ecosystems and ecosystem health due to human activities, make climate projections and help developing better strategies for mitigation. However, these models are subject to flaws and require rigorous validation processes to test their predictive skills. The model's evaluations have long been damped by the lack of *in situ* observations, which has certainly slowed the development and the improvement of BGC models. The number of observations collected by the BGC-Argo program is now greater than any other *in situ* data set (Claustre et al., 2020) and thus, offers new opportunities for the validation of BGC models.

In this study, we use the global data set of BGC-Argo observations to validate a state-of-the-art BGC model simulation. Our aim was to demonstrate the invaluable opportunities offered by the BGC-Argo observations for evaluating global BGC model solutions. To ease the comparison between model and observations at global scale, we proposed 22 assessment metrics, based on the model evaluation framework developed by Hipsey et al. (Hipsey et al., 2020). These metrics either evaluate the model state accuracy or the skill of the model in capturing emergent properties. We did not propose BGC-Argo-based phenology metrics (Gittings et al., 2019), because the numbers of observations per month and per bin is still presently too low to derive such robust metrics. We suggested 4 diagnostic plots, which we believe are particularly suitable for displaying the metrics in support of the identification of model-data difference and subsequent analysis of model representativity. We also discuss the promising avenue of BGC-Argo-based metrics relative to optical properties in the ocean for the validation of the new generation of BGC model equipped with a multispectral light module.

We assumed that the differences between the observed and predicted BGC values were only attributable to the BGC model, PISCES. However, BGC models are coupled to ocean general circulation systems and the quality of the BGC predictions strongly depends on the accuracy of the physical properties that control the BGC state variables. In our case, the dynamical component has been extensively validated (Lellouche et al., 2018, 2013), and correctly represented variables that are constrained by observations (e. g., temperature and salinity). However, unconstrained variables in the physical system (e.g., vertical velocities) can generate unrealistic biases in various biogeochemical variables, especially in the Equatorial Belt area (Fennel et al., 2019; Park et al., 2018).

We have restricted the number of diagnostic plots as well the statistical indices to the ones that are most commonly used in the modelling community. More complex statistical indicators (Stow et al., 2009) can be computed with the proposed metrics, depending on the context and the skill level necessary. Likewise, similar or more elaborate diagrams can also be used, such as Target diagram (Salon et al., 2019), zonal mean diagrams (Doney et al., 2009), or interannual time series (Doney et al., 2009).

The comparison between BGC-Argo data and model simulations is not only beneficial for the modelling community but also for the BGC-Argo community. Observation System Simulation Experiments (OSSEs) are generally used to inform, *a priori*, observing network design (Ford, 2020). Here, we showed that the spatial maps of model-observations comparison are also informative *a posteriori,* with respect to the network design, as they highlight sensitive areas where BGC-Argo observations are critical and where sustained BGC-Argo observations are required to better constrain the model. These maps correspond to the regions where the model uncertainty (see RMSD spatial maps in Figs. A22-A44) is the highest, i.e., the Equatorial belt with respect to the carbonate system variables, the Southern Ocean with respect to the nutrients and the DCM variables, and the western boundary currents and OMZs with respect to oxygen.

Tables

Table 1. Data mode and QC flags of the BGC-Argo observations used in this study. In the Argo data-system, the data are available in three data modes, “Real-Time”, ”Adjusted” and ”Delayed”. See section 2a for a brief description of each data mode. The flags “3” and “4” refers to “potentially bad data “ and “bad data”, respectively. See also Bittig et al. (2019), for a more detailed description of Argo data modes and flags.

Parameter	Data mode	Data mode of associated pressure, temperature and salinity profiles	QC flags
Chl*a*	Adjusted and Delayed	Real time, Adjusted and Delayed	• Real time: All flags except 4 • Adjusted or Delayed: All flags except 3 and 4
O_2	Delayed	Delayed	• All flags except 3 and 4
NO_3	Adjusted and Delayed	Real time, Adjusted and Delayed	• Real time: All flags except 4 • Adjusted or Delayed: All flags except 3 and 4
pH	Adjusted and Delayed	Real time, Adjusted and Delayed	• Real time: All flags except 4 • Adjusted or Delayed: All flags except 3 and 4
b_{bp}	Real time and Delayed	Real time, Adjusted and Delayed	• Real time: All flags except 4 • Adjusted or Delayed (P,T,S): All flags except 3 and 4

- Adjusted or Delayed (b_{bp}): All flags 4

Table 2. Assessment metrics used to assess the model simulation with BGC-Argo data . For each metric, the level of assessment, as described in Hipsey et al. (2020) is also indicated.

Process	Metric	Definition	units	Assessment level
Air-sea $CO_{2\ flux}$	$spCO_2$	Depth-averaged pCO_2 in the mixed layer	µatm	State variable
Oceanic pH	spH	Depth-averaged pH in the mixed layer	total	State variable
	$pH_{200\text{-}400}$	Depth-averaged pH in the 200-400 m layer	total	State variable
Biological carbon pump	sChl	Depth-averaged Chl*a* in the mixed layer	mg m^{-3}	State variable
	sNO_3	Depth-averaged NO_3 in the mixed layer	µmol kg^{-1}	State variable
	sPO_4	Depth-averaged PO_4 in the mixed layer	µmol kg^{-1}	State variable
	sSi	Depth-averaged Si in the mixed layer	µmol kg^{-1}	State variable
	sDIC	Depth-averaged DIC in the mixed layer	µmol kg^{-1}	State variable
	$NO_{3\ meso}$	Depth-averaged NO_3 in the mesopelagic layer	µmol kg^{-1}	State variable
	$PO_{4\ meso}$	Depth-averaged PO_4 in the mesopelagic layer	µmol kg^{-1}	State variable
	Si_{meso}	Depth-averaged Si in the mesopelagic layer	µmol kg^{-1}	State variable

	DIC_{meso}	Depth-averaged DIC in the mesopelagic layer	µmol kg^{-1}	State variable
	sPOC	Depth-averaged POC in the mixed layer	mg m^{-3}	State variable
	POC_{meso}	Depth-averaged POC in the mesopelagic layer	mg m^{-3}	State variable
	Chl_{DCM}	Magnitude of DCM	mg m^{-3}	Emergent property
	H_{DCM}	Depth of DCM	m	Emergent property
	H_{nit}	Depth of nitracline	m	Emergent property
Oxygen levels and OMZs	sO_2	Depth-averaged O_2 in the lixed layer	µmol kg^{-1}	State variable
	$O_{2\ 300}$	O_2 at 300 m	µmol kg^{-1}	State variable
	$O_{2\ 1000}$	O_2 at 1000 m	µmol kg^{-1}	State variable
	O_{2min}	value of O_2 minimum	µmol kg^{-1}	Emergent property
	H_{O2min}	Depth of O_2 minimum	m	Emergent property

Table 3. Global model skill assessment. The assessment metrics are defined in Table 2.

Metric	BGC-Argo mean	BGC-Argo SD	Model mean	Model SD	Bias	RMSD	Pearson correlation coefficient
$spCO_2$ (μatm)	374	29	370	28	-5	29	0.50
spH (total)	8.056	0.030	8.058	0.028	0.001	0.028	0.54
$pH_{200\text{-}400}$ (total)	7.933	0.125	7.949	0.114	0.016	0.038	0.96
log_{10}(sChl (mg m^{-3}))	-0.7	0.7	-0.6	0.4	0.1	0.5	0.69
sNO_3 (μmol kg^{-1})	9.4	10.1	9.1	9.6	-0.3	2.5	0.97
sPO_4 (μmol kg^{-1})	0.75	0.64	0.81	0.62	0.07	0.15	0.98
sSi (μmol kg^{-1})	8.5	14.6	10.5	14.6	2.0	4.7	0.96
sDIC (μmol kg^{-1})	2077.0	69.7	2077.5	65.8	0.4	19.0	0.96
$NO_{3\ meso}$ (μmol kg^{-1})	20.6	9.6	19.9	8.7	-0.8	2.2	0.98
$PO_{4\ meso}$ (μmol kg^{-1})	1.50	0.7	1.5	0.6	0.0	0.1	0.98
Si_{meso} (μmol kg^{-1})	30.0	28.8	30.7	26.9	0.7	4.6	0.99
DIC_{meso} (μmol kg^{-1})	2170.5	57.2	2161.1	53.7	-9.4	15.8	0.98
log_{10}(sPOC (mg m^{-3})	1.73	0.29	1.55	0.27	-0.18	0.32	0.57
log_{10}(POC_{meso} (mg m^{-3})	1.41	0.23	0.99	0.32	-0.42	0.53	0.35
Chl_{DCM} (m)	-0.3	0.4	-0.4	0.2	-0.1	0.3	0.55
H_{DCM} (m)	79	36	75	36	-3	21	0.84
H_{nit} (m)	43	63	41	57	-2	27	0.89

sO_2 (μmol kg^{-1})	266.9	47.8	267.3	47.9	0.4	12.8	0.96
$O_{2\ 300}$ (μmol kg^{-1})	208.3	68.8	211.4	61.9	3.1	18.9	0.96
O_{2min} (μmol kg^{-1})	208.3	68.8	211.4	61.9	3.1	18.9	0.96
H_{O2min} (m)	725	362	813	332	87	165	0.92

Figures

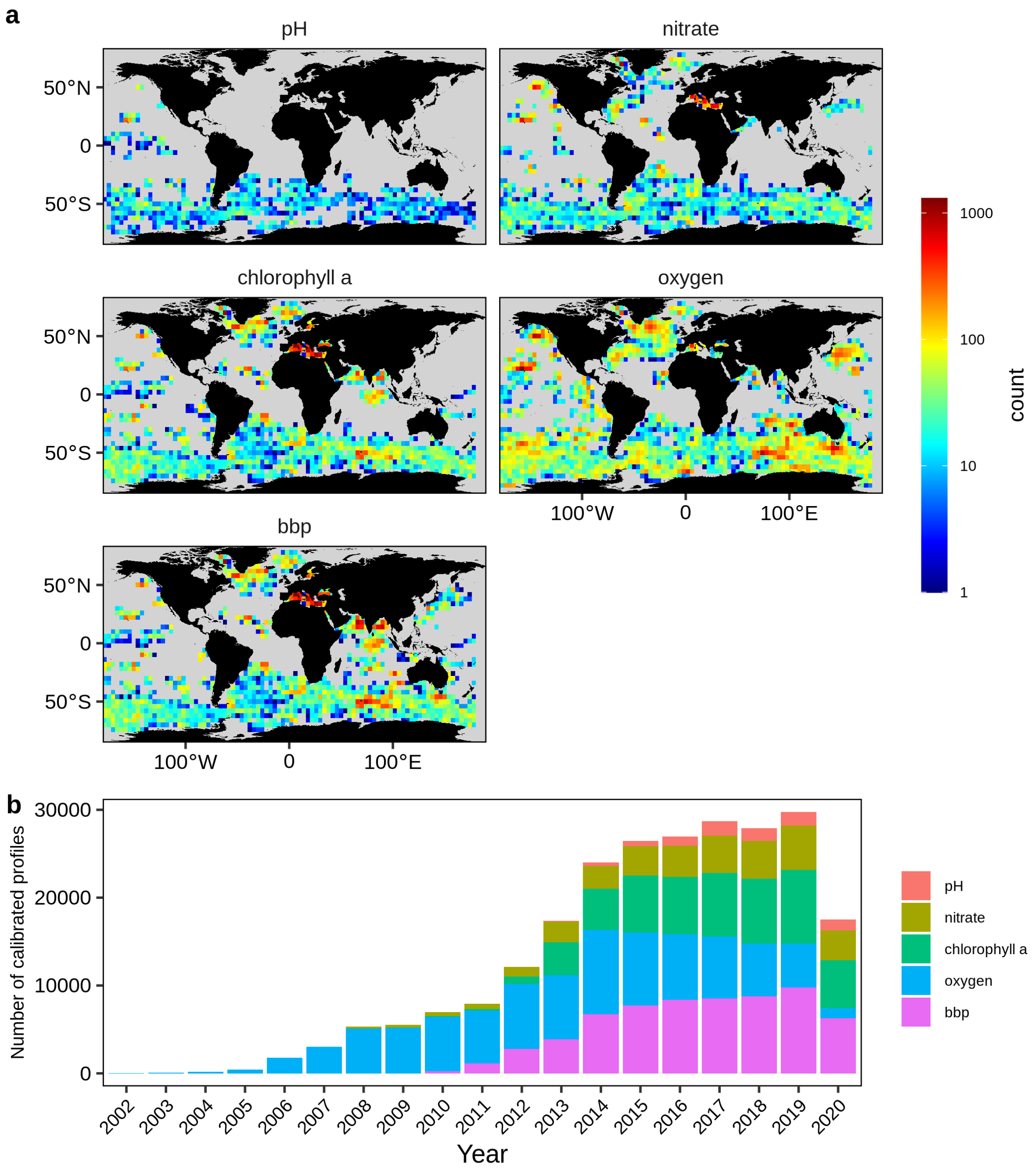


Figure 1. Spatial and temporal coverage of quality-controlled BGC-Argo pH, NO_3^-, Chl*a,* O_2, and b_{bp} profiles. **(a)** Number of quality-controlled profiles for the entire period per 4°x4° bin. **(b)** Number of quality-controlled profiles per year.

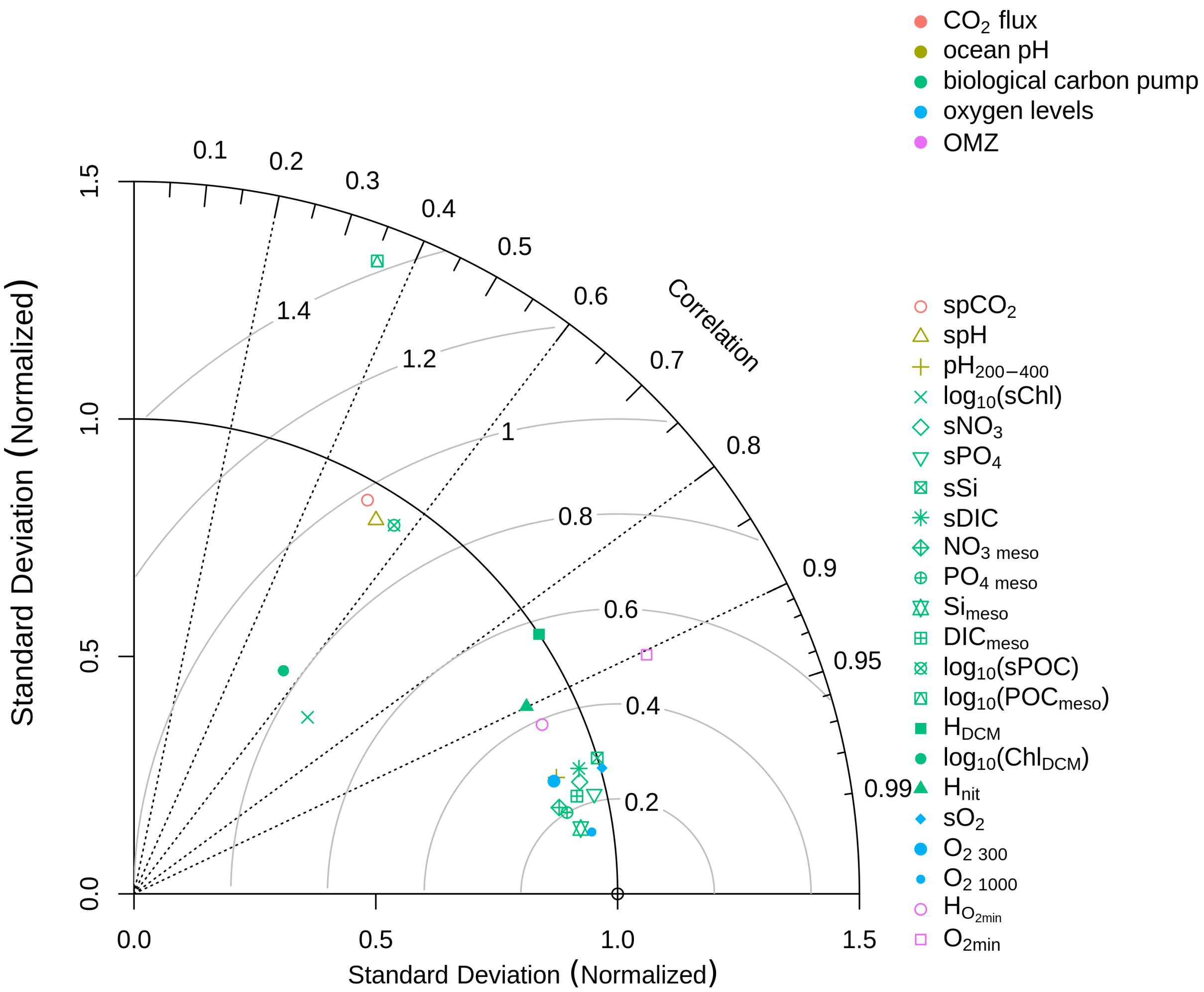


Figure 2. Comparison of BGC-Argo floats' observations and model values for all metrics using Taylor diagram. The symbols correspond to the metrics and the colours represent the BGC processes with which they are associated. Note that the metrics calculated from the float pH and NO_3 used both the direct observations of the floats and as well as the estimations from CANYON-B. The metrics related to Chl*a* and POC, namely sChl, Chl_{DCM}, sPOC, POC_{meso} were log_{10}-transformed because they cover several orders of magnitude and they are lognormally distributed. Observed DCMs and nitracline deeper than 250 m are not included.

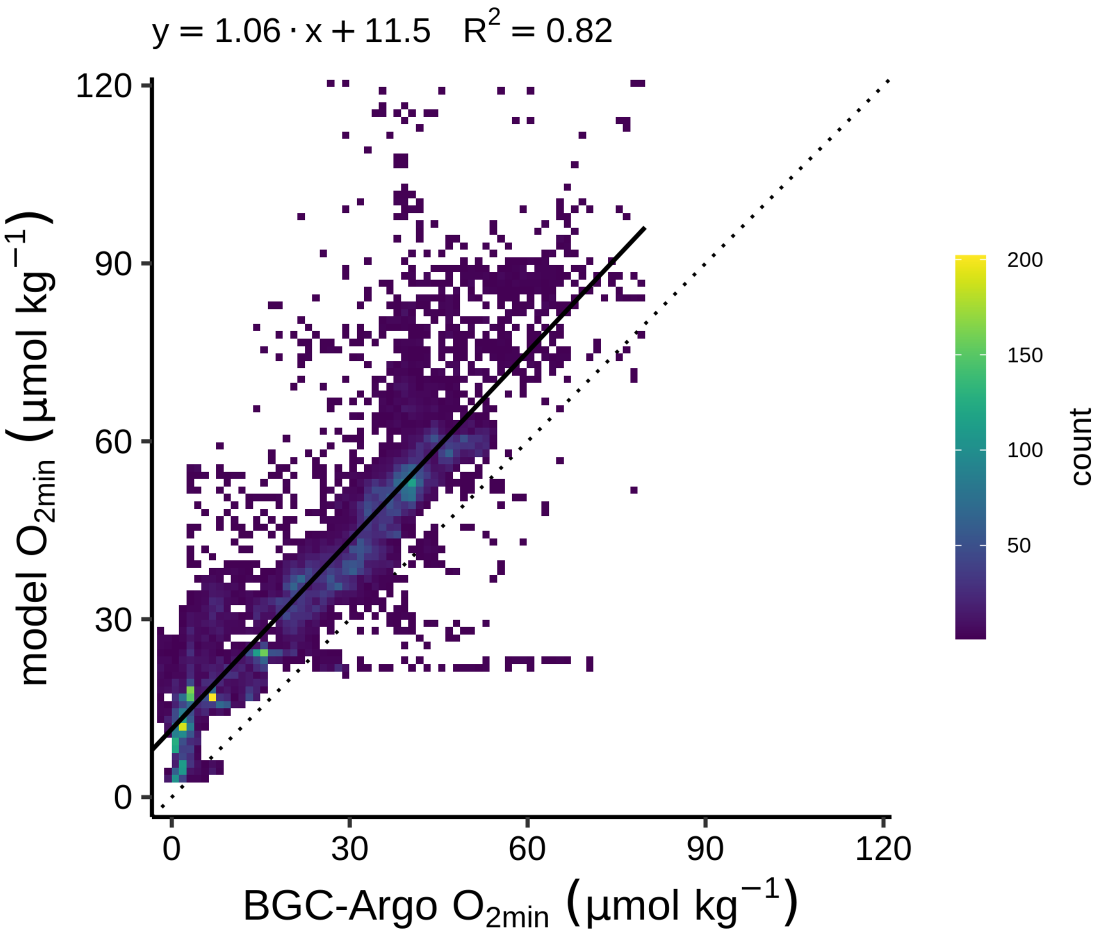


Figure 3. Density plots of BGC-Argo floats' observations and model O_{2min} . Each axis is divided in 100 bins and the colour represents the number of points in each bin. The dashed line represents the 1:1 line. The plain line represents the linear regression line between the two data sets**.** The coefficients of the linear regression line (gain and offset) as well the coefficient of determination (R^2) are indicated on the top of the plot.

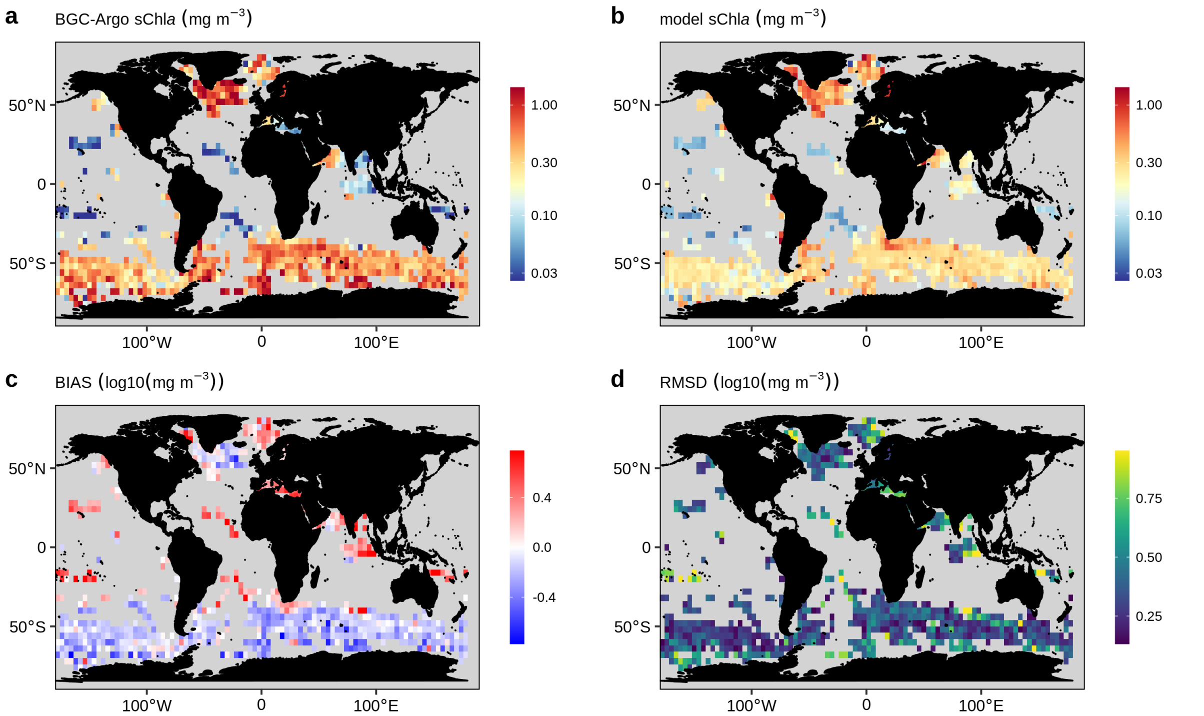


Figure 4. Spatial distribution maps of BGC-Argo floats' observations of sChl **(a),** model sChl **(b)**, the BIAS **(c)** and the RMSD **(d)**. The data are averaged in 4°x4° bins. Bins containing less than 4 points are excluded. The BIAS and RMSD are computed on the $\log_{10}$-transformed data to account that sChl covers several orders of magnitude and is lognormally distributed (Campbell, 1995).

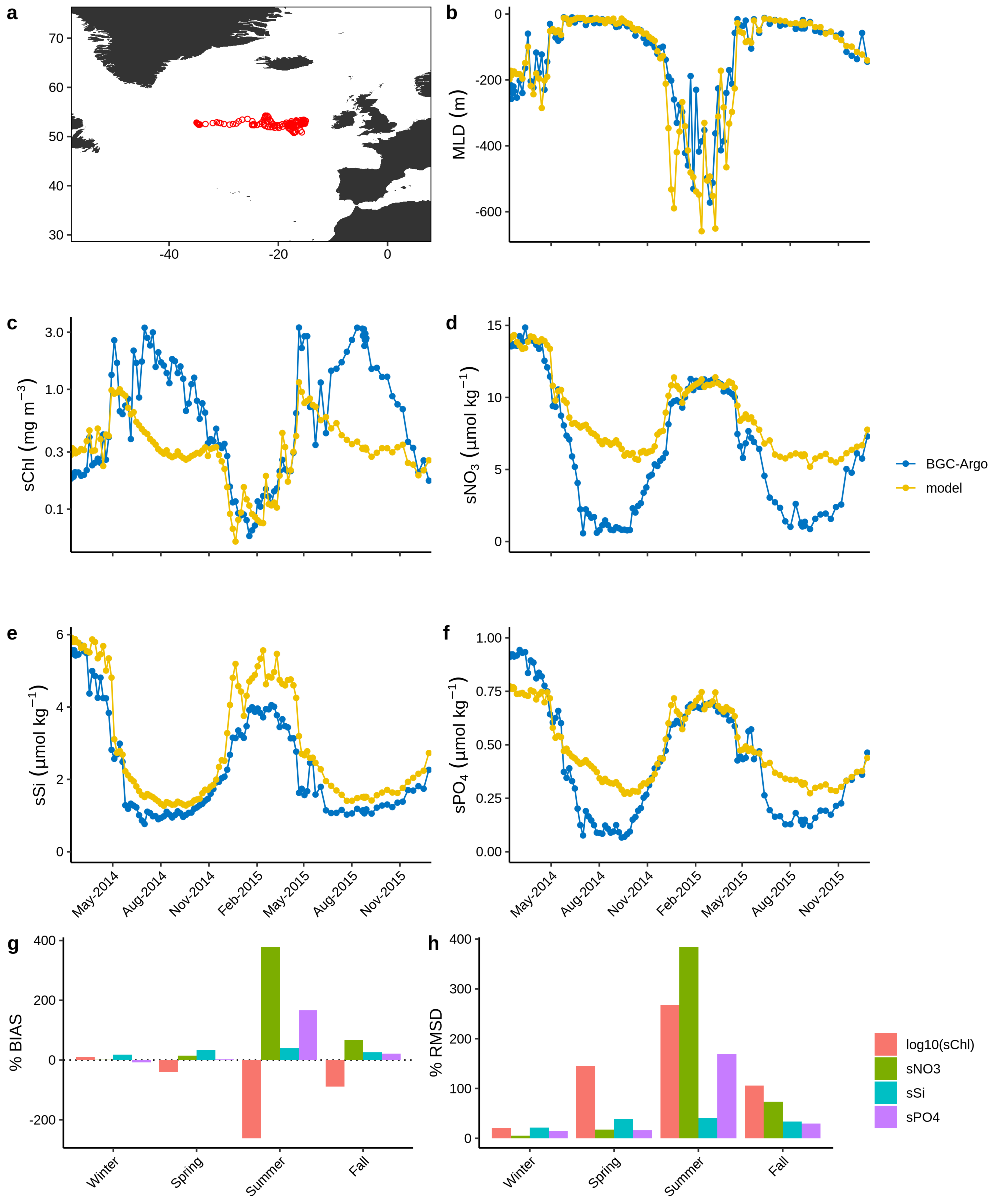


Figure 5. (**a**) Float trajectory of a BGC-Argo float located in the North Atlantic (WMO number: 5904479). Time series of **(b)**, mixed layer depth, **(c)**, sChl, **(d)**, sNO_3, **(c)**, sSi , **(f),** sPO_4 derived from the BGC-Argo floats observations (blue) and from the model simulation (yellow). **(g),** Percent BIAS $\left(100 \times \frac{\frac{1}{N}\sum_{i=1}^{N}(model_i - obs_i)}{|\overline{obs}|}\right)$ and **(h),** percent RMSD $\left(100 \times \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(model_i - obs_i)^2}}{|\overline{obs}|}\right)$ as a function of season. The float sChl and sNO_3 are calculated from the direct observations of the floats, whereas the float sSi and sPO_4 result from CANYON-B predictions.

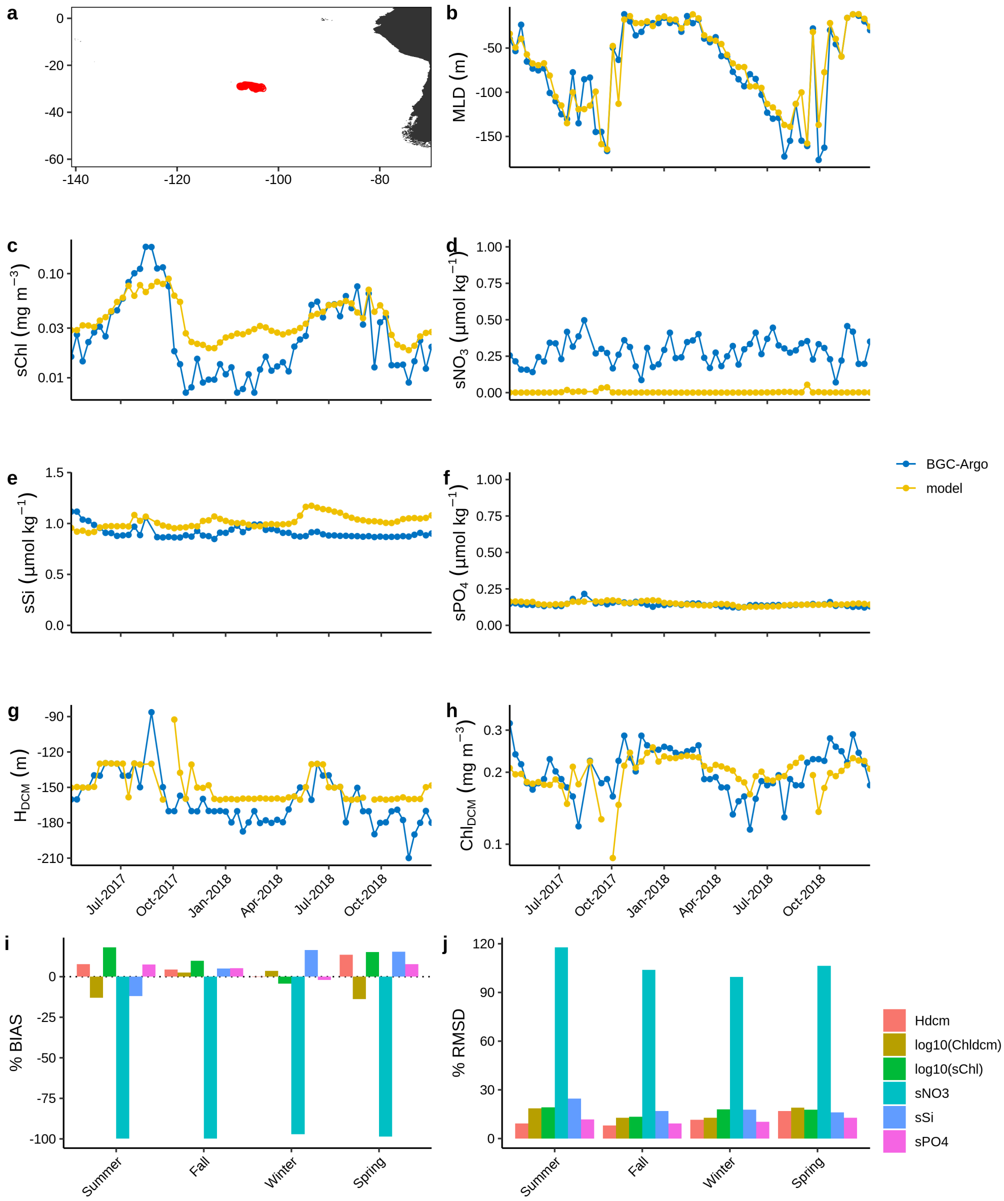


Figure 6. a) Float trajectory of a BGC-Argo float located in the South Pacific subtropical gyre (WMO number: 5904479). Time series of **(b)**, mixed layer depth, **(c)**, sChl, **(d)**, sNO_3, **(c)**, sSi , **(f),** sPO_4, **(g)**, H_{DCM}, **(h),** Chl_{DCM} derived from the BGC-Argo floats observations (blue) and from the model simulation (yellow). Time series of **(i),** percent BIAS $\left(100 \times \frac{\frac{1}{N}\sum_{i=1}^{N}(model_i - obs_i)}{|\overline{obs}|}\right)$ and **(j)** percent RMSD $\left(100 \times \frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(model_i - obs_i)^2}}{|\overline{obs}|}\right)$. The float sChl, H_{DCM}, Chl_{DCM} and sNO_3 are calculated from the direct observations of the floats, whereas the float sSi and sPO_4 result from CANYON-B predictions.

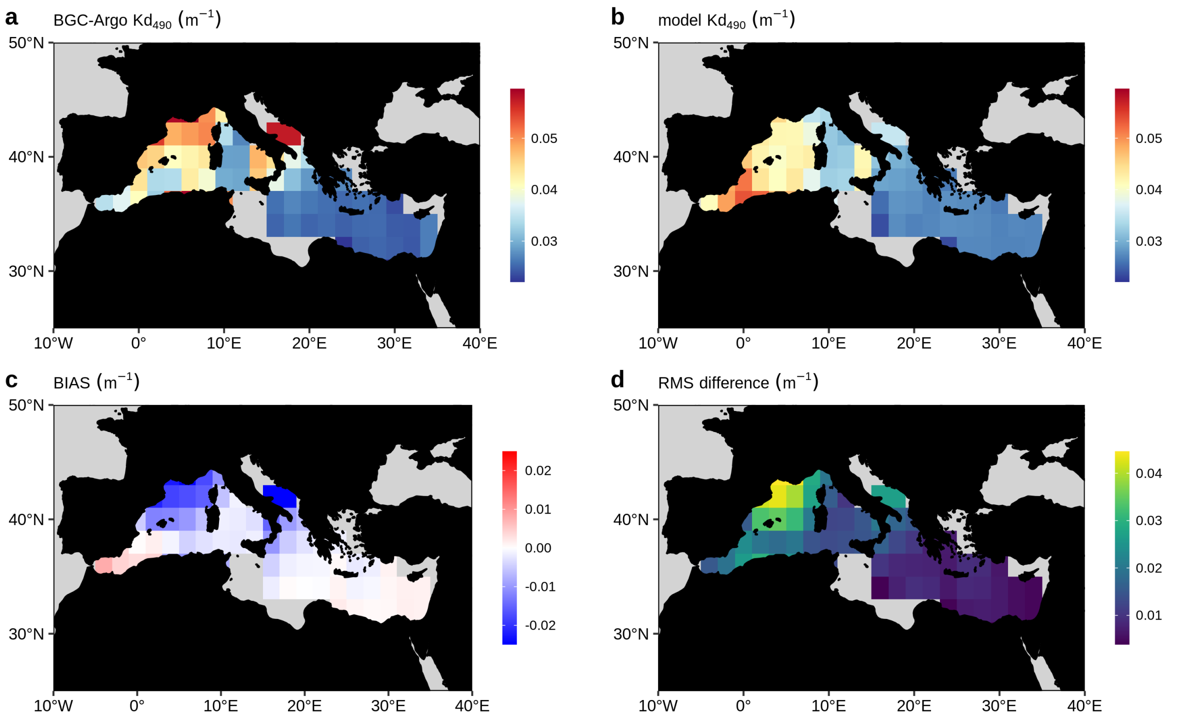


Figure. 7 . Spatial distribution maps of BGC-Argo floats' observations K_d at 490 nm **(a),** modelled K_d at 490 nm from the Mediterranean BGC model **(b)**, the BIAS **(c)** and the RMSD **(d)**. The data are averaged in 2°x2° bins. Bins containing less than 4 points are excluded.

Appendix

A.1 The CMEMS global hydrodynamic-biogeochemical model

The model used in this study features the offline coupled NEMO–PISCES model, with a 1/4° horizontal resolution 50 vertical levels (with 22 levels in the upper 100 m, the vertical resolution is 1m near the surface and decreases to 450m resolution near the bottom) and daily temporal resolution, covering the period from 2009 to 2017.

The biogeochemical model PISCES v2 (Aumont et al., 2015) is a model of intermediate complexity designed for global ocean applications, and is part of NEMO modelling platform. It features 24 prognostic variables and includes five nutrients that limit phytoplankton growth (nitrate, ammonium, phosphate, silicate and iron) and four living compartments: two phytoplankton size classes (nanophytoplankton and diatoms, resp. small and large) and two zooplankton size classes (microzooplankton and mesozooplankton, resp. small and large); the bacterial pool is not explicitly modelled. PISCES distinguishes three non-living detrital pools for organic carbon, particles of calcium carbonate and biogenic silicate. Additionally, the model simulates the carbonate system and dissolved oxygen. PISCES has been successfully used in a variety of biogeochemical studies, both at regional and global scale (Bopp et al., 2005; Gehlen et al., 2006, 2007; Gutknecht et al., 2019; Lefèvre et al., 2019; Schneider et al., 2008; Séférian et al., 2013; Steinacher et al., 2010; Tagliabue et al., 2010).

The dynamical component is the latest Mercator Ocean global 1/12° high-resolution ocean model system, extensively described and validated in Lellouche et al. (2013, 2018). This system provides daily and 1/4°-coarsened fields of horizontal and vertical current velocities, vertical eddy diffusivity, mixed layer depth, sea ice fraction, potential temperature, salinity, sea surface height, surface wind speed, freshwater fluxes and net surface solar shortwave irradiance that drive the transport of biogeochemical tracers. This system also features a reduced-order Kalman filter based on the Singular Evolutive Extended Kalman filter (SEEK) formulation introduced by Pham et al. (1998), that assimilates, on a 7-day assimilation cycle, along-track altimeter data, satellite Sea Surface Temperature and Sea-Ice Concentration from

OSTIA, and *in situ* temperature and salinity vertical profiles from the CORA 4.2 in situ database.

In addition, the biogeochemical component of the coupled system also embeds a reduced order Kalman filter (similar to the above mentioned) that operationally assimilates daily L4 remotely sensed surface chlorophyll (https://resources.marine.copernicus.eu/documents/QUID/CMEMS-GLO-QUID-001-028.pdf). In parallel, a climatological-damping is applied to nitrate, phosphate, oxygen, silicate - with World Ocean Atlas 2013 - to dissolved inorganic carbon and alkalinity – with GLODAPv2 climatology (Key et al., 2015) - and to dissolved organic carbon and iron - with a 4000-year PISCES climatological run. This relaxation is set to mitigate the impact of the physical data assimilation in the offline coupled hydrodynamic-biogeochemical system, leading significant rises of nutrients in the Equatorial Belt area, and resulting in an unrealistic drift of various biogeochemical variables e.g. chlorophyll, nitrate, phosphate (Fennel et al., 2019; Park et al., 2018). The time-scale associated with this climatological damping is set to 1 year and allows a smooth constraint that has been shown to be efficient to reduce the model drift.

A.2 The Mediterranean Sea biogeochemical model MedBFM

The Mediterranean Sea biogeochemical model MedBFM, is based on the system described in Teruzzi et al. (2014) and Salon et al. (2019).

The physical forcing fields needed to compute the transport include the 3-d horizontal and vertical current velocities, vertical eddy diffusivity, potential temperature, and salinity and 2-d data surface data for wind stress. These forcing datasets are simulated by the Mediterranean Sea Monitoring and Forecasting Centre (MED–MFC) in the Copernicus Marine Environmental Monitoring Service (CMEMS, http://marine.copernicus.eu). The biogeochemical model is then offline forced adopting the output computed by the CMEMS MED-MFC. In the present application, we switched off the biogeochemical assimilation scheme that is currently used in the operational MED-MFC system.

The light propagation is resolved coupling an atmospheric multispectral radiative transfer model (Lazzari et al., 2020) with an in-water radiative model (Dutkiewicz et al., 2015) featuring bands at 25 nm resolution in the UV and visible wavelengths.

The horizontal resolution is approximately 6 km and there are 72 vertical levels with 3 m resolution at surface coarsening at 300 m for the deeper layers. The biogeochemical model here adopted (Biogeochemical Flux Model -- BFM -- ; (Vichi et al., 2015)) has been already applied to simulate primary producers biogeochemistry (Lazzari et al., 2012), alkalinity spatial and temporal variability (Cossarini et al., 2015), and CO_2 fluxes (Canu et al., 2015) for the Mediterranean Sea, and has been corroborated using *in situ* data for the operational purposes within CMEMS (Salon et al., 2019). The BFM model has been expanded in the present configuration adding the dynamics of coloured dissolved organic carbon (CDOM) by assuming a constant CDOM:DOC production ratio (i.e. 2%, as in (Dutkiewicz et al., 2015)). The absorption of CDOM, is described using reference absorption at 450 nm of 0.015 m2/mgC (Dutkiewicz et al., 2015) and an exponential slope of 0.017 nm^{-1} (Babin et al., 2003; Organelli et al., 2014).

A.3 BGC-Argo K_d estimates

The data used to compute the K_d metrics are quality checked according to Organelli et al. (2017). Moreover, for the K_d logarithmic interpolation, the following selection rules were applied: the profile must have at least 5 BGC Argo float sampling in the first optical depth, the gap between the two shallower acquisitions must be less than 10 meters, and there must be at least one measurement deeper than 15 meters.

A.4 Figures

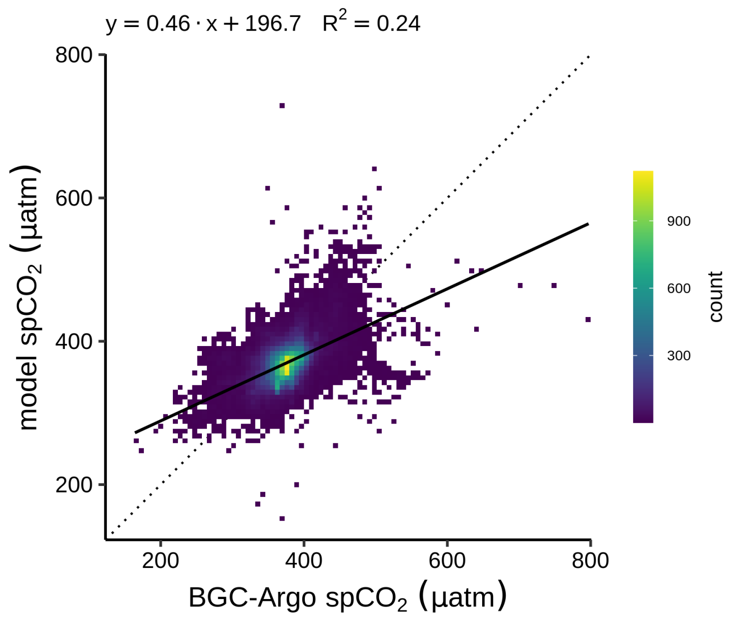


Figure A1. Same as Figure 3 but for $spCO_2$.

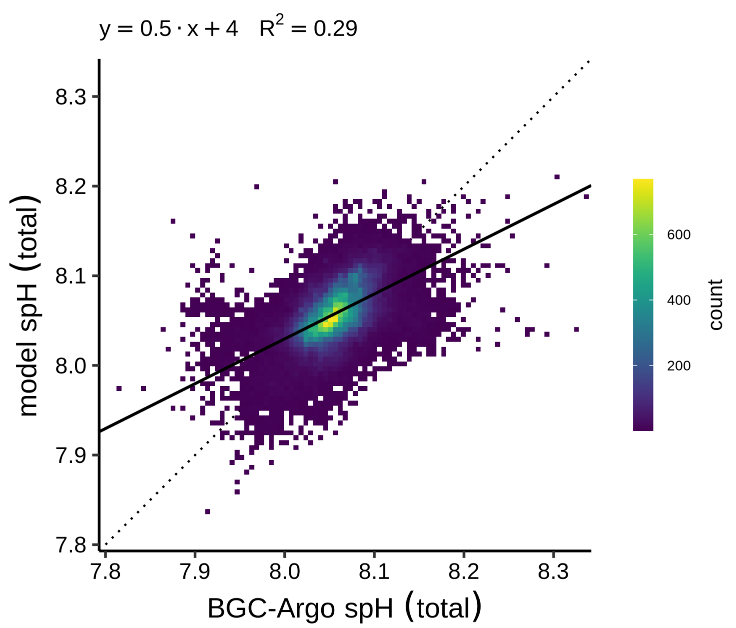


Figure A2. Same as Figure 3 but for spH. Note that spH is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

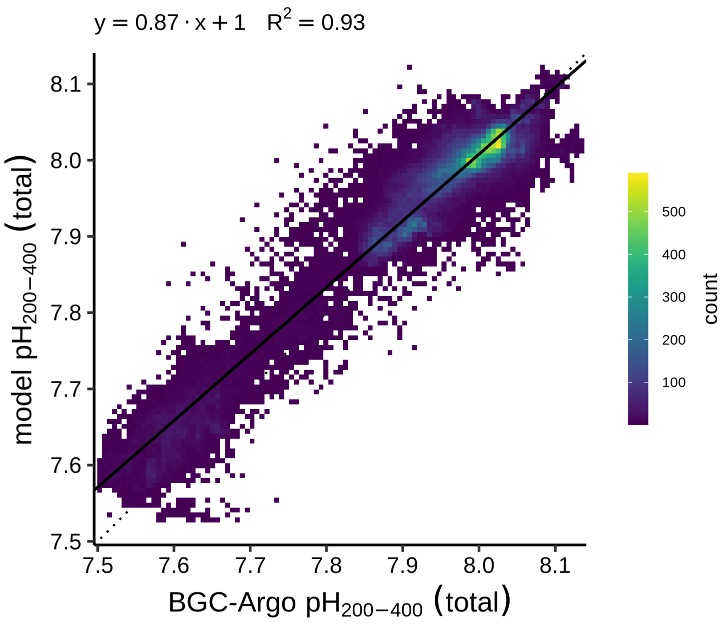


Figure A3. Same as Figure 3 but for $pH_{200-400}$. Note that $pH_{200-400}$ is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

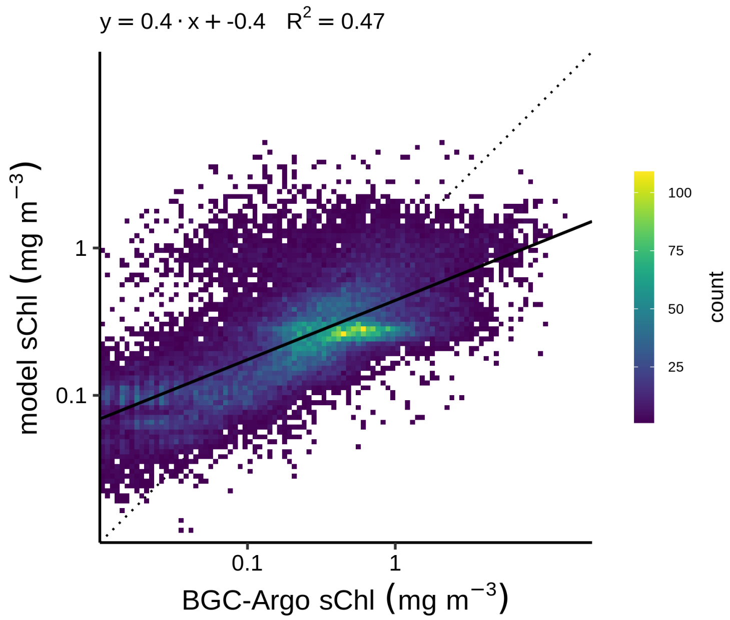


Figure A4. Same as Figure 3 but for sChl. Note that the least squares regression is computed on the $\log_{10}$-transformed data to account that sChl covers several orders of magnitude and it is lognormally distributed (Campbell, 1995). Data lower than 0.01 mg m^{-3} are not included.

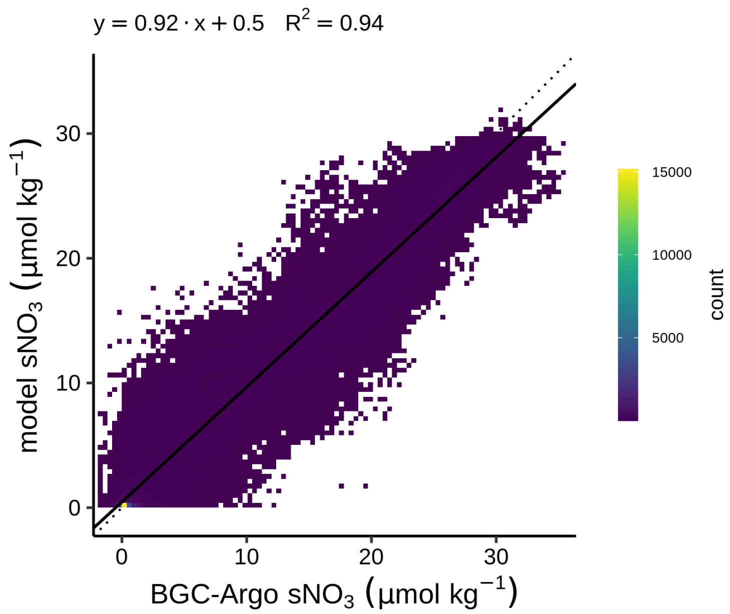


Figure A5. Same as Figure 3 but for sNO_3. Note that sNO_3 is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

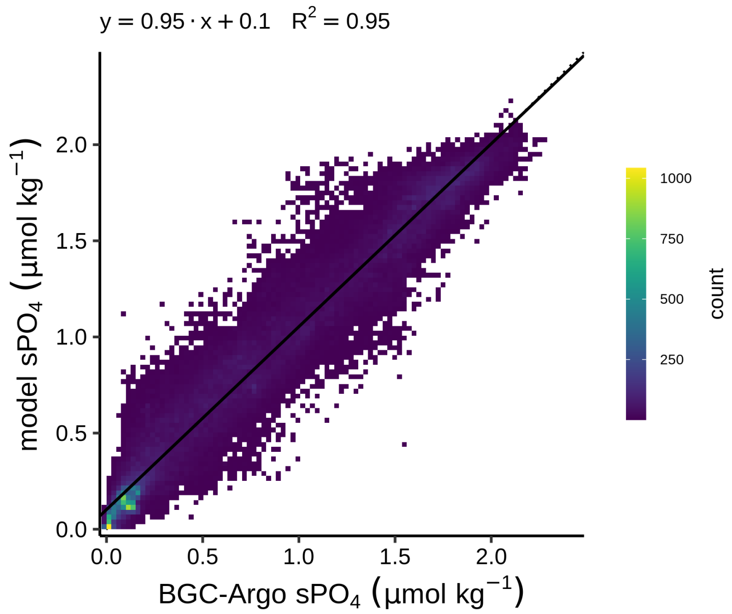


Figure A6. Same as Figure 3 but for sPO_4.

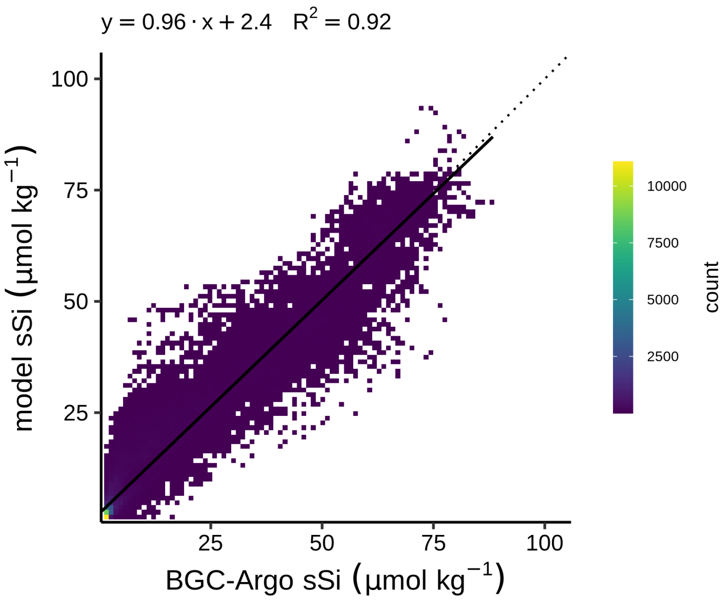


Figure A7. Same as Figure 3 but for sSi.

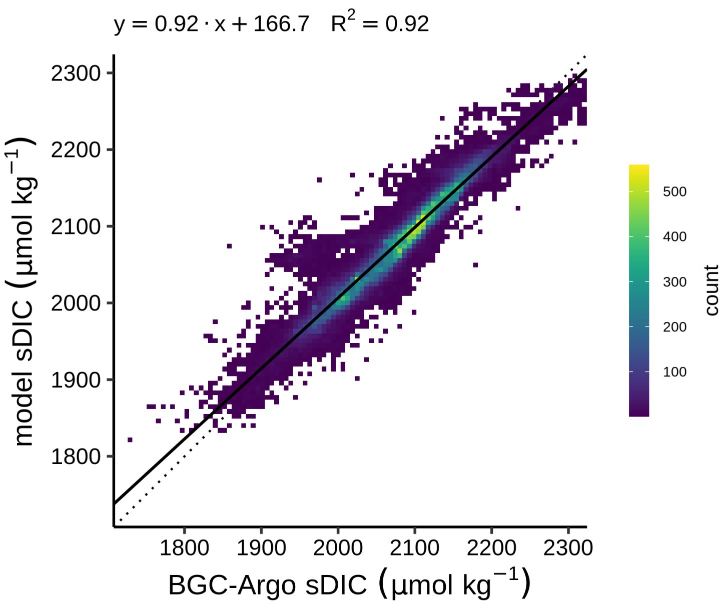


Figure A8. Same as Figure 3 but for sDIC.

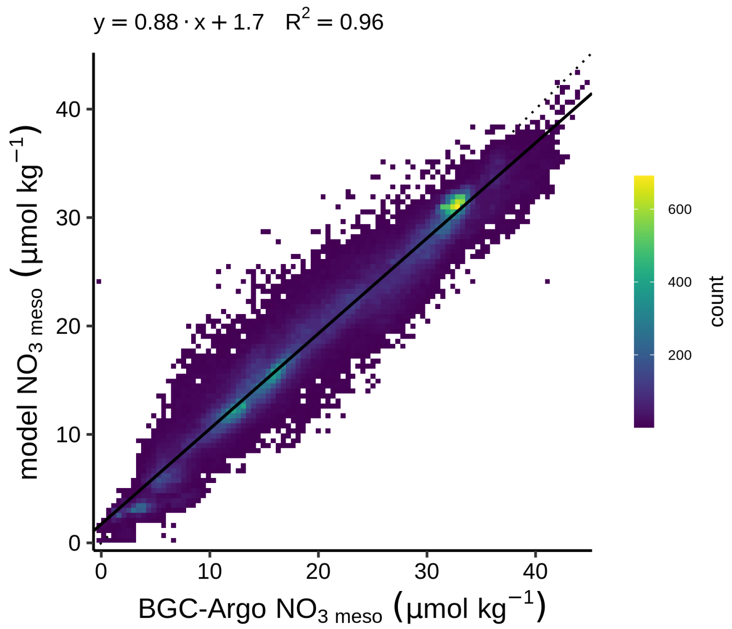


Figure A9. Same as Figure 3 but for NO_3 meso. Note that NO_3 meso is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

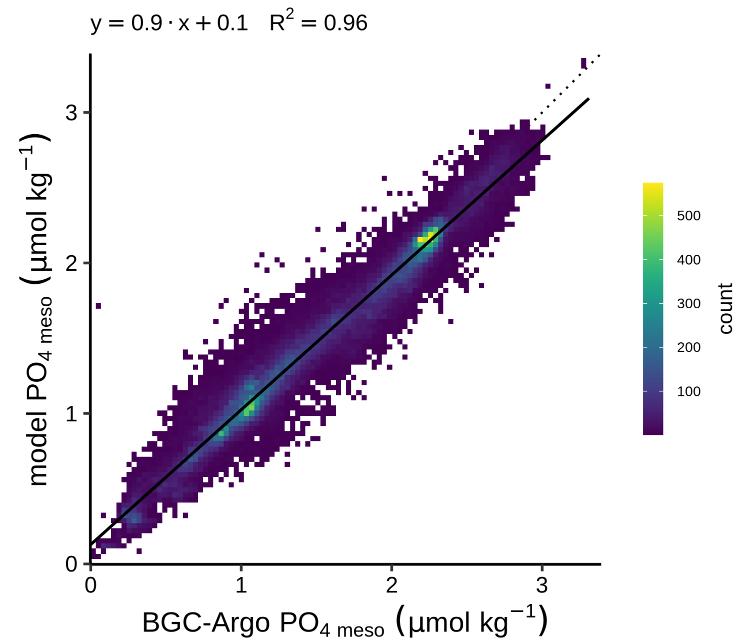


Figure A10. Same as Figure 3 but for PO_4 meso.

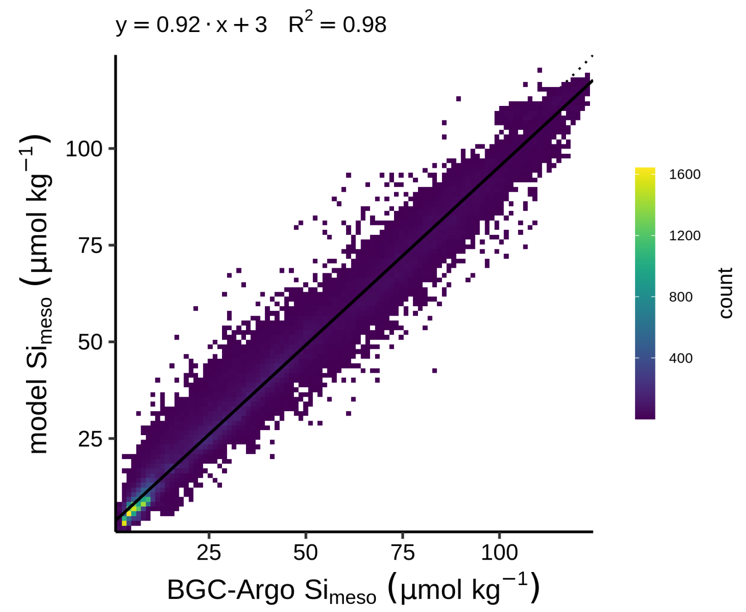


Figure A11. Same as Figure 3 but for Si_{meso}.

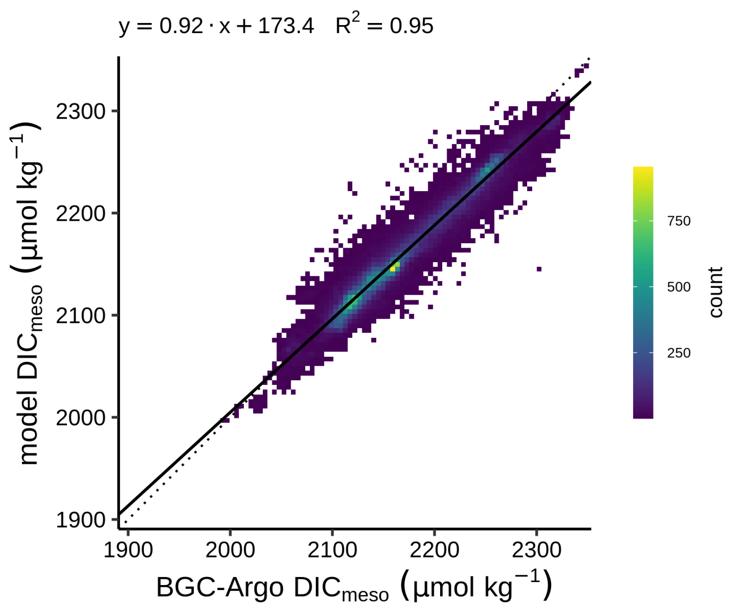


Figure A12. Same as Figure 3 but for DIC_{meso}.

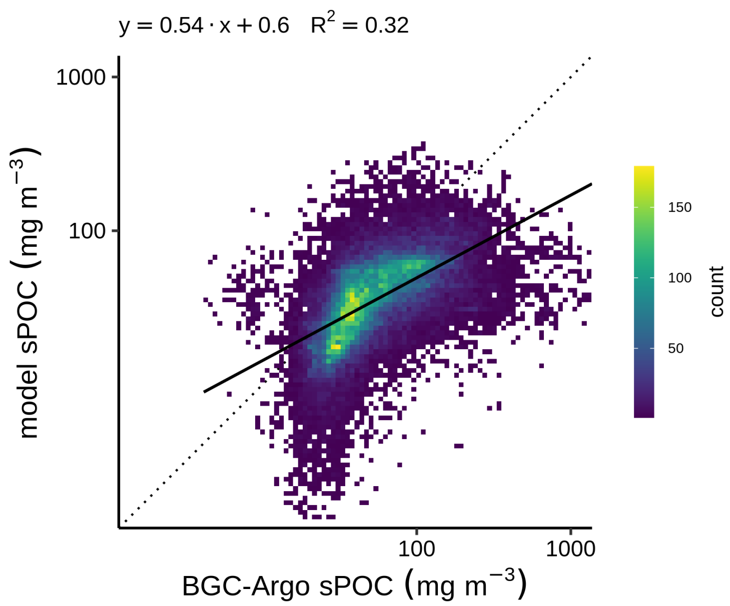


Figure A13. Same as Figure 3 but for sPOC. Note that the least squares regression is computed on the $\log_{10}$-transformed data to account that sPOC covers several orders of magnitude and it is lognormally distributed (Campbell, 1995). Data lower than 0.01 mg m^{-3} are not included.

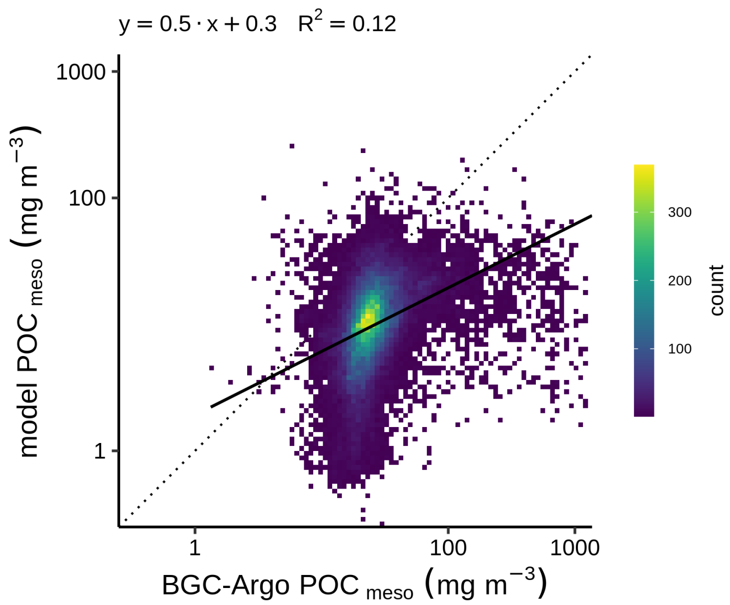


Figure A14. Same as Figure 3 but for POC_{meso}. Note that the least squares regression is computed on the log_{10}-transformed data to account that POC_{meso} covers several orders of magnitude and it is lognormally distributed (Campbell, 1995). Data lower than 0.01 mg m^{-3} are not included.

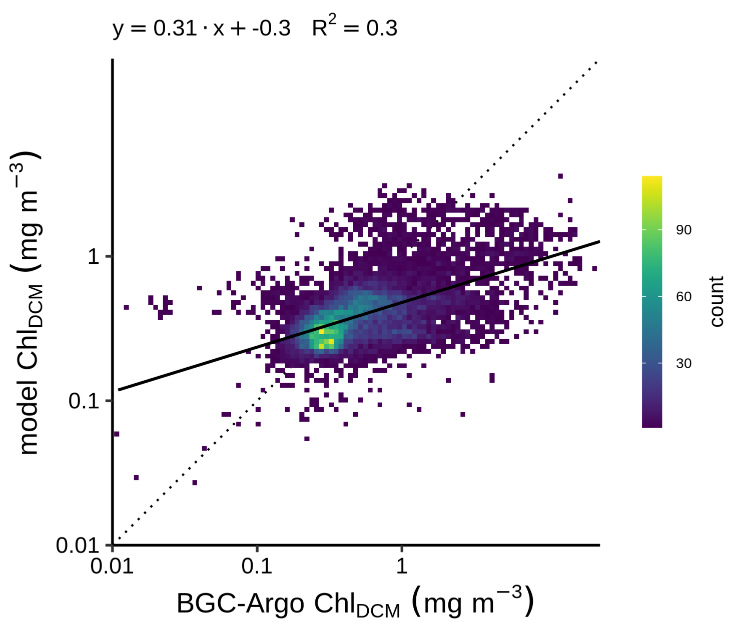


Figure A15. Same as Figure 3 but for Chl_{DCM}. Note that the least squares regression is computed on the log_{10}-transformed data to account that Chl_{DCM} covers several orders of magnitude and it is lognormally distributed (Campbell, 1995). Data lower than 0.01 mg m^{-3} are not included. Observed DCMs deeper than 250 m are not included.

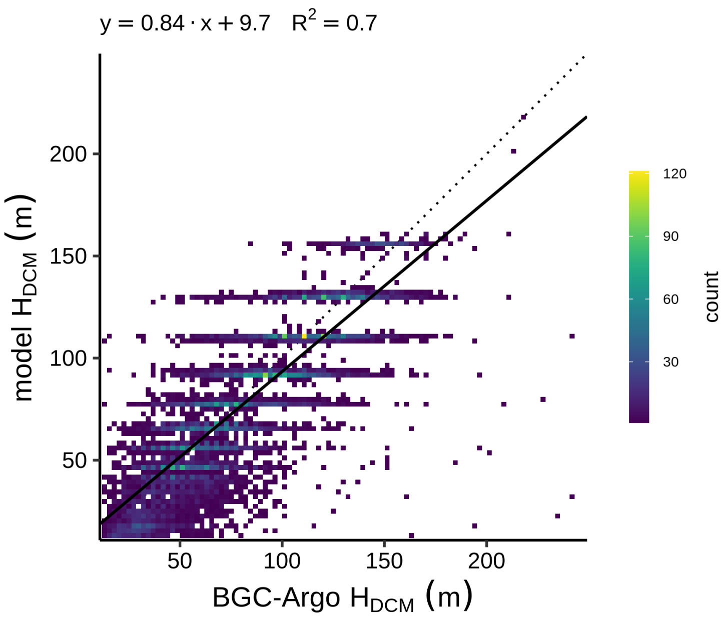


Figure A16. Same as Figure 3 but for H_{DCM}. Observed DCMs deeper than 250 m are not included.

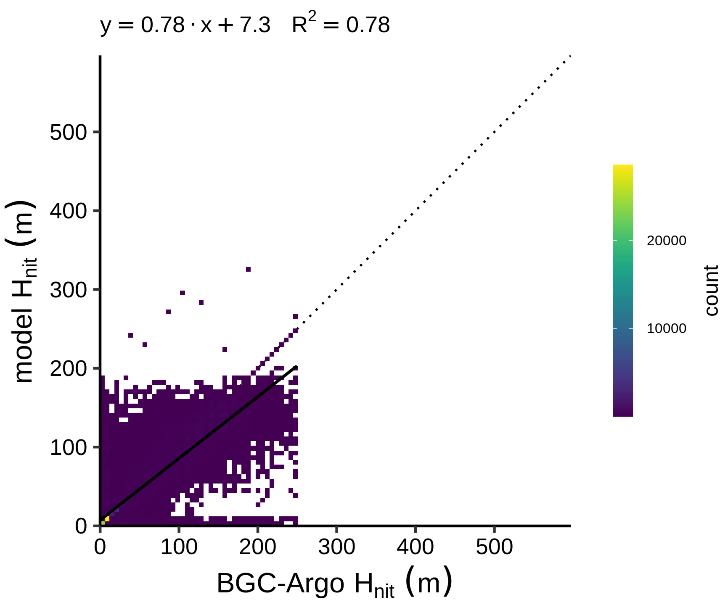


Figure A17. Same as Figure 3 but for H_{nit}. Observed nitracline deeper than 250 m are not included.

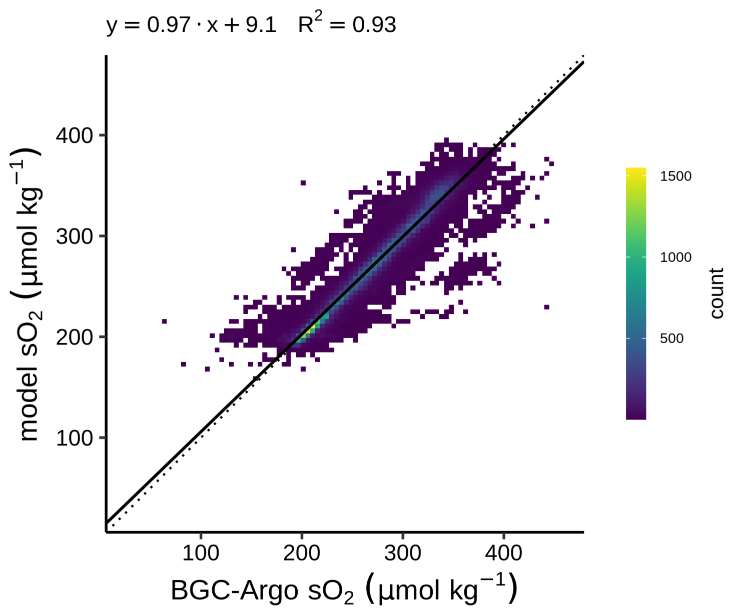


Figure A18. Same as Figure 3 but for sO_2.

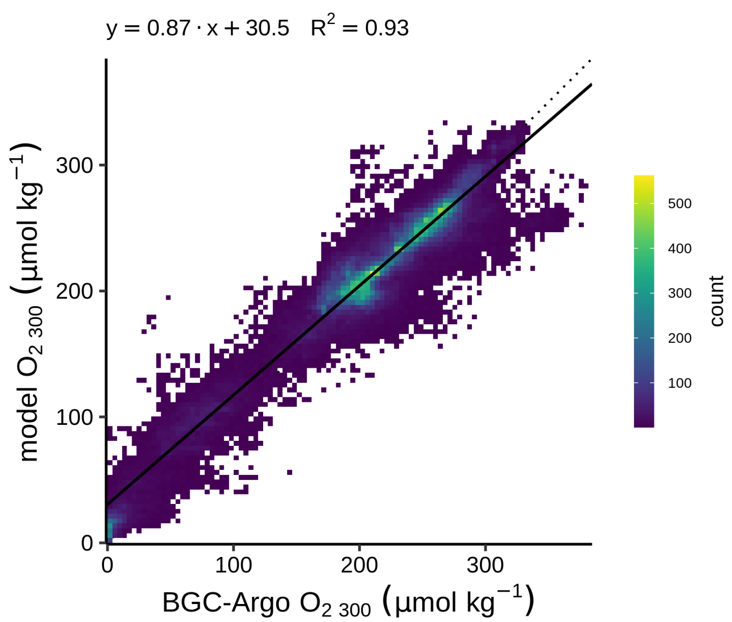


Figure A19. Same as Figure 3 but for $O_{2\ 300}$.

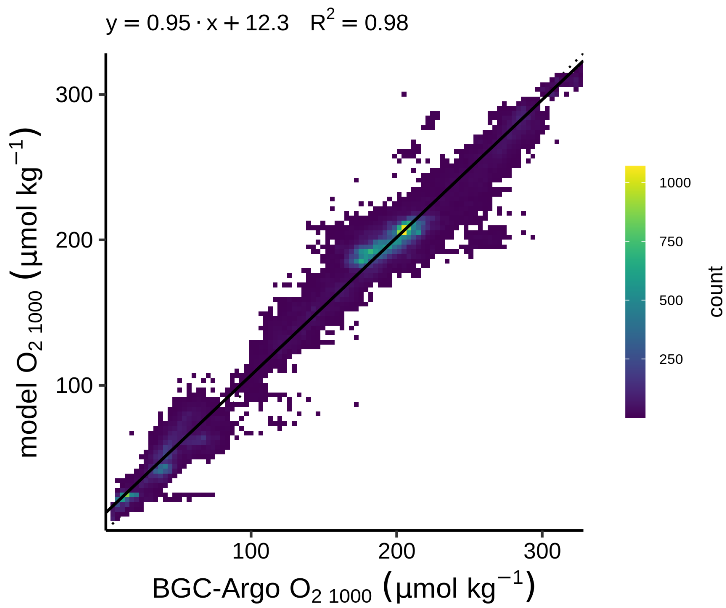


Figure A20. Same as Figure 3 but for $O_{2\ 1000}$.

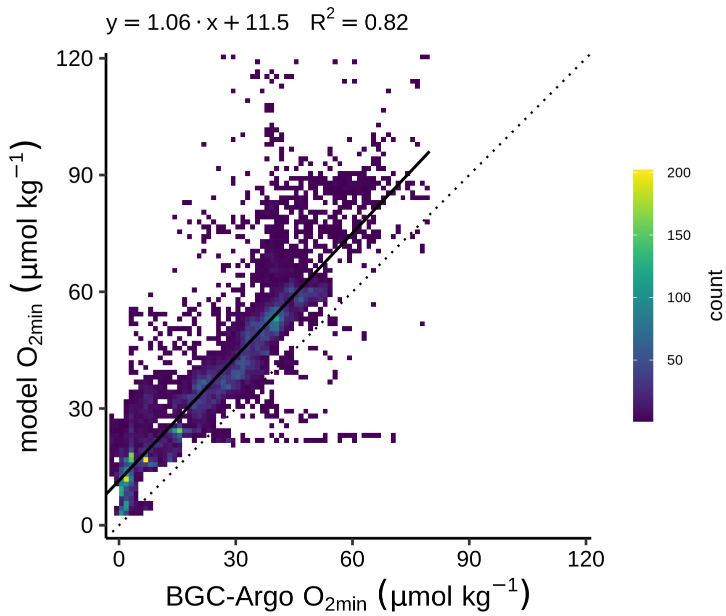
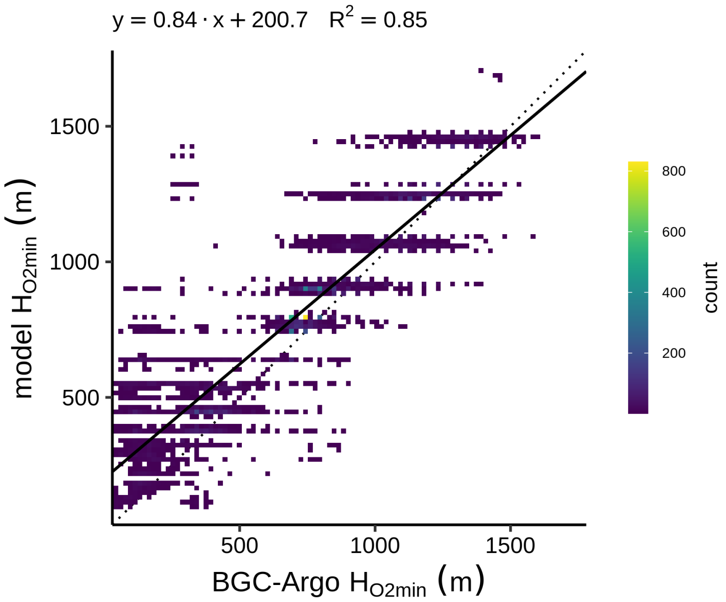


Figure A21. Same as Figure 3..



1

2 **Figure A22.** Same as Figure 3 but for H_{O2min}.

3

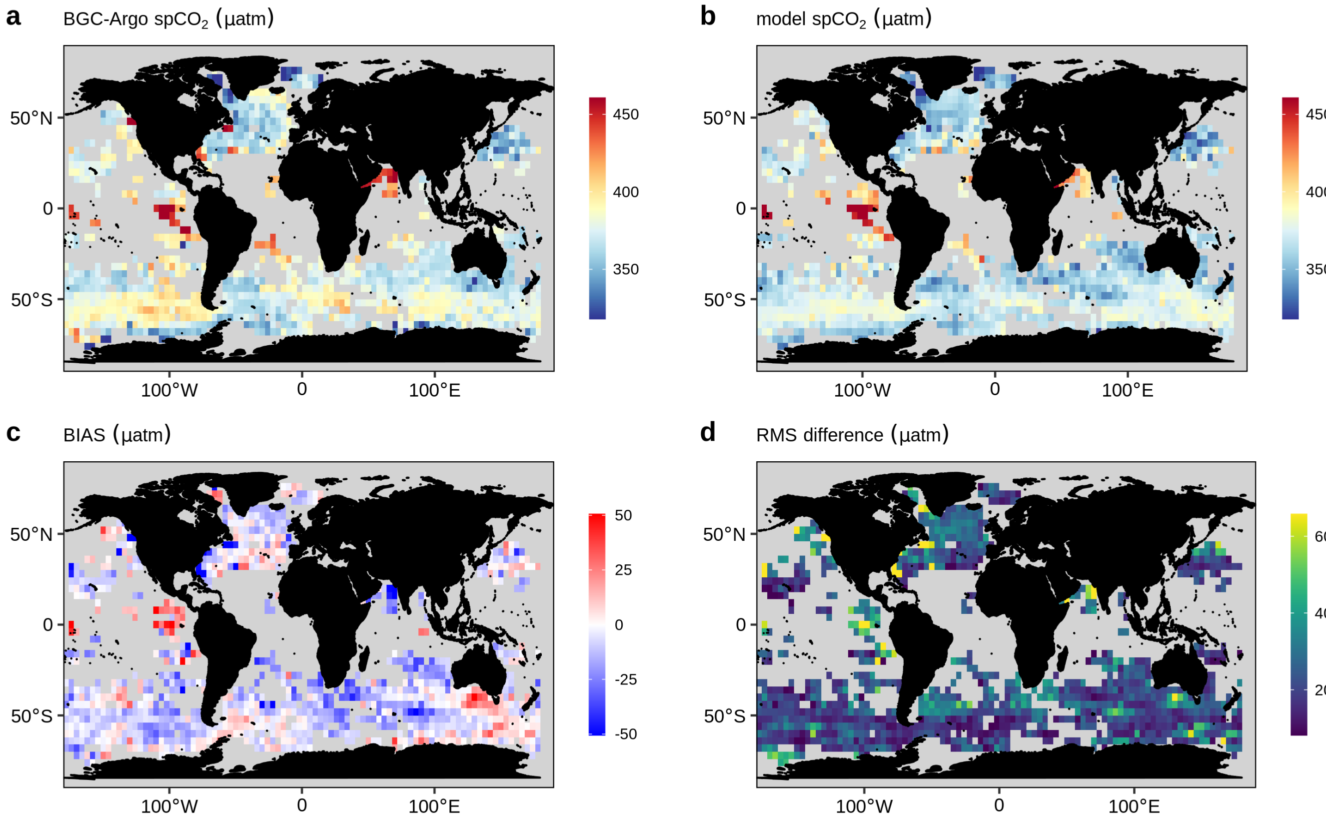


Figure A23. Same as Figure 4 but for $spCO_2$.

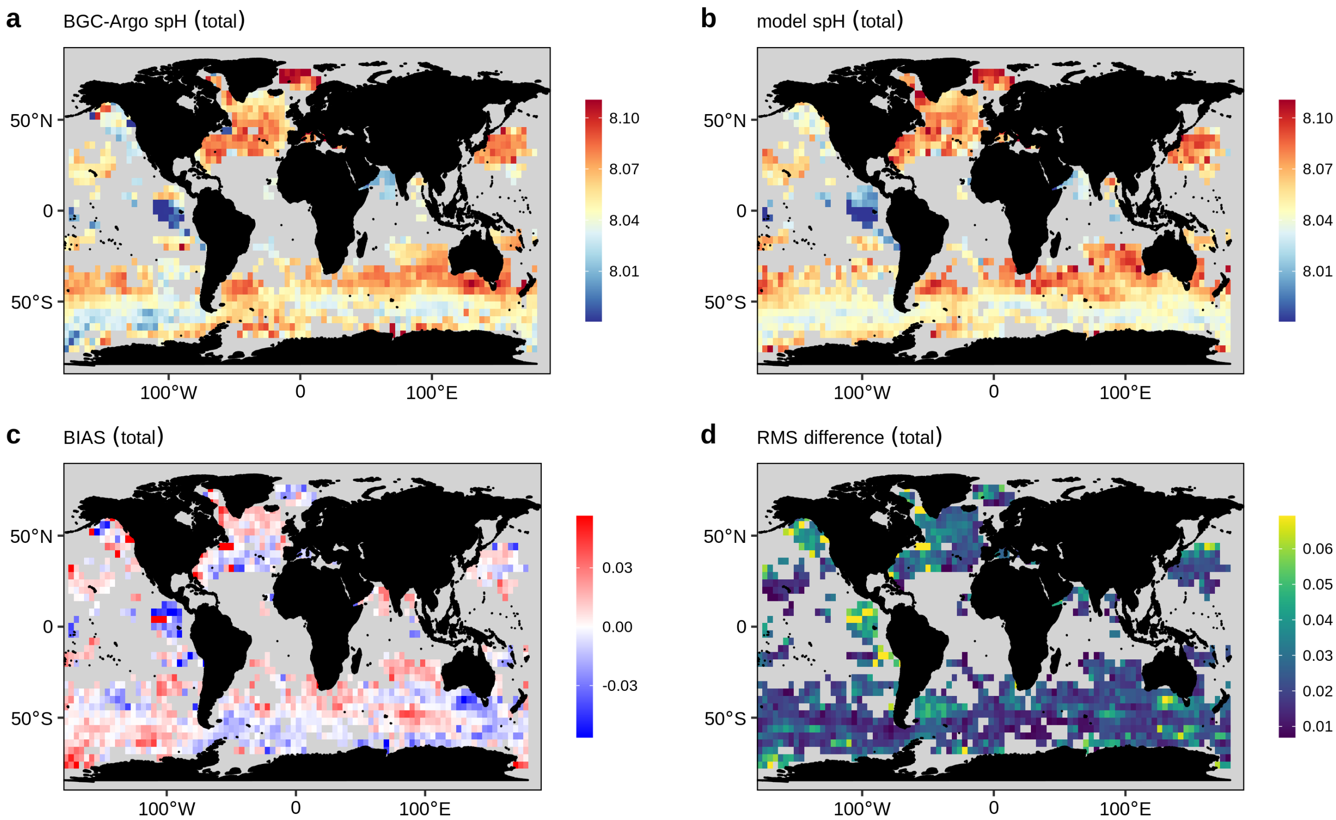


Figure A24. Same as Figure 4 but for spH. Note that spH is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

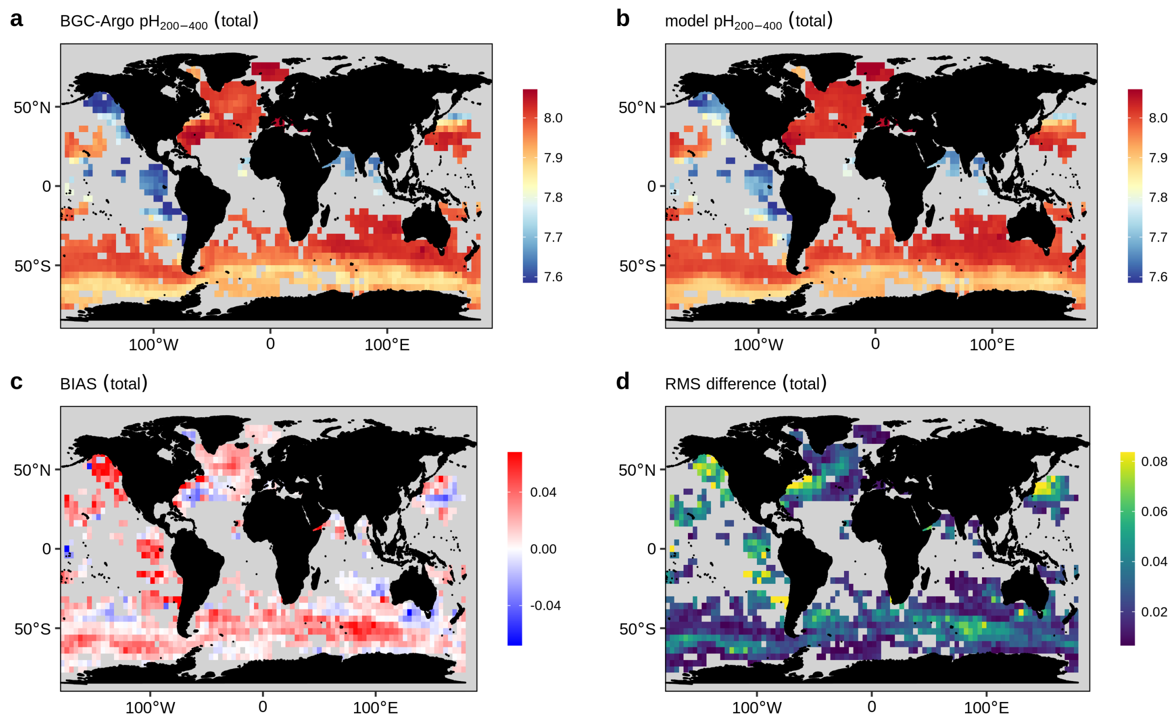


Figure A25. Same as Figure 4 but for $pH_{200\text{-}400}$. Note that $pH_{200\text{-}400}$ is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

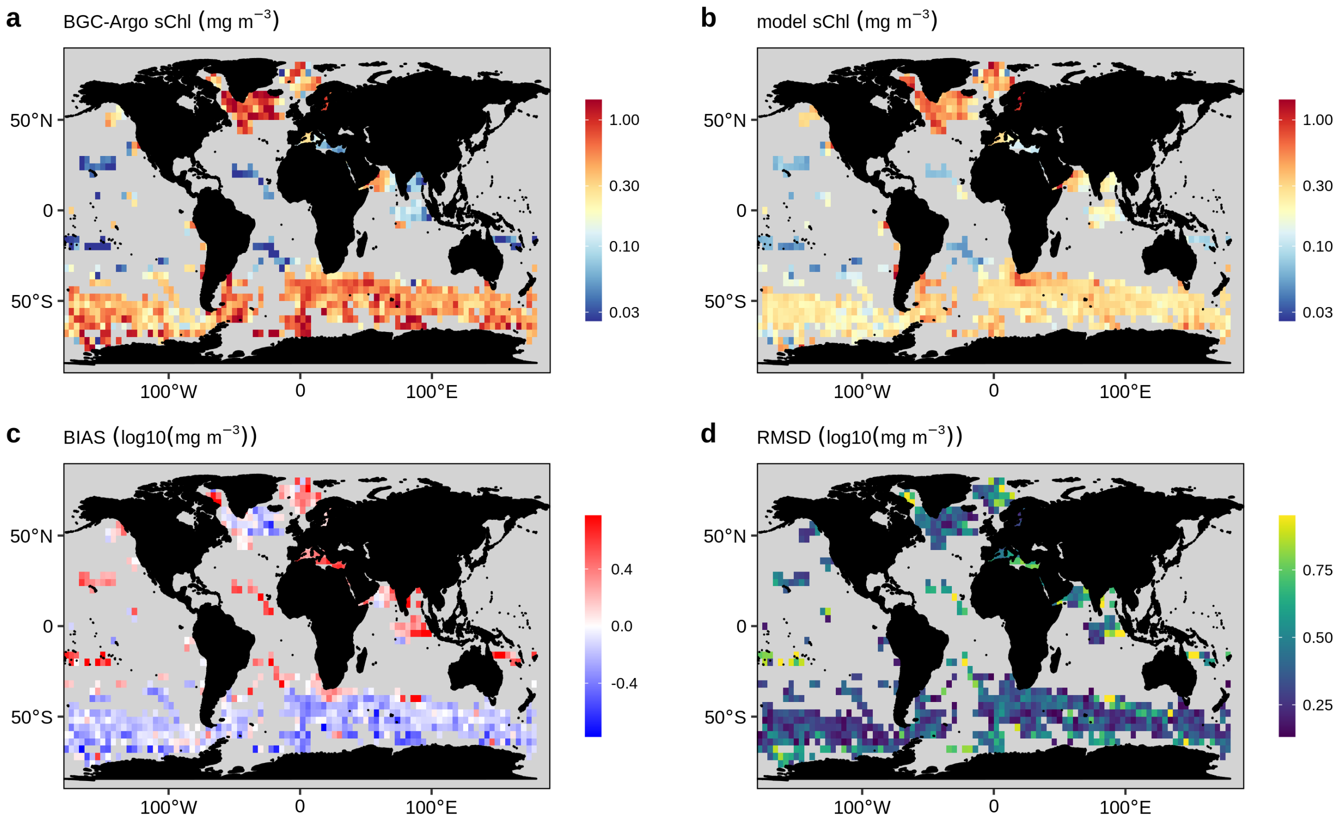


Figure A26. Same as Figure 4.

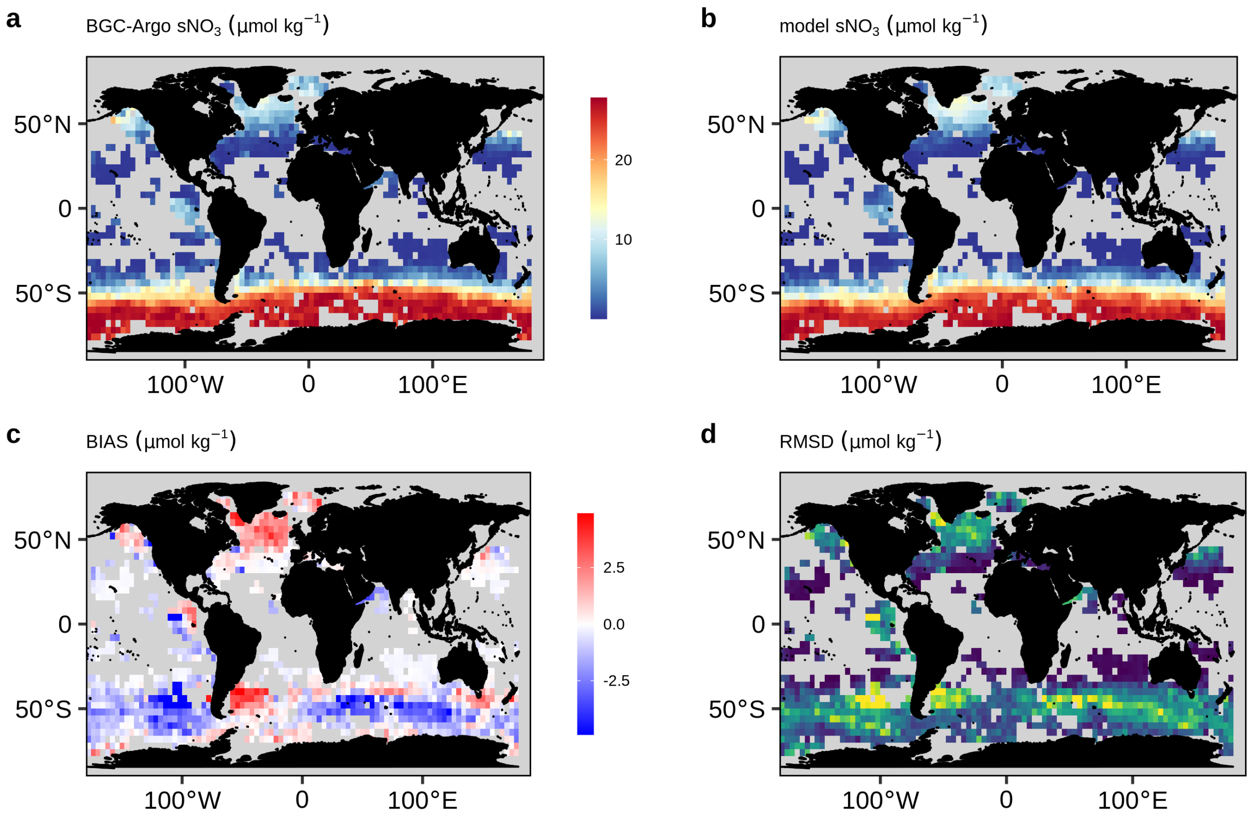


Figure A27. Same as Figure 4 but for sNO_3. Note that sNO_3 is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

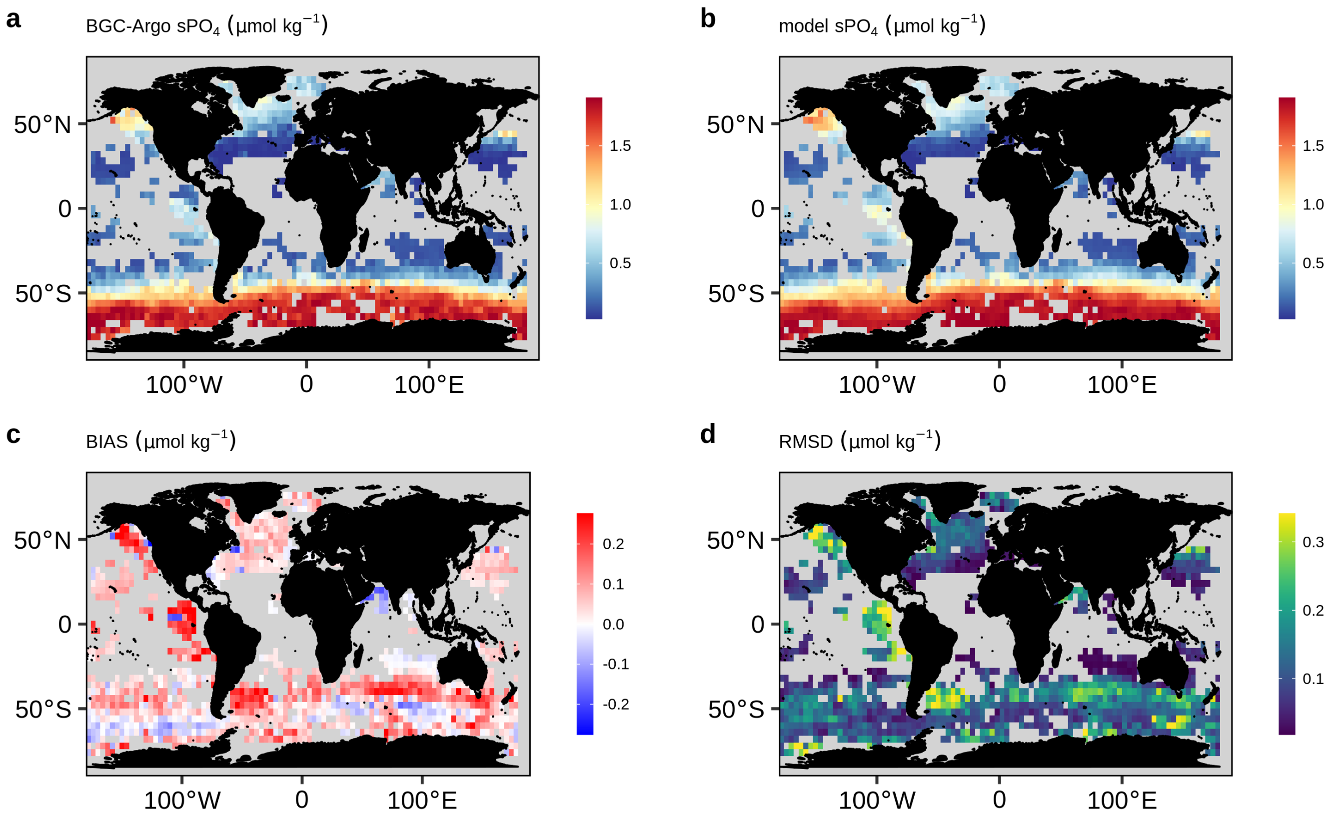


Figure A28. Same as Figure 4 but for sPO_4.

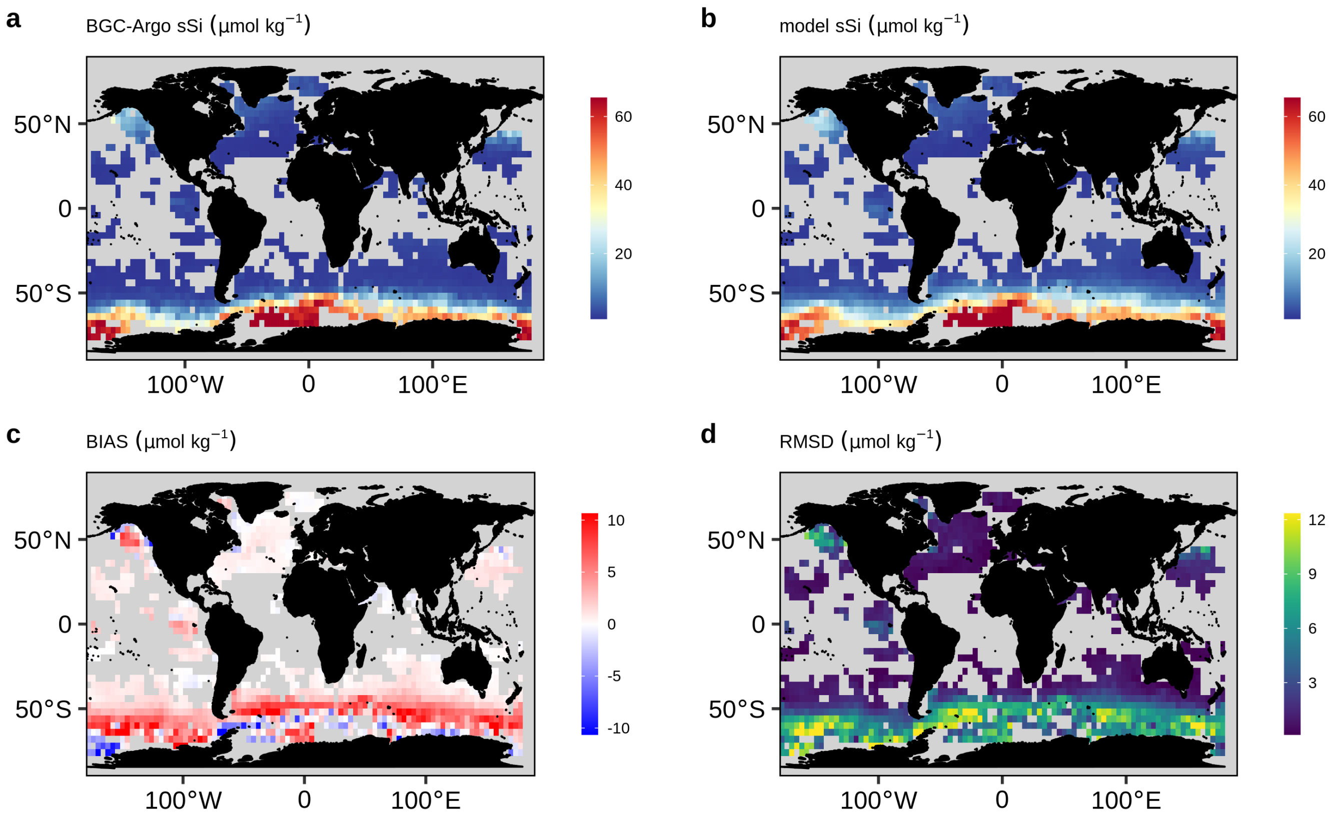


Figure A29. Same as Figure 4 but for sSi.

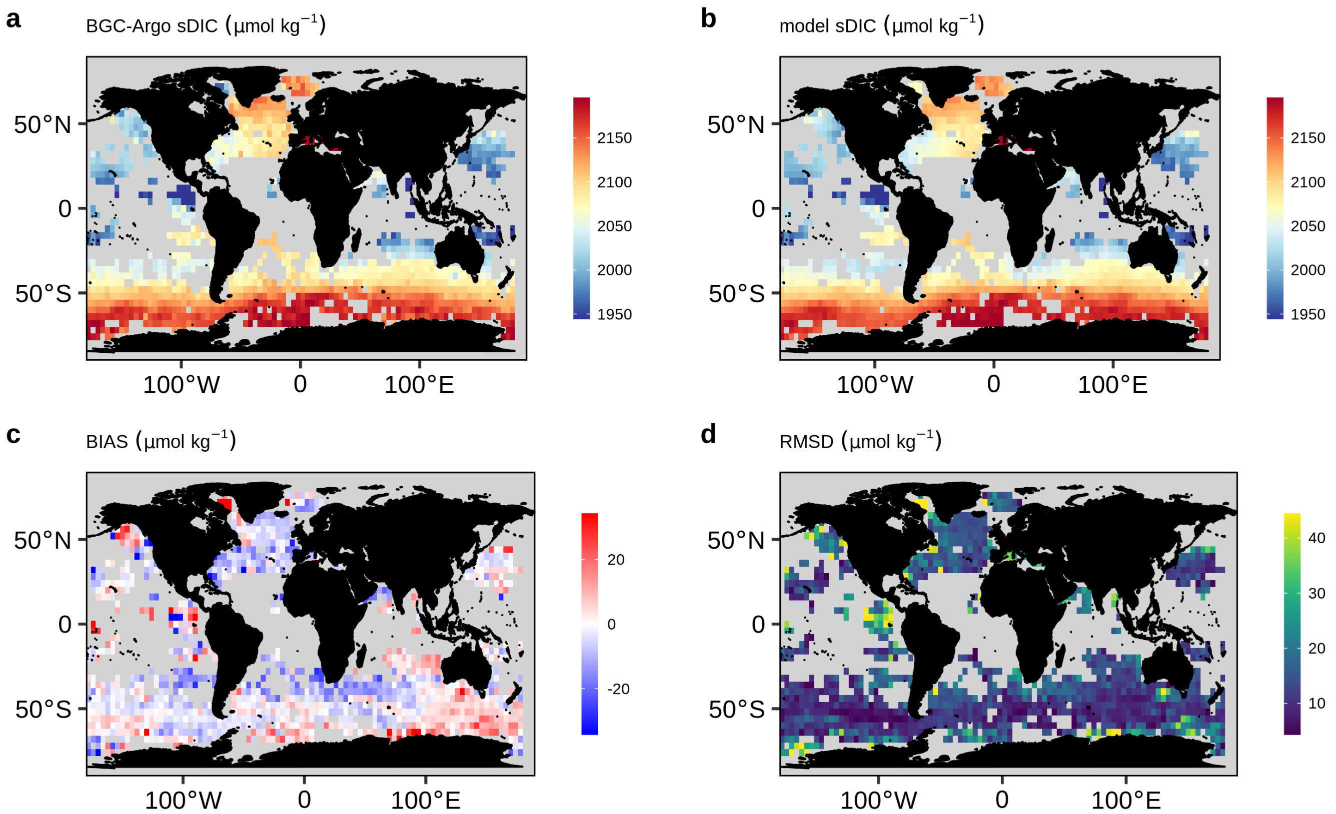


Figure A30. Same as Figure 4 but for sDIC.

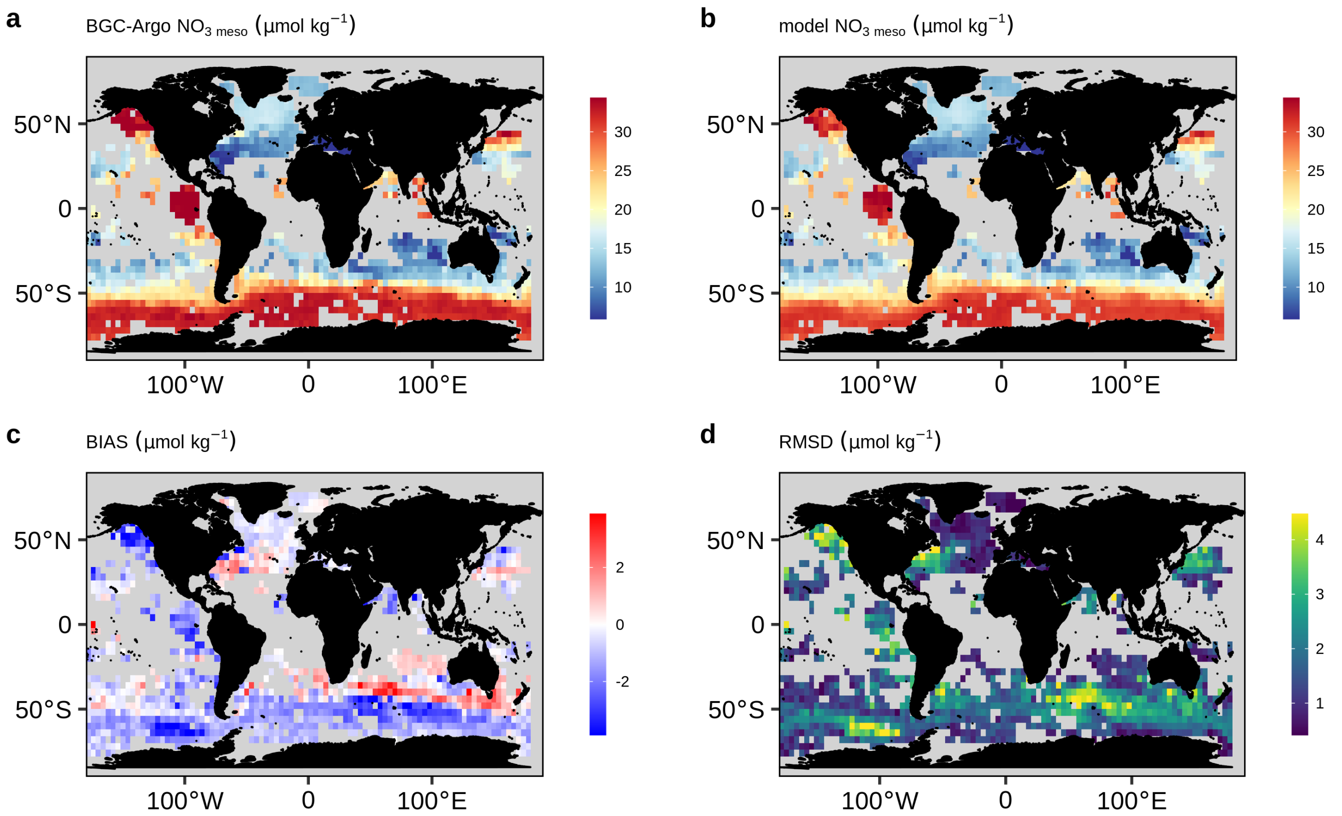


Figure A31. Same as Figure 4 but for $NO_{3\ meso}$. Note that $NO_{3\ meso}$ is calculated from both the direct observations of the floats and as well as the estimations from CANYON-B.

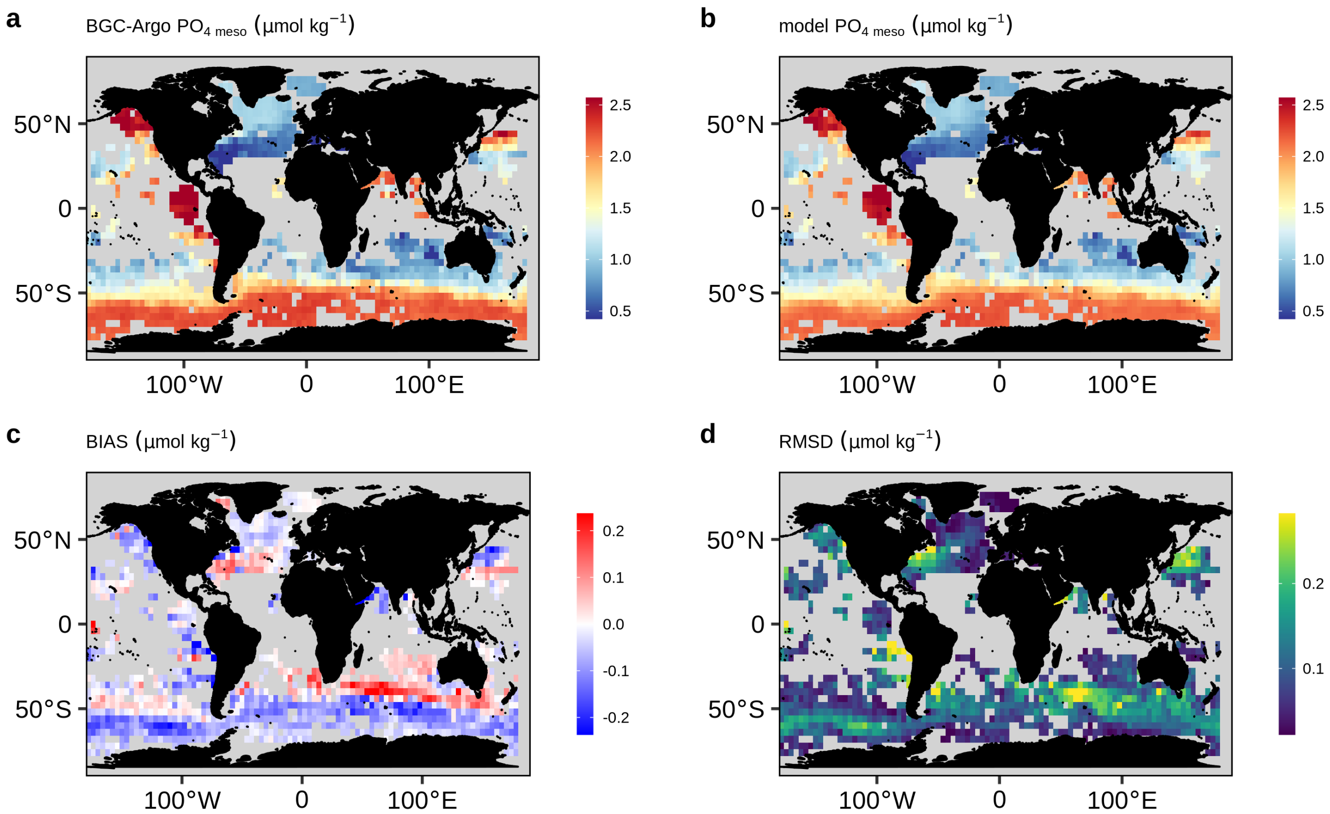


Figure A32. Same as Figure 4 but for $PO_{4\ \mathrm{meso}}$.

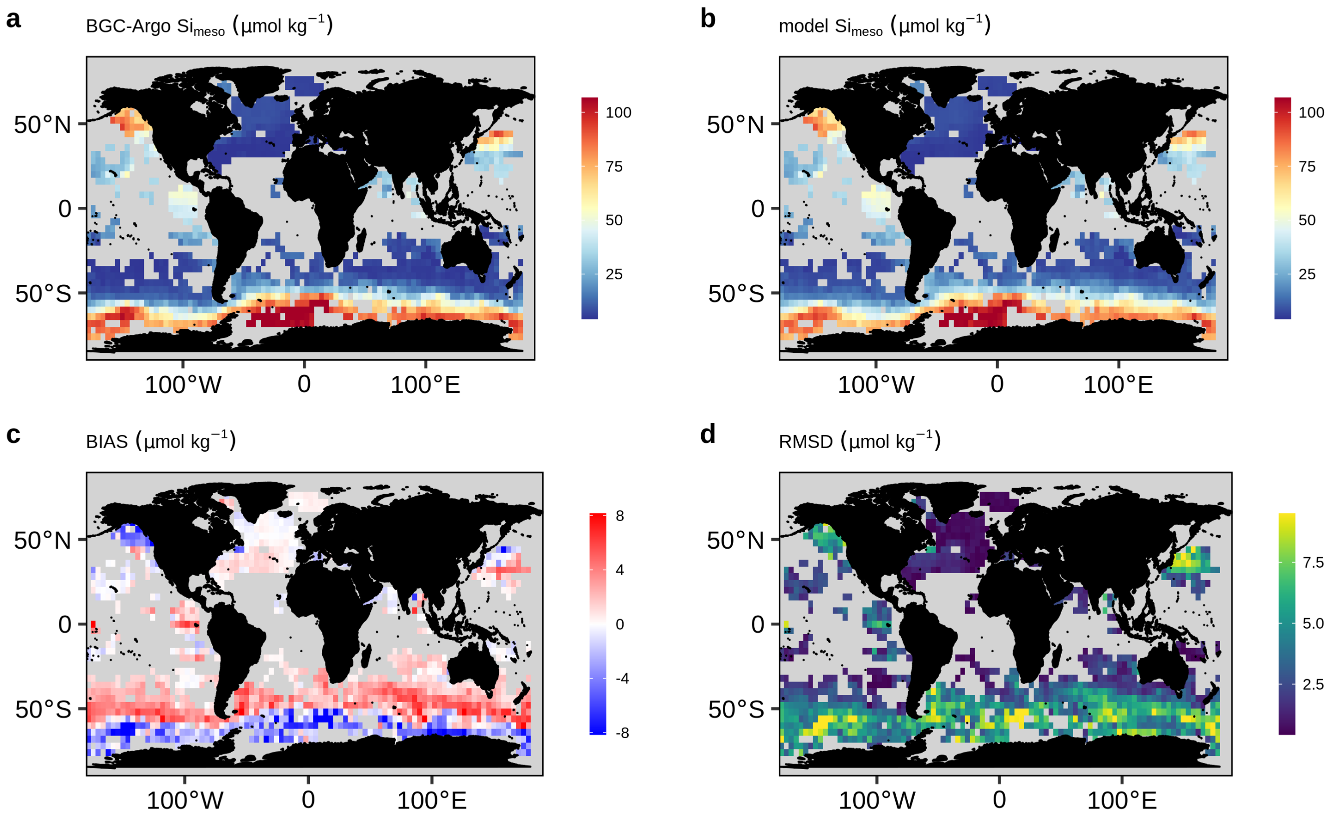


Figure A33. Same as Figure 4 but for Si_{meso}.

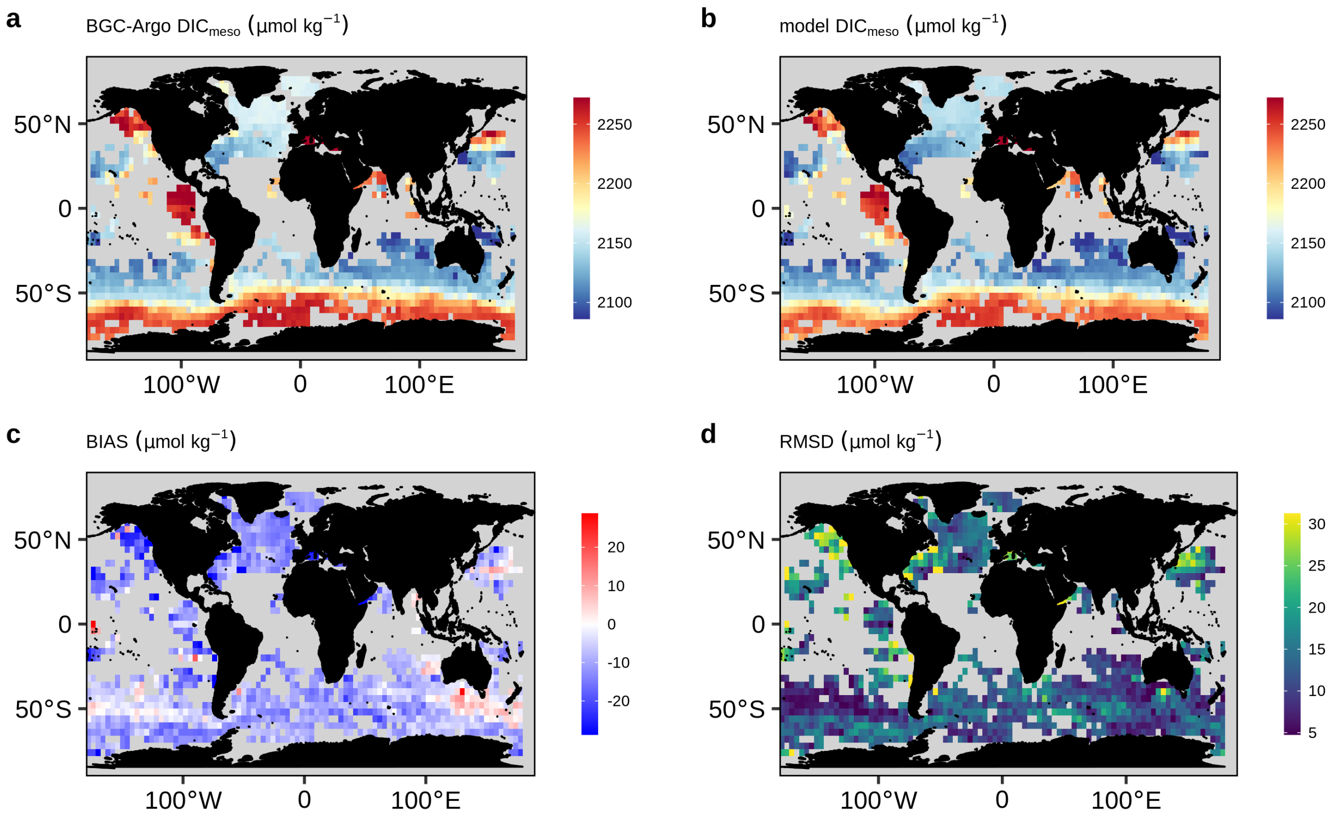


Figure A34. Same as Figure 4 but for DIC_{meso}.

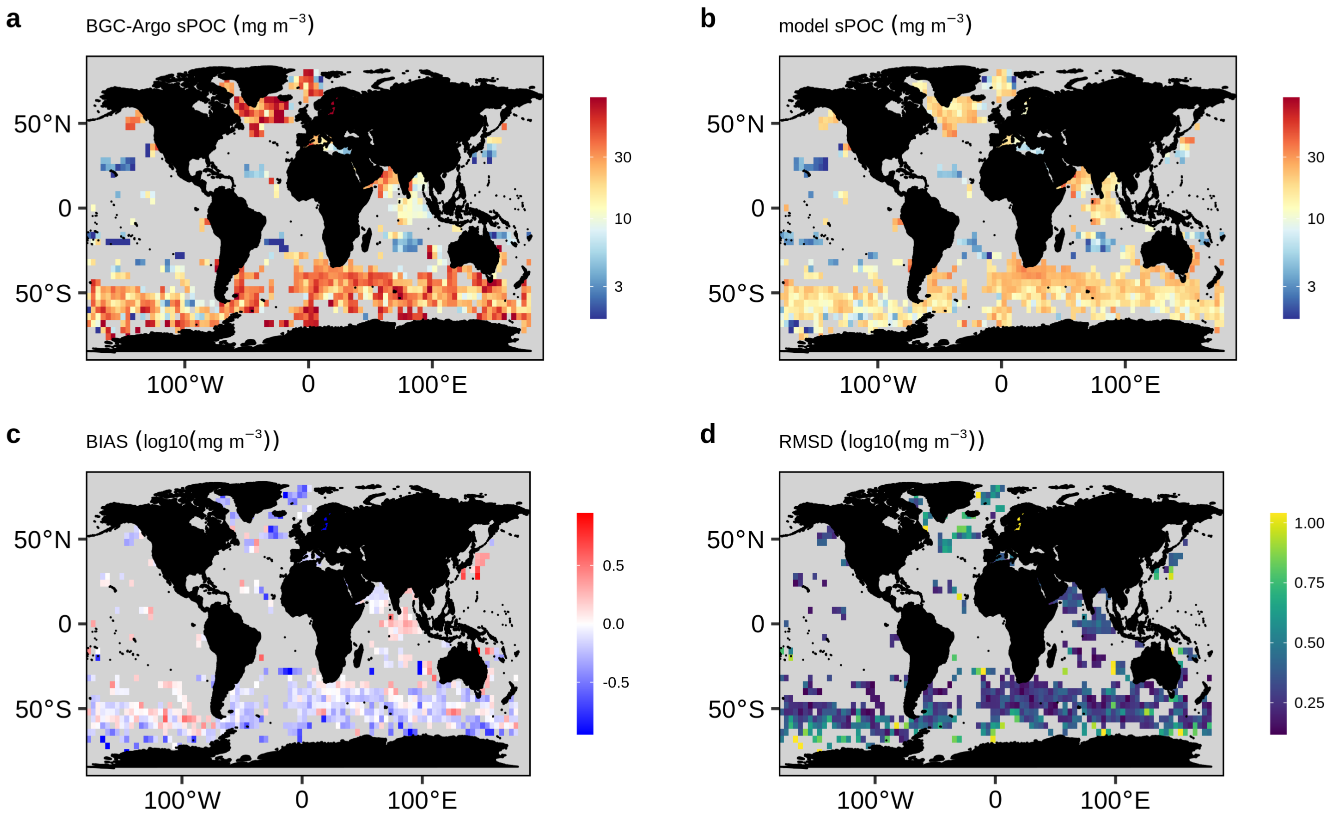


Figure A35. Same as Figure 4 but for sPOC. The BIAS and RMSD are computed on the $\log_{10}$-transformed data to account that sPOC covers several orders of magnitude and it is lognormally distributed (Campbell, 1995)

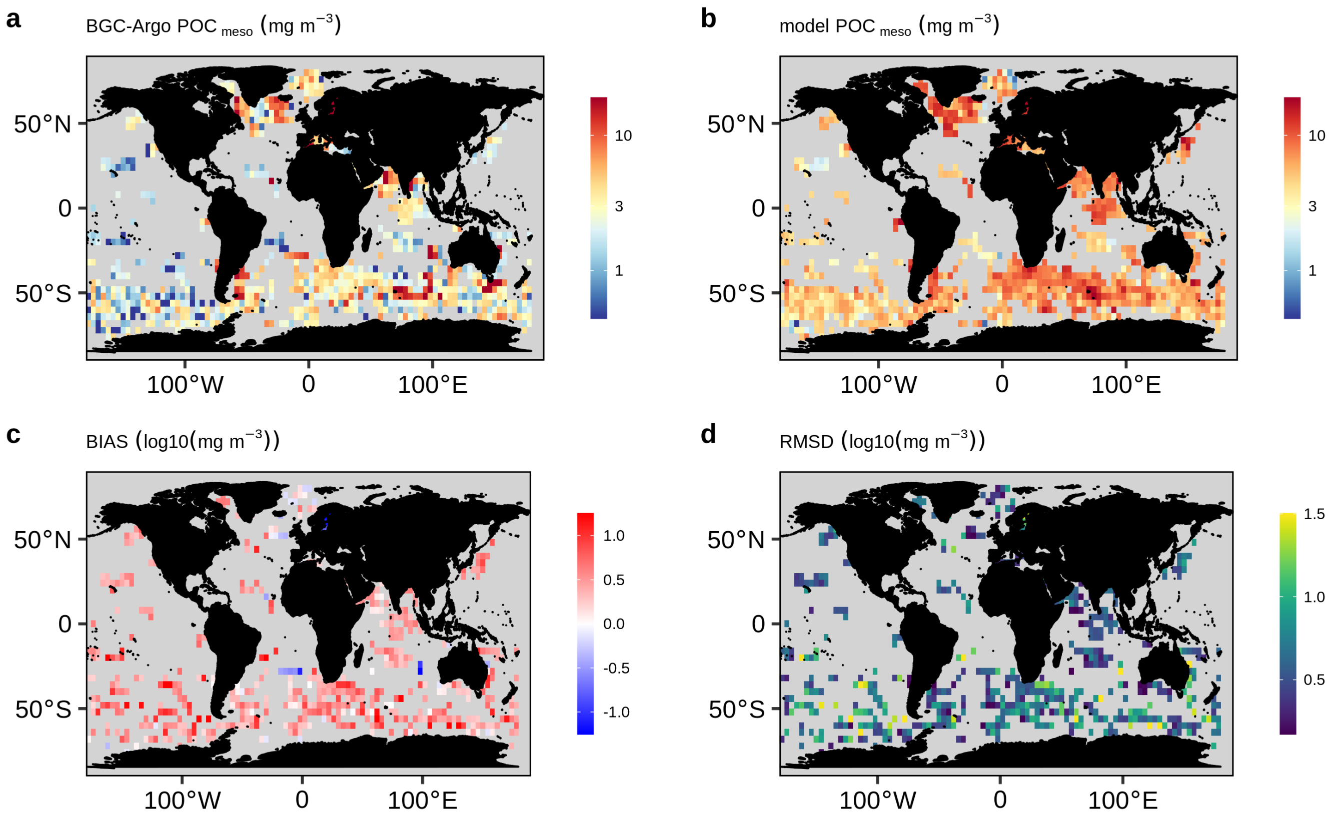


Figure A36. Same as Figure 4 but for POC_{meso}. The BIAS and RMSD are computed on the log_{10}-transformed data to account that POC_{meso} covers several orders of magnitude and it is lognormally distributed (Campbell, 1995)

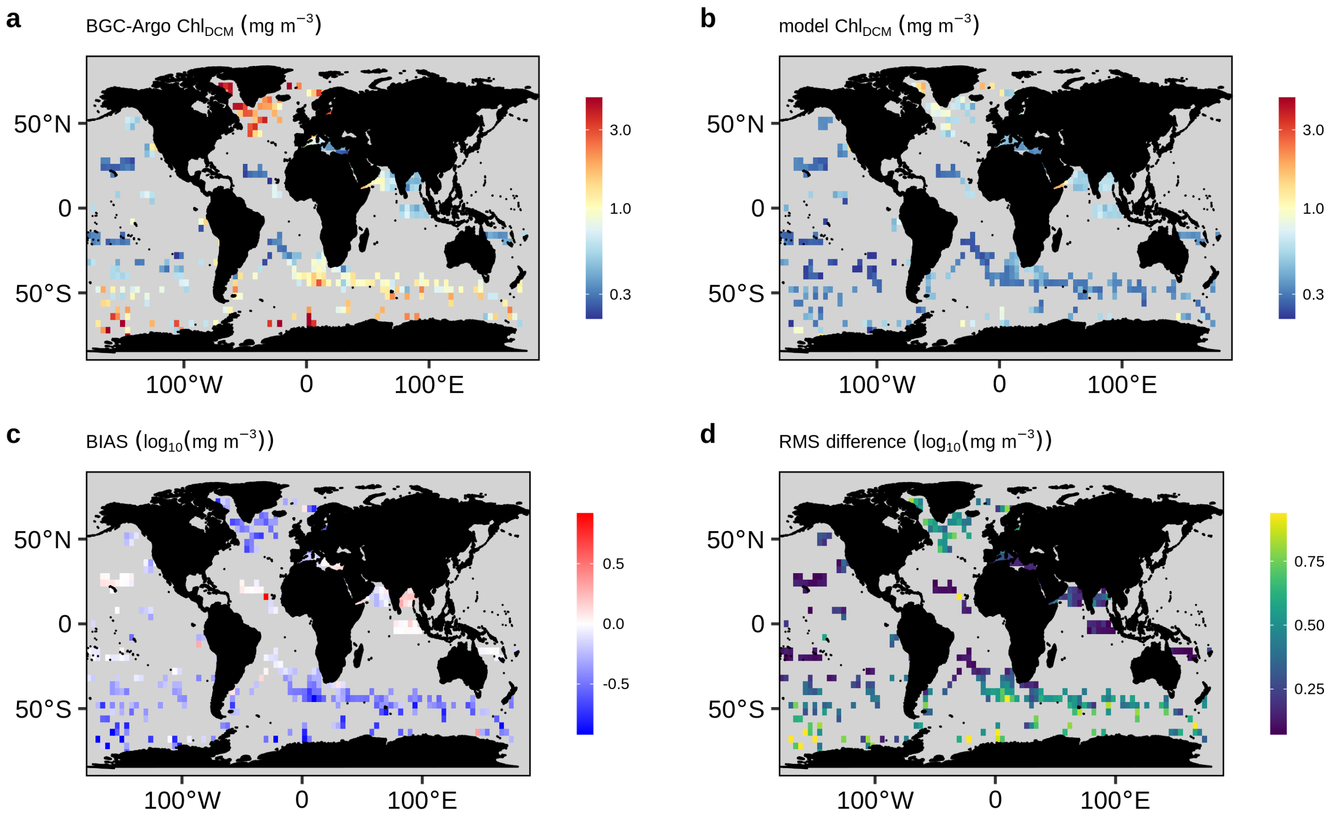


Figure A37. Same as Figure 4 but for Chl$_{DCM}$. Note that the BIAS and RMSD are computed on the log$_{10}$-transformed data to account that Chl$_{DCM}$ covers several orders of magnitude and it is lognormally distributed (Campbell, 1995).

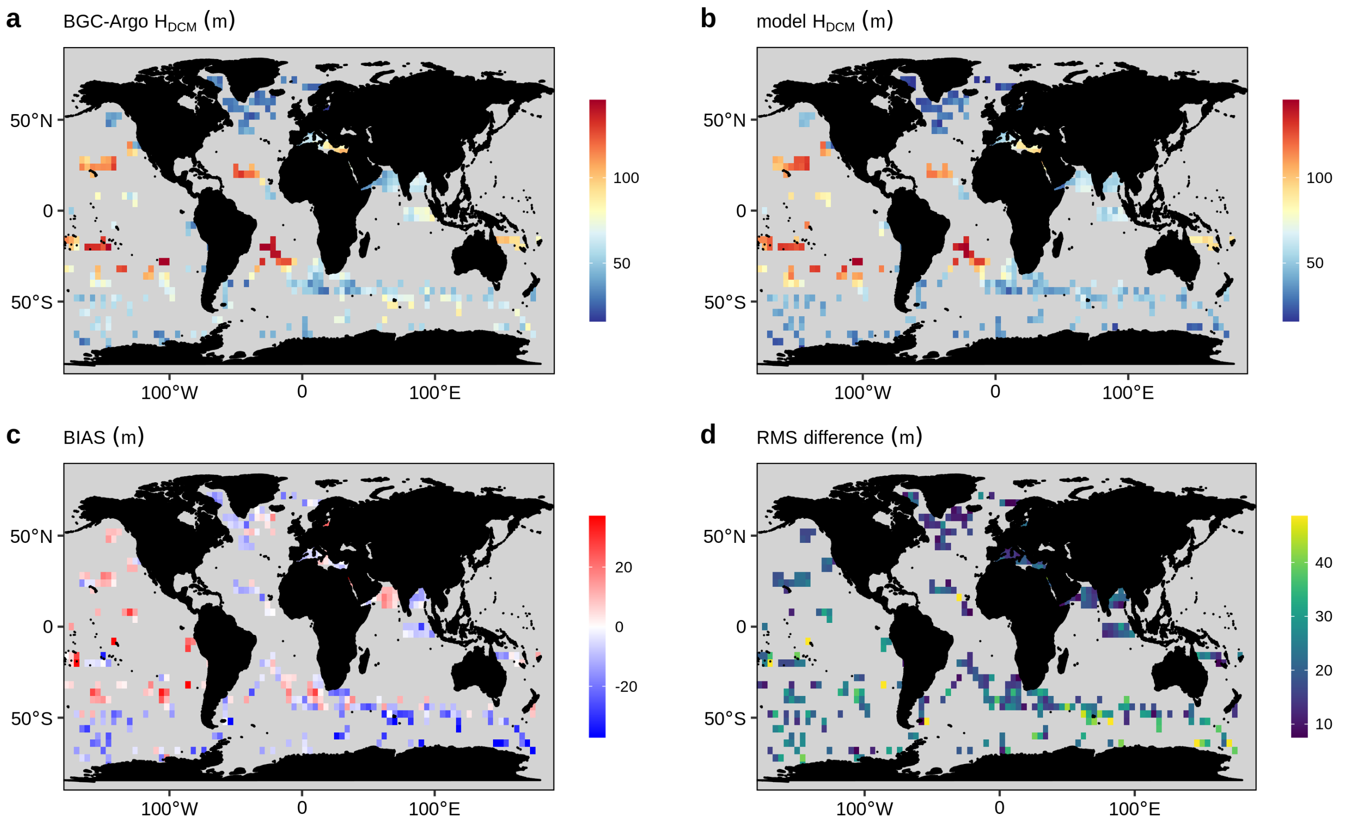


Figure A38. Same as Figure 4 but for H_{DCM}. Observed DCMs deeper than 250 m are not included.

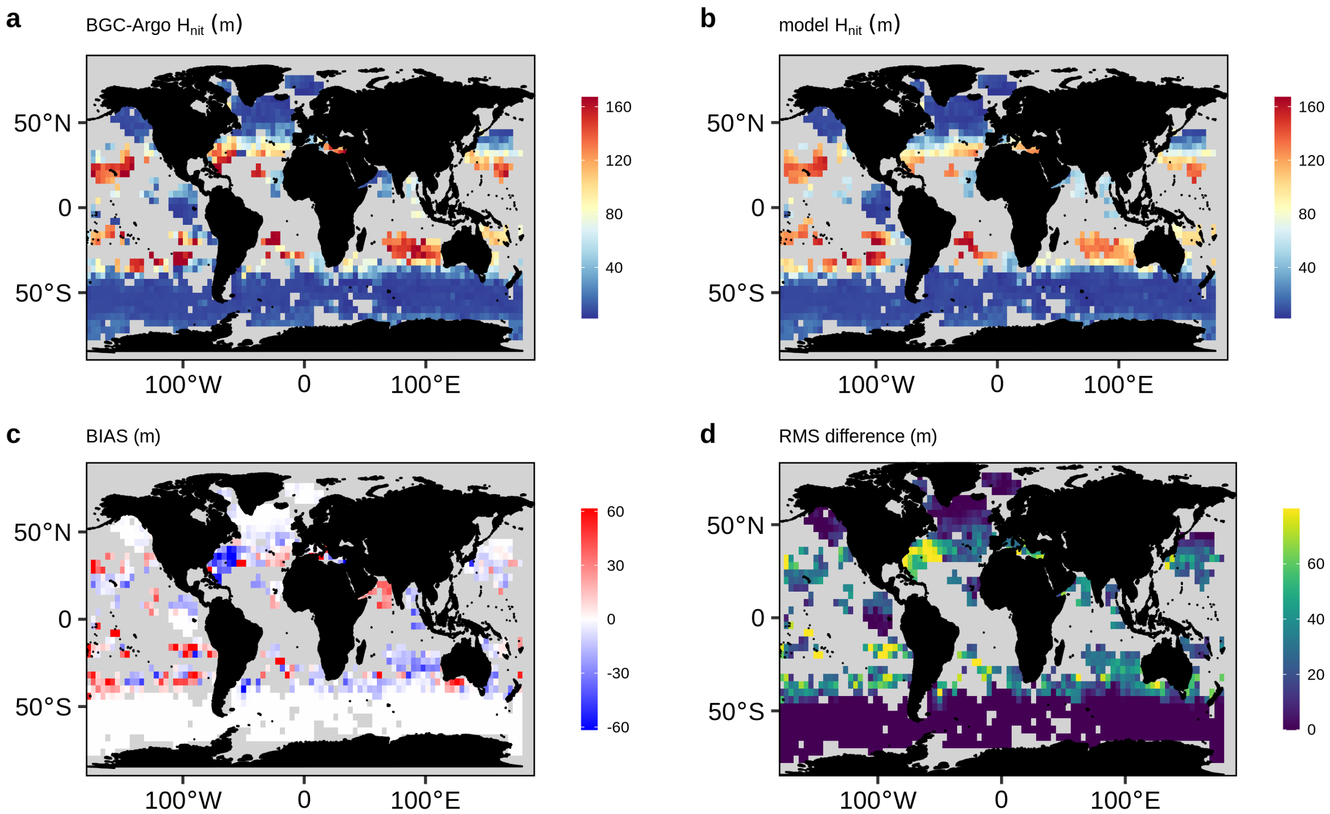


Figure A39. Same as Figure 4 but for H_{nit}. Observed nitracline deeper than 250 m are not included.

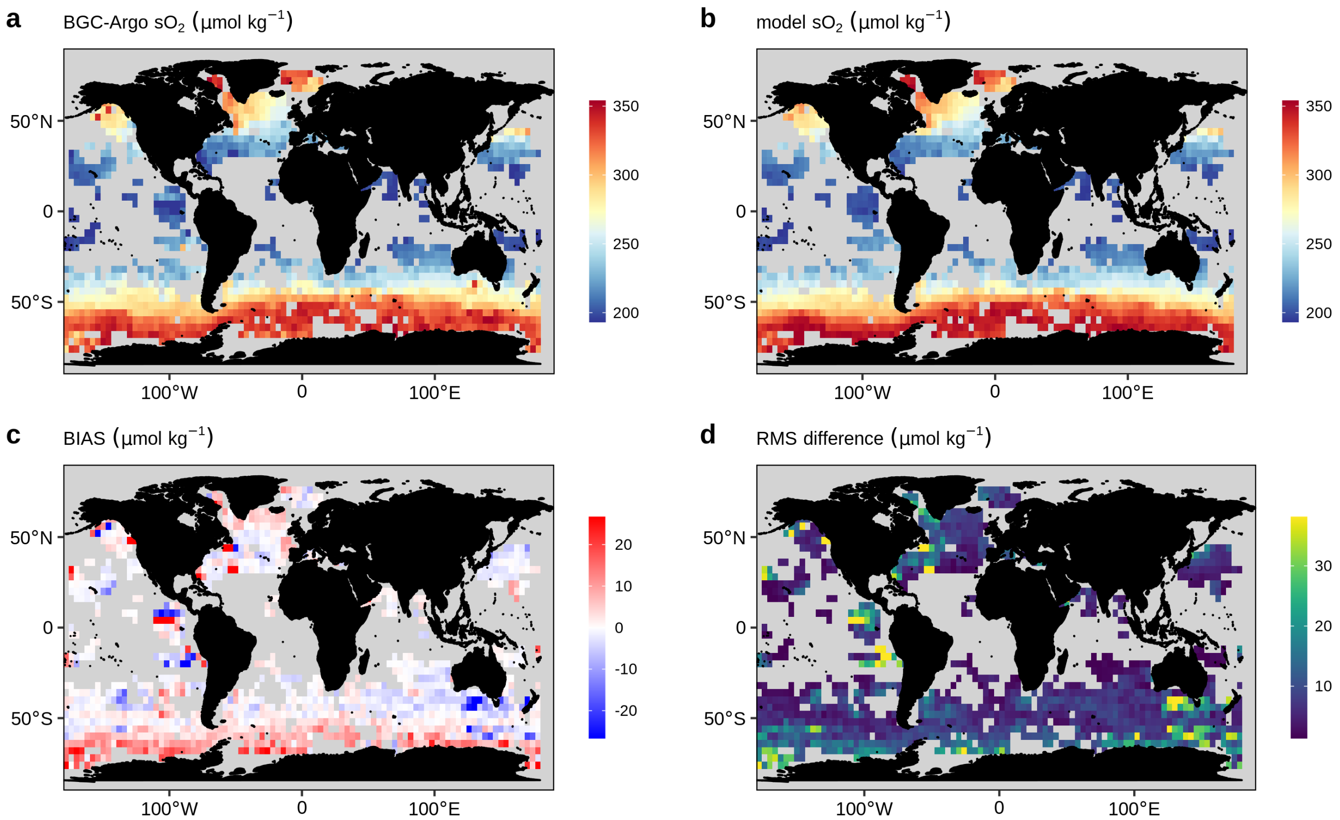


Figure A40. Same as Figure 4 but for sO_2.

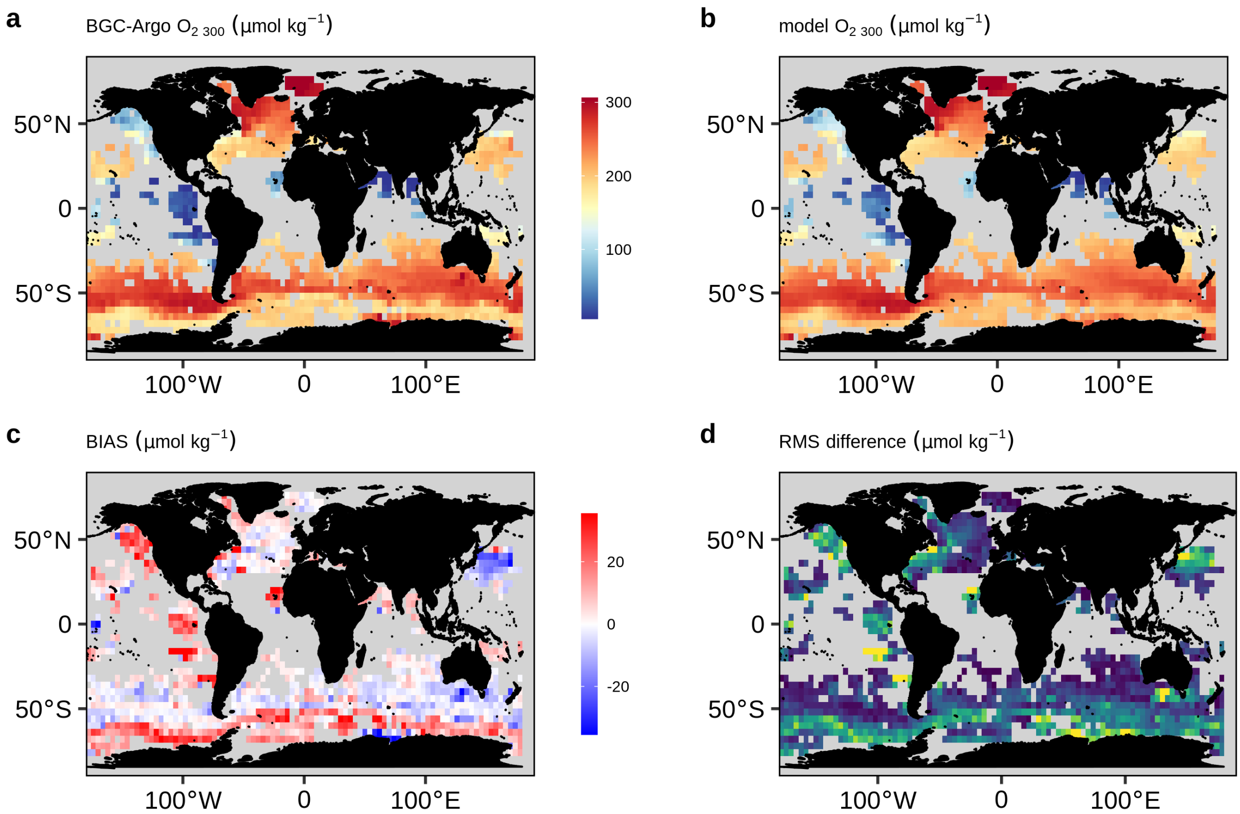


Figure A41. Same as Figure 4 but for $O_{2\ 300}$.

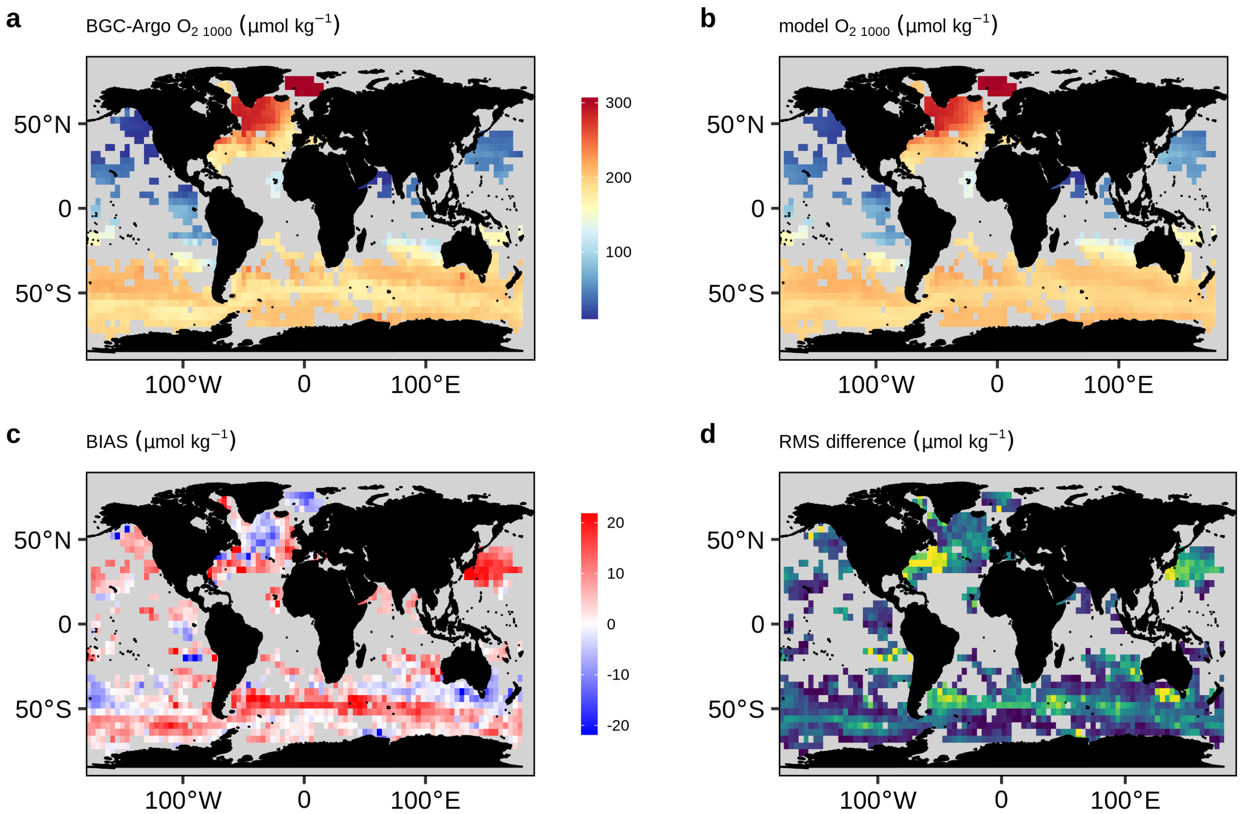


Figure A42. Same as Figure 4 but for $O_{2\ 1000}$.

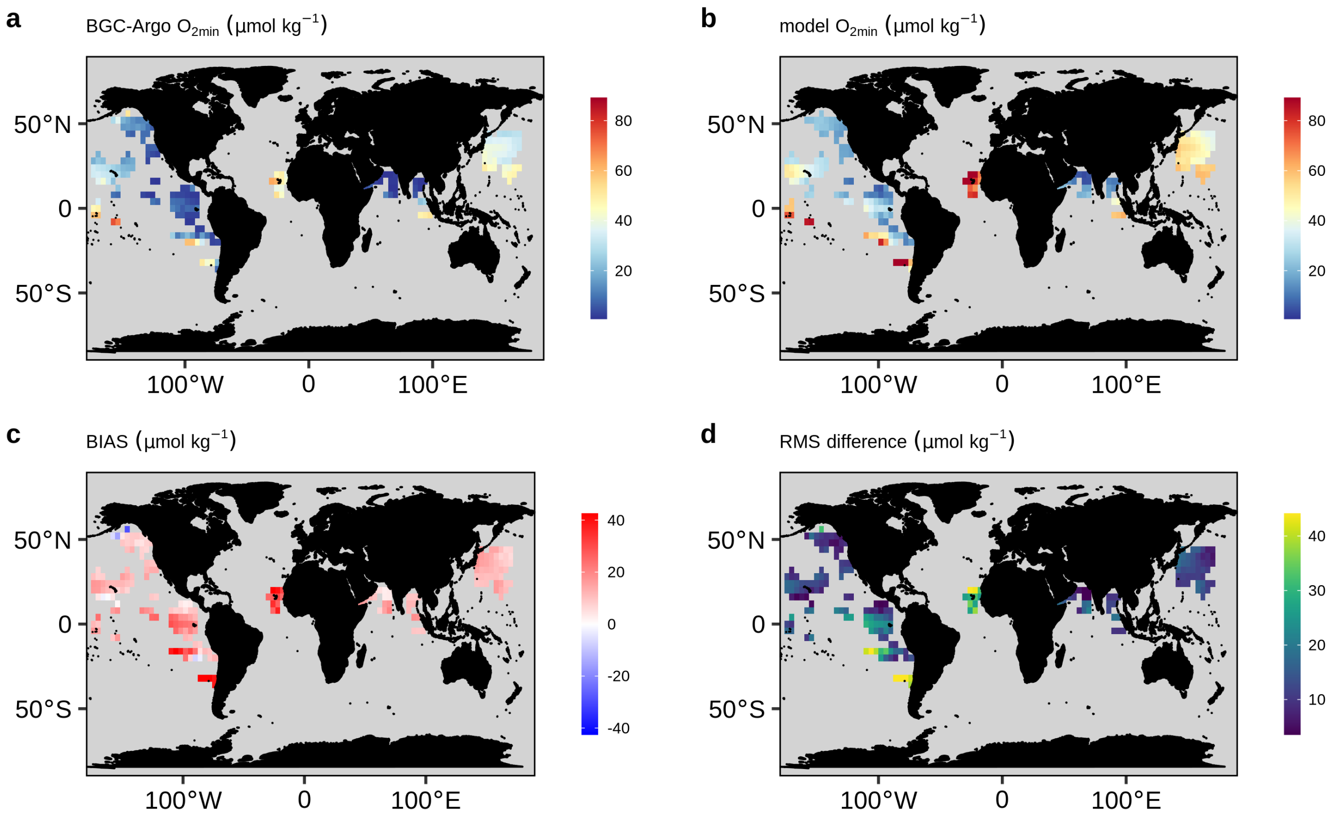


Figure A43. Same as Figure 4 but for O_{2min}.

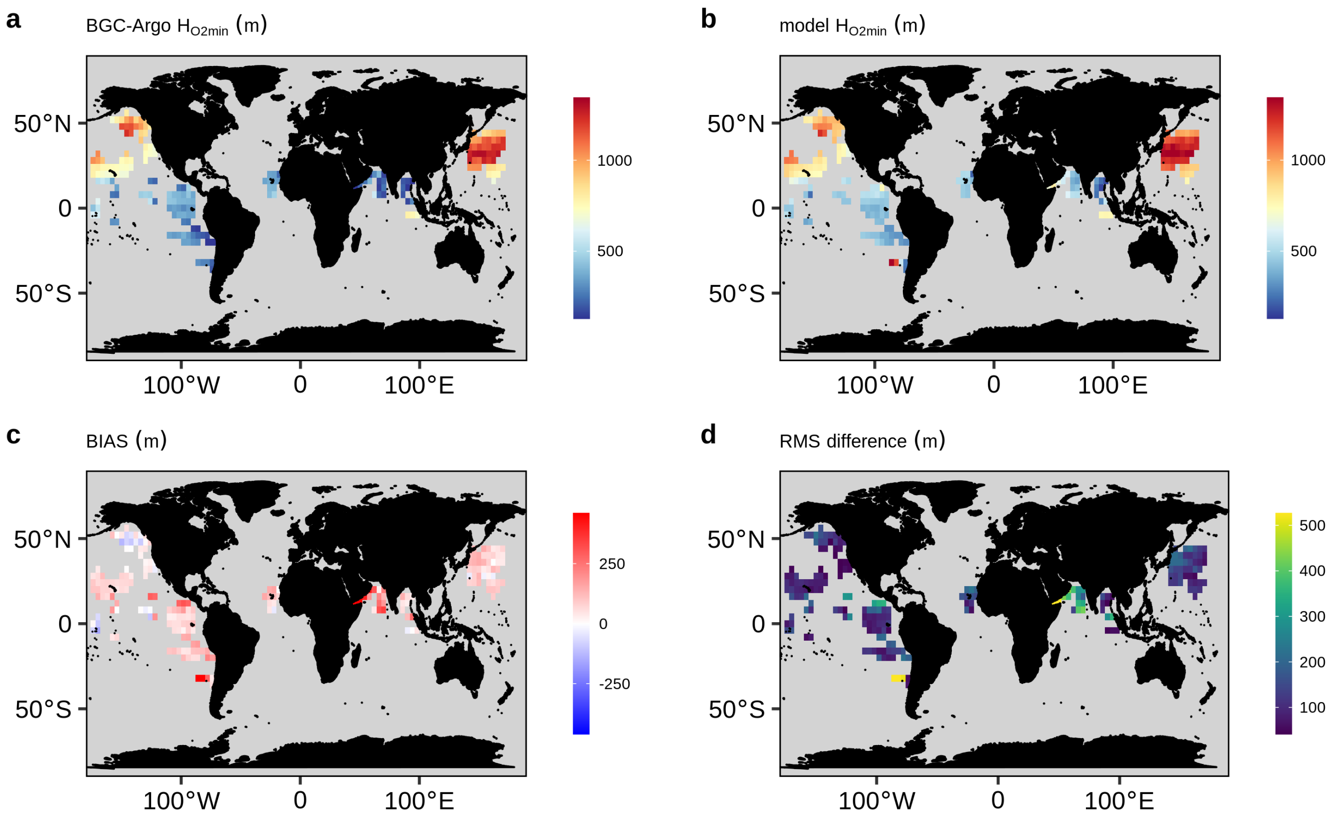


Figure A44. Same as Figure 4 but for H_{O2min}.

Data availability. The BGC model data can be downloaded from the Copernicus Marine Environmental Monitoring Service (https://resources.marine.copernicus.eu/?option=com_csw&view=details&product_id=GLOBAL_ANALYSIS_FORECAST_BIO_001_028). The BGC-Argo data were downloaded from the Argo Global Data Assembly Centre in France (ftp://ftp.ifremer.fr/argo/).

Authors Contribution: AM, GC, FD, SS and VT originated the study. AM, HC, FD, RS and VT designated the study. AM and RS process the BGC-Argo floats data. PL processed the BGC-Argo float in the Mediterranean Sea and run the Mediterranean BGC model.AM analysed the data. AM wrote the first draft of the manuscript. HC, GC, FD, EG, PL, CP, SS,RS,VT and AT contributed to the subsequent drafts. All authors read and approved the final draft.

Competing Interests: The authors declare no competing financial interests.

Materials and correspondence: Correspondence and request for material should be addressed to mignot@mercator-ocean.fr

Acknowledgements: This study has been conducted using the Copernicus Marine Service products (CMEMS). The BGC-Argo data were collected and made freely available by the International Argo program and the national programs that contribute to it (https://www.argo.jcommops. org). The Argo program is part of the Global Ocean Observing System. Part of this work was performed within the framework of the BIOOPTIMOD and MASSIMILI CMEMS Service Evolution Projects. This paper represents a contribution to the following research projects: NAOS (funded by the Agence Nationale de la Recherche in the framework of the French "Equipement d'avenir" program, grant ANR J11R107-F), remOcean (funded by the European Research Council, grant 246777), and the French Bio-Argo program (BGC-Argo France; funded by CNES-TOSCA, LEFE-GMMC).

References

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., and Gehlen, M.: PISCES-v2: an ocean biogeochemical model for carbon and ecosystem studies, Geosci. Model Dev., 8, 2465–2513, https://doi.org/10.5194/gmd-8-2465-2015, 2015.

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., and Hoepffner, N.: Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe, J. Geophys. Res. Oceans, 108, 2003.

Baird, M. E., Cherukuru, N., Jones, E., Margvelashvili, N., Mongin, M., Oubelkheir, K., Ralph, P. J., Rizwi, F., Robson, B. J., Schroeder, T., Skerratt, J., Steven, A. D. L., and Wild-Allen, K. A.: Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: Comparison with satellite data, Environ. Model. Softw., 78, 79–96, https://doi.org/10.1016/j.envsoft.2015.11.025, 2016.

Barbieux, M., Uitz, J., Gentili, B., Pasqueron de Fommervault, O., Mignot, A., Poteau, A., Schmechtig, C., Taillandier, V., Leymarie, E., Penkerc'h, C., D'Ortenzio, F., Claustre, H., and Bricaud, A.: Bio-optical characterization of subsurface chlorophyll maxima in the Mediterranean Sea from a Biogeochemical-Argo float database, Biogeosciences, 16, 1321–1342, https://doi.org/10.5194/bg-16-1321-2019, 2019.

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

Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., Körtzinger, A., and Gattuso, J.-P.: An alternative to static climatologies: robust estimation of open ocean CO2 variables and nutrient concentrations from T, S, and O2 data using Bayesian neural networks, Front. Mar. Sci., 5, 328, 2018.

Bittig, H. C., Maurer, T. L., Plant, J. N., Wong, A. P., Schmechtig, C., Claustre, H., Trull, T. W., Udaya Bhaskar, T. V. S., Boss, E., and Dall'Olmo, G.: A BGC-Argo guide: Planning, deployment, data handling and usage, Front. Mar. Sci., 6, 502, 2019.

Bopp, L., Aumont, O., Cadule, P., Alvain, S., and Gehlen, M.: Response of diatoms distribution to global warming and potential implications: A global model study, Geophys. Res. Lett., 32, https://doi.org/10.1029/2005GL023653, 2005.

Bosc, E., Bricaud, A., and Antoine, D.: Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations, Glob. Biogeochem. Cycles, 18, https://doi.org/10.1029/2003GB002034, 2004.

Breitburg, D., Levin, L. A., Oschlies, A., Grégoire, M., Chavez, F. P., Conley, D. J., Garçon, V., Gilbert, D., Gutiérrez, D., Isensee, K., Jacinto, G. S., Limburg, K. E., Montes, I., Naqvi, S. W. A., Pitcher, G. C., Rabalais, N. N., Roman, M. R., Rose, K. A., Seibel, B. A., Telszewski, M., Yasuhara, M., and Zhang, J.: Declining oxygen in the global ocean and coastal waters, Science, 359, https://doi.org/10.1126/science.aam7240, 2018.

Briggs, N., Perry, M. J., Cetinić, I., Lee, C., D'Asaro, E., Gray, A. M., and Rehm, E.: High-resolution observations of aggregate flux during a sub-polar North Atlantic spring bloom, Deep Sea Res. Part Oceanogr. Res. Pap., 58, 1031–1039, https://doi.org/10.1016/j.dsr.2011.07.007, 2011.

Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, J. Geophys. Res. Oceans, 100, 13237–13254, https://doi.org/10.1029/95JC00458, 1995.

Canu, D. M., Ghermandi, A., Nunes, P. A., Lazzari, P., Cossarini, G., and Solidoro, C.: Estimating the value of carbon sequestration ecosystem services in the Mediterranean Sea: An ecological economics approach, Glob. Environ. Change, 32, 87–95, 2015.

Capuzzo, E., Lynam, C. P., Barry, J., Stephens, D., Forster, R. M., Greenwood, N., McQuatters-Gollop, A., Silva, T., van Leeuwen, S. M., and Engelhard, G. H.: A decline in primary production in the North Sea over 25 years, associated with reductions in zooplankton abundance and fish stock recruitment, Glob. Change Biol., 24, e352–e364, https://doi.org/10.1111/gcb.13916, 2018.

Cermeno, P., Dutkiewicz, S., Harris, R. P., Follows, M., Schofield, O., and Falkowski, P. G.: The role of nutricline depth in regulating the ocean carbon cycle, Proc. Natl. Acad. Sci., 105, 20344–20349, https://doi.org/10.1073/pnas.0811302106, 2008.

Claustre, H., Johnson, K. S., and Takeshita, Y.: Observing the Global Ocean with Biogeochemical-Argo, Annu. Rev. Mar. Sci., 12, annurev-marine-010419-010956, https://doi.org/10.1146/annurev-marine-010419-010956, 2020.

Cossarini, G., Lazzari, P., and Solidoro, C.: Spatiotemporal variability of alkalinity in the Mediterranean Sea, Biogeosciences, 12, 1647–1658, https://doi.org/10.5194/bg-12-1647-2015, 2015.

Cossarini, G., Mariotti, L., Feudale, L., Mignot, A., Salon, S., Taillandier, V., Teruzzi, A., and D'Ortenzio, F.: Towards operational 3D-Var assimilation of chlorophyll Biogeochemical-Argo float data into a biogeochemical model of the Mediterranean Sea, Ocean Model., 133, 112–128, https://doi.org/10.1016/j.ocemod.2018.11.005, 2019.

Crowder, L. B., Hazen, E. L., Avissar, N., Bjorkland, R., Latanich, C., and Ogburn, M. B.: The Impacts of Fisheries on Marine Ecosystems and the Transition to Ecosystem-Based Management, Annu. Rev. Ecol. Evol. Syst., 39, 259–278, https://doi.org/10.1146/annurev.ecolsys.39.110707.173406, 2008.

Cullen, J. J.: Subsurface Chlorophyll Maximum Layers: Enduring Enigma or Mystery Solved?, Annu. Rev. Mar. Sci., 7, 207–239, https://doi.org/10.1146/annurev-marine-010213-135111, 2015.

Dale, T., Rey, F., and Heimdal, B. R.: Seasonal development of phytoplankton at a high latitude oceanic site, Sarsia, 84, 419–435, 1999.

Dall'Olmo, G. and Mork, K. A.: Carbon export by small particles in the Norwegian Sea, Geophys. Res. Lett., 41, 2921–2927, https://doi.org/10.1002/2014GL059244, 2014.

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

D'Ortenzio, F., Lavigne, H., Besson, F., Claustre, H., Coppola, L., Garcia, N., Laes-Huon, A., Le Reste, S., Malarde, D., Migon, C., Morin, P., Mortier, L., Poteau, A., Prieur, L., Raimbault, P., and Testor, P.: Observing mixed layer depth, nitrate and chlorophyll concentrations in the northwestern Mediterranean: A combined satellite and NO3 profiling floats experiment, Geophys. Res. Lett., 41, 2014GL061020, https://doi.org/10.1002/2014GL061020, 2014.

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J.: Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model, Biogeosciences, 12, 4447–4481, https://doi.org/10.5194/bg-12-4447-2015, 2015.

Eriksen, M., Lebreton, L. C. M., Carson, H. S., Thiel, M., Moore, C. J., Borerro, J. C., Galgani, F., Ryan, P. G., and Reisser, J.: Plastic Pollution in the World's Oceans: More than 5 Trillion Plastic Pieces Weighing over 250,000 Tons Afloat at Sea, PLoS ONE, 9, e111913, https://doi.org/10.1371/journal.pone.0111913, 2014.

Evans, G. T. and Parslow, J. S.: A Model of Annual Plankton Cycles, Biol. Oceanogr., 3, 327–347, https://doi.org/10.1080/01965581.1985.10749478, 1985.

Evers-King, H., Martinez-Vicente, V., Brewin, R. J. W., Dall'Olmo, G., Hickman, A. E., Jackson, T., Kostadinov, T. S., Krasemann, H., Loisel, H., Röttgers, R., Roy, S., Stramski, D., Thomalla, S., Platt, T., and Sathyendranath, S.: Validation and Intercomparison of Ocean Color Algorithms for Estimating Particulate Organic Carbon in the Oceans, Front. Mar. Sci., 4, 251, https://doi.org/10.3389/fmars.2017.00251, 2017.

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., and the GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team: Advancing Marine Biogeochemical and Ecosystem Reanalyses and Forecasts as Tools for Monitoring and Managing Ecosystem Health, Front. Mar. Sci., 6, 89, 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, Biogeochemistry: Open Ocean, https://doi.org/10.5194/bg-2020-152, 2020.

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Hauck, J., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Le Quéré, C., Bakker, D. C. E., Canadell, J. G., Ciais, P., Jackson, R. B., Anthoni, P., Barbero, L., Bastos, A., Bastrikov, V., Becker, M., Bopp, L., Buitenhuis, E., Chandra, N., Chevallier, F., Chini, L. P., Currie, K. I., Feely, R. A., Gehlen, M., Gilfillan, D., Gkritzalis, T., Goll, D. S., Gruber, N., Gutekunst, S., Harris, I., Haverd, V., Houghton, R. A., Hurtt, G., Ilyina, T., Jain, A. K., Joetzjer, E., Kaplan, J. O., Kato, E., Klein Goldewijk, K., Korsbakken, J. I., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lienert, S., Lombardozzi, D., Marland, G., McGuire, P. C., Melton, J. R., Metzl, N., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Neill, C., Omar, A. M., Ono, T., Peregon, A., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Séférian, R., Schwinger, J., Smith, N., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F. N., van der Werf,

G. R., Wiltshire, A. J., and Zaehle, S.: Global Carbon Budget 2019, Earth Syst. Sci. Data, 11, 1783–1838, https://doi.org/10.5194/essd-11-1783-2019, 2019.

Galí, M., Falls, M., Claustre, H., Aumont, O., and Bernardello, R.: Bridging the gaps between particulate backscattering measurements and modeled particulate organic carbon in the ocean, Biogeochemistry: Open Ocean, https://doi.org/10.5194/bg-2021-201, 2021.

Gasparin, F., Cravatte, S., Greiner, E., Perruche, C., Hamon, M., Van Gennip, S., and Lellouche, J.-M.: Excessive productivity and heat content in tropical Pacific analyses: Disentangling the effects of in situ and altimetry assimilation, Ocean Model., 160, 101768, https://doi.org/10.1016/j.ocemod.2021.101768, 2021.

Gehlen, M., Bopp, L., Emprin, N., Aumont, O., Heinze, C., and Ragueneau, O.: Reconciling surface ocean productivity, export fluxes and sediment composition in a global biogeochemical ocean model, Biogeosciences, 3, 521–537, https://doi.org/10.5194/bg-3-521-2006, 2006.

Gehlen, M., Gangstø, R., Schneider, B., Bopp, L., Aumont, O., and Ethe, C.: The fate of pelagic $CaCO_3$ production in a high CO_2 ocean: a model study, Biogeosciences, 4, 505–519, https://doi.org/10.5194/bg-4-505-2007, 2007.

Gittings, J. A., Raitsos, D. E., Kheireddine, M., Racault, M.-F., Claustre, H., and Hoteit, I.: Evaluating tropical phytoplankton phenology metrics using contemporary tools, Sci. Rep., 9, 1–9, 2019.

Gregg, W. W. and Rousseaux, C. S.: Directional and spectral irradiance in ocean models: effects on simulated global phytoplankton, nutrients, and primary production, Front. Mar. Sci., 3, 240, 2016.

Gutknecht, E., Reffray, G., Mignot, A., Dabrowski, T., and Sotillo, M. G.: Modelling the marine ecosystem of Iberia-Biscay-Ireland (IBI) European waters for CMEMS operational applications, Ocean Sci., 15, 1489–1516, https://doi.org/10.5194/os-15-1489-2019, 2019.

Hipsey, M. R., Gal, G., Arhonditsis, G. B., Carey, C. C., Elliott, J. A., Frassl, M. A., Janse, J. H., de Mora, L., and Robson, B. J.: A system of metrics for the assessment and improvement of aquatic ecosystem models, Environ. Model. Softw., 128, 104697, https://doi.org/10.1016/j.envsoft.2020.104697, 2020.

Iida, Y., Takatani, Y., Kojima, A., and Ishii, M.: Global trends of ocean CO 2 sink and ocean acidification: an observation-based reconstruction of surface ocean inorganic carbon variables, J. Oceanogr., 1–36, 2020.

Johnson, Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., Swift, D. D., Williams, N. L., Boss, E., Haëntjens, N., Talley, L. D., and Sarmiento, J. L.: Biogeochemical sensor performance in the SOCCOM profiling float array: SOCCOM BIOGEOCHEMICAL SENSOR PERFORMANCE, J. Geophys. Res. Oceans, 122, 6416–6436, https://doi.org/10.1002/2017JC012838, 2017.

Johnson, Plant, J. N., and Maurer, T. L.: Processing BGC-Argo pH data at the DAC level, 2018a.

Johnson, Pasqueron De Fommervault, O., Serra, R., D'Ortenzio, F., Schmechtig, C., Claustre, H., and Poteau, A.: Processing Bio-Argo nitrate concentration at the DAC Level, 2018b.

Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., and Hoppema, M.: Global Ocean Data Analysis Project, Version 2 (GLODAPv2), Carbon Dioxide Information Analysis Center, Oak Ridge Nat Lab, 2015.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., Dunne, J. P., Gehlen, M., Ilyina, T., John, J. G., Lenton, A., Li, H., Lovenduski, N. S., Orr, J. C., Palmieri, J., Santana-Falcón, Y., Schwinger, J., Séférian, R., Stock, C. A., Tagliabue, A., Takano, Y., Tjiputra, J., Toyama, K., Tsujino, H., Watanabe, M., Yamamoto, A., Yool, A., and Ziehn, T.: Twenty-first century ocean warming, acidification, deoxygenation, and upper-ocean nutrient and primary production decline from CMIP6 model projections, Biogeosciences, 17, 3439–3470, https://doi.org/10.5194/bg-17-3439-2020, 2020.

Lavigne, H., D'Ortenzio, F., Migon, C., Claustre, H., Testor, P., d'Alcalà, M. R., Lavezza, R., Houpert, L., and Prieur, L.: Enhancing the comprehension of mixed layer depth control on the Mediterranean phytoplankton phenology: Mediterranean Phytoplankton Phenology, J. Geophys. Res. Oceans, 118, 3416–3430, https://doi.org/10.1002/jgrc.20251, 2013.

Lazzari, Solidoro, C., Ibello, V., Salon, S., Teruzzi, A., Béranger, K., Colella, S., and Crise, A.: Seasonal and inter-annual variability of plankton chlorophyll and primary production in the Mediterranean Sea: a modelling approach, Biogeosciences, 9, 217–233, https://doi.org/10.5194/bg-9-217-2012, 2012.

Lazzari, Solidoro, C., Salon, S., and Bolzon, G.: Spatial variability of phosphate and nitrate in the Mediterranean Sea: A modeling approach, Deep Sea Res. Part Oceanogr. Res. Pap., 108, 39–52, https://doi.org/10.1016/j.dsr.2015.12.006, 2016.

Lazzari, Salon, S., Terzić, E., Gregg, W. W., D'Ortenzio, F., Vellucci, V., Organelli, E., and Antoine, D.: Assessment of the spectral downward irradiance at the surface of theMediterranean Sea using the OASIM ocean-atmosphere radiative model, Surface/Numerical Models/Mediterranean Sea/Air-sea fluxes/Oceanic ecosystems, https://doi.org/10.5194/os-2020-108, 2020.

Lefèvre, N., Veleda, D., Tyaquiçã, P., Perruche, C., Diverrès, D., and Ibánhez, J. S. P.: Basin-Scale Estimate of the Sea-Air CO_2 Flux During the 2010 Warm Event in the Tropical North Atlantic, J. Geophys. Res. Biogeosciences, 124, 973–986, https://doi.org/10.1029/2018JG004840, 2019.

Lellouche, Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., Benkiran, M., Testut, C.-E., Bourdalle-Badie, R., Gasparin, F., Hernandez, O., Levier, B., Drillet, Y., Remy, E., and Le Traon, P.-Y.: Recent updates to the Copernicus Marine Service global ocean monitoring and forecasting real-time 1⁄12° high-resolution system, Ocean Sci., 14, 1093–1126, https://doi.org/10.5194/os-14-1093-2018, 2018.

Lellouche, J.-M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., Ferry, N., Desportes, C., Testut, C.-E., Bricaud, C., Bourdallé-Badie, R., Tranchant, B., Benkiran, M., Drillet, Y., Daudin, A., and De Nicola, C.: Evaluation of global monitoring and forecasting systems at Mercator Océan, Ocean Sci., 9, 57–81, https://doi.org/10.5194/os-9-57-2013, 2013.

Letelier, R. M., Karl, D. M., Abbott, M. R., and Bidigare, R. R.: Light driven seasonal patterns of chlorophyll and nitrate in the lower euphotic zone of the North Pacific Subtropical Gyre, Limnol. Oceanogr., 49, 508–519, 2004.

Lynch, D. R., McGillicuddy, D. J., and Werner, F. E.: Skill assessment for coupled biological/physical models of marine systems, J. Mar. Syst., 1, 1–3, 2009.

Macías, D., Stips, A., and Garcia-Gorriz, E.: The relevance of deep chlorophyll maximum in the open Mediterranean Sea evaluated through 3D hydrodynamic-biogeochemical coupled simulations, Ecol. Model., 281, 26–37, 2014.

Mignot, Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., and Xing, X.: Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: A Bio-Argo float investigation, Glob. Biogeochem. Cycles, 28, 856–876, https://doi.org/10.1002/2013GB004781, 2014.

Mignot, Ferrari, R., and Claustre, H.: Floats with bio-optical sensors reveal what processes trigger the North Atlantic bloom, Nat. Commun., 9, https://doi.org/10.1038/s41467-017-02143-6, 2018.

Mignot, A., Claustre, H., D'Ortenzio, F., Xing, X., Poteau, A., and Ras, J.: From the shape of the vertical profile of in vivo fluorescence to Chlorophyll-a concentration, Biogeosciences, 8, 2391–2406, https://doi.org/10.5194/bg-8-2391-2011, 2011.

Mignot, A., D'Ortenzio, F., Taillandier, V., Cossarini, G., and Salon, S.: Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll *a* Concentrations, Geophys. Res. Lett., 46, 4330–4337, https://doi.org/10.1029/2018GL080541, 2019.

Omand, M. M. and Mahadevan, A.: The shape of the oceanic nitracline, Biogeosciences, 12, 3273–3287, https://doi.org/10.5194/bg-12-3273-2015, 2015.

Organelli, E., Bricaud, A., Antoine, D., and Matsuoka, A.: Seasonal dynamics of light absorption by chromophoric dissolved organic matter (CDOM) in the NW Mediterranean Sea (BOUSSOLE site), Deep Sea Res. Part Oceanogr. Res. Pap., 91, 72–85, 2014.

Organelli, E., Barbieux, M., Claustre, H., Schmechtig, C., Poteau, A., Bricaud, A., Boss, E. B., Briggs, N., Dall'Olmo, G., and d'Ortenzio, F.: Two databases derived from BGC-Argo float measurements for marine biogeochemical and bio-optical applications, Earth Syst. Sci. Data, 9, 861–880, 2017.

Osman, M. B., Das, S. B., Trusel, L. D., Evans, M. J., Fischer, H., Grieman, M. M., Kipfstuhl, S., McConnell, J. R., and Saltzman, E. S.: Industrial-era decline in subarctic Atlantic productivity, Nature, 569, 551–555, https://doi.org/10.1038/s41586-019-1181-8, 2019.

Park, J.-Y., Stock, C. A., Yang, X., Dunne, J. P., Rosati, A., John, J., and Zhang, S.: Modeling Global Ocean Biogeochemistry With Physical Data Assimilation: A Pragmatic Solution to the Equatorial Instability, J. Adv. Model. Earth Syst., 10, 891–906, https://doi.org/10.1002/2017MS001223, 2018.

Paulmier, A. and Ruiz-Pino, D.: Oxygen minimum zones (OMZs) in the modern ocean, Prog. Oceanogr., 80, 113–128, 2009.

Plant, J. N., Johnson, K. S., Sakamoto, C. M., Jannasch, H. W., Coletti, L. J., Riser, S. C., and Swift, D. D.: Net community production at Ocean Station Papa observed with nitrate and oxygen sensors on profiling floats, Glob. Biogeochem. Cycles, 30, 859–879, https://doi.org/10.1002/2015GB005349, 2016.

Richardson, K. and Bendtsen, J.: Vertical distribution of phytoplankton and primary production in relation to nutricline depth in the open ocean, Mar. Ecol. Prog. Ser., 620, 33–46, https://doi.org/10.3354/meps12960, 2019.

Riley, G.: Factors Controlling Phytoplankton Populations on Georges Bank, J. Mar. Res., 6, 54–73, 1946.

Roxy, M. K., Modi, A., Murtugudde, R., Valsala, V., Panickal, S., Prasanna Kumar, S., Ravichandran, M., Vichi, M., and Lévy, M.: A reduction in marine primary productivity driven by rapid warming over the tropical Indian Ocean, Geophys. Res. Lett., 43, 826–833, https://doi.org/10.1002/2015GL066979, 2016.

Salon, S., Cossarini, G., Bolzon, G., Feudale, L., Lazzari, P., Teruzzi, A., Solidoro, C., and Crise, A.: Novel metrics based on Biogeochemical Argo data to improve the model uncertainty evaluation of the CMEMS Mediterranean marine ecosystem forecasts, Ocean Sci., 15, 997–1022, https://doi.org/10.5194/os-15-997-2019, 2019.

Sauzède, R., Bittig, H. C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J.-P., Legendre, L., and Johnson, K. S.: Estimates of Water-Column Nutrient Concentrations and Carbonate System Parameters in the Global Ocean: A Novel Approach Based on Neural Networks, Front. Mar. Sci., 4, https://doi.org/10.3389/fmars.2017.00128, 2017.

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.

Schmechtig, C., Poteau, A., Claustre, H., D’Ortenzio, F., and Boss, E.: Processing bio-Argo chlorophyll-A concentration at the DAC level, Ifremer, https://doi.org/10.13155/39468, 2015.

Schmechtig, C., Claustre, H., Poteau, A., and D’Ortenzio, F.: Bio-Argo quality control manual for the Chlorophyll-A concentration, Ifremer, https://doi.org/10.13155/35385, 2018.

Schmidtko, S., Stramma, L., and Visbeck, M.: Decline in global oceanic oxygen content during the past five decades, Nature, 542, 335–339, https://doi.org/10.1038/nature21399, 2017.

Schneider, B., Bopp, L., Gehlen, M., Segschneider, J., Frölicher, T. L., Cadule, P., Friedlingstein, P., Doney, S. C., Behrenfeld, M. J., and Joos, F.: Climate-induced interannual variability of marine primary and export production in three global coupled climate carbon cycle models, Biogeosciences, 5, 597–614, https://doi.org/10.5194/bg-5-597-2008, 2008.

Séférian, R., Bopp, L., Gehlen, M., Orr, J. C., Ethé, C., Cadule, P., Aumont, O., Salas y Mélia, D., Voldoire, A., and Madec, G.: Skill assessment of three earth system models with common marine biogeochemistry, Clim. Dyn., 40, 2549–2573, https://doi.org/10.1007/s00382-012-1362-8, 2013.

Skákala, J., Bruggeman, J., Brewin, R. J. W., Ford, D. A., and Ciavatta, S.: Improved Representation of Underwater Light Field and Its Impact on Ecosystem Dynamics: A Study in the North Sea, J. Geophys. Res. Oceans, 125, https://doi.org/10.1029/2020JC016122, 2020.

Snowden, D., Tsontos, V. M., Handegard, N. O., Zarate, M., O' Brien, K., Casey, K. S., Smith, N., Sagen, H., Bailey, K., Lewis, M. N., and Arms, S. C.: Data Interoperability Between Elements of the Global Ocean Observing System, Front. Mar. Sci., 6, 442, https://doi.org/10.3389/fmars.2019.00442, 2019.

Sosik, H. M.: Characterizing seawater constituents from optical properties, Real-Time Coast. Obs. Syst. Ecosyst. Dyn. Harmful Algal Blooms Ed. Babin M Roesler CS Cullen JJ UNESCO, 281–329, 2008.

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Cocco, V., Doney, S. C., Gehlen, M., Lindsay, K., Moore, J. K., Schneider, B., and Segschneider, J.: Projected 21st century decrease in marine productivity: a multi-model analysis, Biogeosciences, 7, 979–1005, https://doi.org/10.5194/bg-7-979-2010, 2010.

Stow, C. A., Jolliff, J., McGillicuddy, D. J., Doney, S. C., Allen, J. I., Friedrichs, M. A. M., Rose, K. A., and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, J. Mar. Syst., 76, 4–15, https://doi.org/10.1016/j.jmarsys.2008.03.011, 2009.

Stramma, L., Johnson, G. C., Sprintall, J., and Mohrholz, V.: Expanding Oxygen-Minimum Zones in the Tropical Oceans, Science, 320, 655–658, https://doi.org/10.1126/science.1153847, 2008.

Tagliabue, A., Bopp, L., Dutay, J.-C., Bowie, A. R., Chever, F., Jean-Baptiste, P., Bucciarelli, E., Lannuzel, D., Remenyi, T., Sarthou, G., Aumont, O., Gehlen, M., and Jeandel, C.: Hydrothermal contribution to the oceanic dissolved iron inventory, Nat. Geosci., 3, 252–256, https://doi.org/10.1038/ngeo818, 2010.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res. Atmospheres, 106, 7183–7192, https://doi.org/10.1029/2000JD900719, 2001.

Teruzzi, A., Dobricic, S., Solidoro, C., and Cossarini, G.: A 3-D variational assimilation scheme in coupled transport-biogeochemical models: Forecast of Mediterranean biogeochemical properties: 3D-VAR IN BIOGEOCHEMICAL MODELS, J. Geophys. Res. Oceans, 119, 200–217, https://doi.org/10.1002/2013JC009277, 2014.

Terzić, E., Lazzari, P., Organelli, E., Solidoro, C., Salon, S., D'Ortenzio, F., and Conan, P.: Merging bio-optical data from Biogeochemical-Argo floats and models in marine biogeochemistry, Biogeosciences, 16, 2527–2542, https://doi.org/10.5194/bg-16-2527-2019, 2019.

Thierry, V. and Bittig, H.: Argo quality control manual for dissolved oxygen concentration, 2018.

Thierry, V., Bittig, H., Gilbert, D., Kobayashi, T., Kanako, S., and Schmid, C.: Processing Argo oxygen data at the DAC level, Ifremer, https://doi.org/10.13155/39795, 2018.

Tuan Pham, D., Verron, J., and Christine Roubaud, M.: A singular evolutive extended Kalman filter for data assimilation in oceanography, J. Mar. Syst., 16, 323–340, https://doi.org/10.1016/S0924-7963(97)00109-7, 1998.

Vichi, M., Lovato, T., Lazzari, P., Cossarini, G., Gutierrez, E., Mattia, G., Masina, S., McKiver, W. J., Pinardi, N., and Solidoro, C.: The Biogeochemical Flux Model (BFM): Equation Description and User Manual, BFM version 5.1, BFM Report series N. 1, Release 1.1, July 2015, Bologna, Italy, 104pp, 2015.

Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean revisited: Gas exchange and wind speed over the ocean, Limnol. Oceanogr. Methods, 12, 351–362, https://doi.org/10.4319/lom.2014.12.351, 2014.

Ward, B. A., Friedrichs, M. A. M., Anderson, T. R., and Oschlies, A.: Parameter optimisation techniques and the problem of underdetermination in marine biogeochemical models, J. Mar. Syst., 81, 34–43, https://doi.org/10.1016/j.jmarsys.2009.12.005, 2010.

Williams, R. G. and Follows, M. J.: Ocean dynamics and the carbon cycle: Principles and mechanisms, Cambridge University Press, 2011.

Wong, Keeley, Robert, Carval, Thierry, and Argo Data Management Team,: Argo Quality Control Manual for CTD and Trajectory Data, https://doi.org/10.13155/33951, 2015.

Yang, B., Fox, J., Behrenfeld, M. J., Boss, E. S., Haëntjens, N., Halsey, K. H., Emerson, S. R., and Doney, S. C.: In Situ Estimates of Net Primary Production in the Western North Atlantic With Argo Profiling Floats, J. Geophys. Res. Biogeosciences, 126, https://doi.org/10.1029/2020JG006116, 2021.