Geoscientific
Model Development
Discussions

Using synthetic float capabilities in E3SMv2 to assess spatio-temporal variability in ocean physics and biogeochemistry

Cara Nissen[1], Nicole S. Lovenduski[1], Mathew Maltrud[2], Alison R. Gray[3], Yohei Takano[2], Kristen Falcinelli[3], Jade Sauvé[3], and Katherine Smith[2]

[1]Department of Atmospheric and Oceanic Sciences and Institute of Arctic and Alpine Research, University of Colorado, Boulder, Colorado, USA
[2]Fluid Dynamics and Solid Mechanics (T-3), Los Alamos National Laboratory, Los Alamos, NM, USA
[3]School of Oceanography, University of Washington, Seattle, WA, USA

Correspondence: Cara Nissen (cara.nissen@colorado.edu)

Abstract. Since their advent over two decades ago, autonomous Argo floats have revolutionized the field of oceanography, and more recently, the addition of biogeochemical and biological sensors to these floats has greatly improved our understanding of carbon, nutrient, and oxygen cycling in the ocean. While Argo floats offer unprecedented horizontal, vertical, and temporal coverage of the global ocean, uncertainties remain about whether Argo sampling frequency and density capture the true spatio-temporal variability of physical, biogeochemical, and biological properties. As the true distributions of, e.g., temperature or oxygen are unknown, these uncertainties remain difficult to address with Argo floats alone. Numerical models with synthetic observing systems offer one potential avenue to address these uncertainties. Here, we implement synthetic biogeochemical Argo floats into the Energy Exascale Earth System Model version 2 (E3SMv2). Since the synthetic floats sample the model fields at model run time, the end-user defines the sampling protocol ahead of any model simulation, including the number and distribution of synthetic floats to be deployed, their sampling frequency, and the prognostic or diagnostic model fields to be sampled. Using a six-year proof-of-concept simulation, we illustrate the utility of the synthetic floats in different case studies. In particular, we quantify the impact of i) sampling density on the float-derived detection of deep-ocean change in temperature or oxygen and on float-derived estimates of phytoplankton phenology, ii) sampling frequency and sea-ice cover on float trajectory lengths and hence float-derived estimates of current velocities, and iii) short-term variability in ecosystem stressors on estimates of seasonal variability.

1 Introduction

Autonomous observing systems, such as profiling floats, sample the global ocean at much higher spatial and temporal resolution than traditional ship-based hydrography (e.g., Talley et al., 2016; Wong et al., 2020; Johnson et al., 2022). The >2 million profiles collected by autonomous Argo floats as part of "Core Argo" over the past two and a half decades have revolutionized our understanding of the spatio-temporal variability of physical ocean properties in the top 2000 m of the water column (e.g., Jayne et al., 2017; Johnson et al., 2022). Complementing the successful implementation of a global physical float array, a "Global Ocean Biogeochemistry (GO-BGC) Array" will be implemented over the upcoming years and decades as part of

"BGC Argo" (Roemmich et al., 2019; Claustre et al., 2020; Matsumoto et al., 2022; Schofield et al., 2022). In recent years, the deployment of >250 floats carrying biogeochemical and biological sensors within the "Southern Ocean Carbon and Climate

25 Observations and Modeling" project (SOCCOM, 2023; Sarmiento et al., 2023) have already permitted the first quantification of spatio-temporal variability in carbon, nutrients and oxygen in the Southern Ocean (e.g., Gray et al., 2018; Arteaga et al., 2020; Su et al., 2021; Prend et al., 2022a, b), promising similar scientific advances on the global scale. Similarly, the successful deployment of regional pilot "Deep Argo" arrays, with floats sampling down to 6000 m, demonstrates the possibility to expand this technology to the sampling of the whole water column (Blunden and Arndt, 2017). While the advent of Argo floats

30 represents a major step forward in the global ocean observing system with many potential uses, important uncertainties remain due to, e.g., the specific spatio-temporal coverage of the resulting dataset and uncertainties in determining float positions.

Most Argo floats profile the upper ocean approximately every 10 days, drifting with ocean circulation in between profiles at a parking depth of ~1000 dbar and being localized via the Global Positioning System (GPS) when transmitting data upon surfacing (Fig. 1 and Jayne et al., 2017). As a float's position is not known while drifting at the parking depth, its exact

35 trajectory in between profiles cannot be resolved, which directly affects float-based velocity estimates (Ollitrault and Rannou, 2013; Wang et al., 2022; Zilberman et al., 2023). Similarly, since the presence of sea ice prevents a float from surfacing to avoid sensor damage and total float loss (Klatt et al., 2007), currently floats can not be routinely localized while under sea ice, thus limiting the utility of Argo floats in high-latitude polar regions (e.g. Chamberlain et al., 2018).

Data from the Core Argo and BGC Argo floats may still not fully characterize the spatio-temporal variability in upper ocean

40 biogeochemical properties with high certainty despite the unprecedented spatio-temporal coverage of the global ocean. As an example, the projected float density in BGC Argo (1000 floats; Roemmich et al., 2019) corresponds to approximately one float per 6° x 6° grid cell. This target density is much larger than typical spatial scales of e.g., variability in phytoplankton chlorophyll (<100 km; see e.g. McKee et al., 2022) and pCO_2 (<1 km; see e.g. Eveleth et al., 2017), complicating the extrapolation from individual float-based observations to larger spatial scales. The 10-day sampling frequency is also longer than the turnover

45 time of phytoplankton biomass of 2-6 days (Behrenfeld and Falkowski, 1997). Similarly, biogeochemical properties, such as nutrients and carbonate chemistry are known to exhibit variability on time scales shorter than 10 days and sub-kilometer spatial scales (Kawai and Wada, 2007; Gruber et al., 2021; Nicholson et al., 2022). Altogether, this complicates the float-based detection of any trends in phytoplankton dynamics and carbon cycling.

Until the Deep Argo float program is fully operational, the majority of the ocean volume will remain under-sampled (Jayne

50 et al., 2017; Roemmich et al., 2019). As a result, the observation-based detection of changes in deep-ocean heat and oxygen content and in deep-ocean ventilation are still mostly based on scarce hydrographic observations (Johnson et al., 2015; Talley et al., 2016; Roemmich et al., 2019), which introduces large uncertainties when extrapolating to the global scale. All Deep Argo floats will be equipped with sensors to measure temperature and salinity, but the fraction of floats onto which an oxygen sensor will be mounted is less clear (King et al., 2021). While available funding will ultimately limit both the number of floats

55 to be deployed and the number of floats including an oxygen sensor, dedicated studies assessing the impact of float density on our ability to track large-scale changes in both deep-ocean temperature and oxygen are lacking.

Overall, in the absence of knowledge on the true distribution of physical, biogeochemical, and biological properties, uncertainties stemming from sampling frequency and density as well as imprecise localization are difficult to address with float-based observations alone, as it remains unknown how representative a given float profile is for a wider area or a longer time scale.

60 Model simulations are one approach to address these uncertainties. In particular, numerical models with synthetic observing systems can provide a known truth for the global distribution of any physical, biogeochemical or biological tracer, so that such models can be used as a ideal testbed to address uncertainties in sampling network design. In the past, this approach has been used to assess variability in oceanic heat content (Johnson et al., 2015; Allison et al., 2019; Garry et al., 2019; Gasparin et al., 2020), salinity distributions (Gasparin et al., 2020), the global oceanic carbon sink (Gloege et al., 2021; Hauck et al.,

65 2023), and chlorophyll concentrations (Ford, 2021), demonstrating the wide range of possible scientific applications. Most published studies extracted the synthetic observations offline, i.e., from time-averaged model fields after the model integration (e.g. Gasparin et al., 2020; Gloege et al., 2021). This approach is storage-intensive, as model fields need to be stored at high temporal frequency (often daily) because real-world observations always represent snapshots of ocean properties rather than time-averages, leading to higher uncertainties if lower-frequency (e.g., monthly) model fields are used to extract the synthetic

70 observations. An alternative approach, which minimizes this uncertainty, is to extract synthetic observations online during the model run time, but only few models have such capabilities to date (see e.g. Brady et al., 2021; Clow et al., 2023).

Here, we present the new synthetic biogeochemical float capabilities of the Energy Exascale Earth System Model version 2 (E3SMv2). To more closely resemble real-world Argo floats, the synthetic floats sample the model fields online during model run time. The number and distribution of the synthetic floats, the sampling frequency, and the sampled variables are defined by

75 the end-user before the start of the model experiment. After describing the implementation of synthetic floats into E3SMv2 in more detail in the following section, we will present its utility for physical, biogeochemical, and biological research questions with several case studies.

2 Methods

2.1 Model description: E3SMv2 with synthetic biogeochemical floats

80 We implement synthetic biogeochemical floats into the ocean component of the E3SMv2 version 2 (E3SMv2; Golaz et al., 2022). The physical component of E3SMv2 consists of the E3SMv2 Atmosphere Model version 2 (EAMv2), the E3SMv2 Land Model version 2 (ELMv2), the Model for Prediction Across Scales Ocean (MPAS-O), and the Model for Prediction Across Scales Sea Ice (MPAS-Seaice; Golaz et al., 2019; Petersen et al., 2019; Turner et al., 2022; Golaz et al., 2022). Both MPAS-O and MPAS-Seaice are run on unstructured multi-resolution model grids, allowing for enhanced model resolution in selected

85 regions (Ringler et al., 2013). Ocean biogeochemistry is described by the Marine Biogeochemistry Library (MARBL; Long et al., 2021), which is based on the Biogeochemical Elemental Cycling module (BEC; Moore et al., 2004, 2013). MARBL describes the biogeochemical cycling of carbon, nitrogen, silicon, phosphorus, iron, and oxygen and allows flexible lower trophic level ecosystem configuration (Long et al., 2021). Lastly, sea-ice biogeochemistry is represented by four dissolved
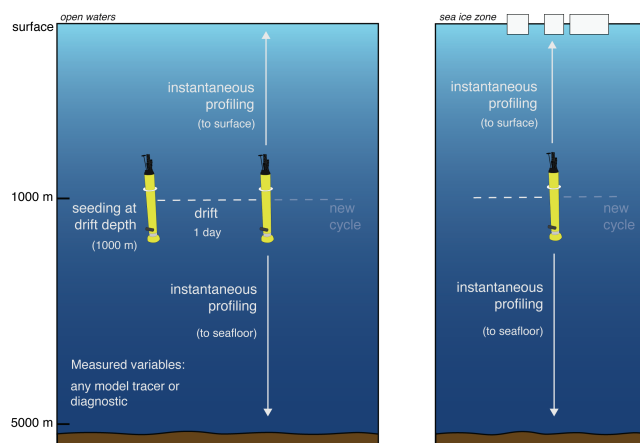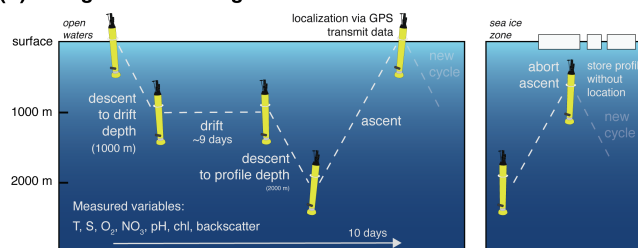
Figure 1. Sketch of a synthetic vs. a real-world biogeochemical float. (a) A synthetic biogeochemical float in E3SMv2 with a 1-day sampling cycle, i.e., drifting at a parking depth of 1000 m for a day and then instantaneously sampling any model tracer or diagnostic throughout the water column both in open waters and under sea ice cover. (b) A biogeochemical Argo float with a 10-day sampling cycle, i.e., drifting at a parking depth of 1000 m for about 9 days, then descending to 2000 m and sampling temperature, salinity, oxygen, nitrate, chlorophyll, and particulate backscatter in the upper 2000 m while ascending. This float relies on localization via GPS, so that its exact position is unknown while under ice. Panel (b) is adapted from Jayne et al. (2017).

inorganic nutrients (silicate, nitrate, ammonium, and iron), dissolved organic nitrogen, and three phytoplankton groups, i.e.,

90   diatoms, small flagellates, and *Phaeocystis* sp. (MPAS-Seaice zbgc; Jeffery et al., 2020).

The implementation of synthetic floats in E3SMv2 builds on the Lagrangian, in Situ, Global, High-Performance Particle Tracking (LIGHT) online particle tracking module developed for MPAS-O (Wolfram et al., 2015; Brady et al., 2021). For our simulations, LIGHT particles are seeded at a depth of 1000 m and advected laterally with ocean circulation using a second order Runge-Kutta scheme; unlike previous studies (e.g., Brady et al., 2021), particles are not permitted to move vertically.

95   During the simulation, the virtual "floats" instantaneously sample the whole water column at their current location and with a prescribed frequency, e.g., daily or 10-day. The synthetic floats are thus not subject to lateral current displacement during ascent or descent (cf. synthetic and Argo floats in Fig. 1). Further, in contrast to Argo floats, whose position can only be registered
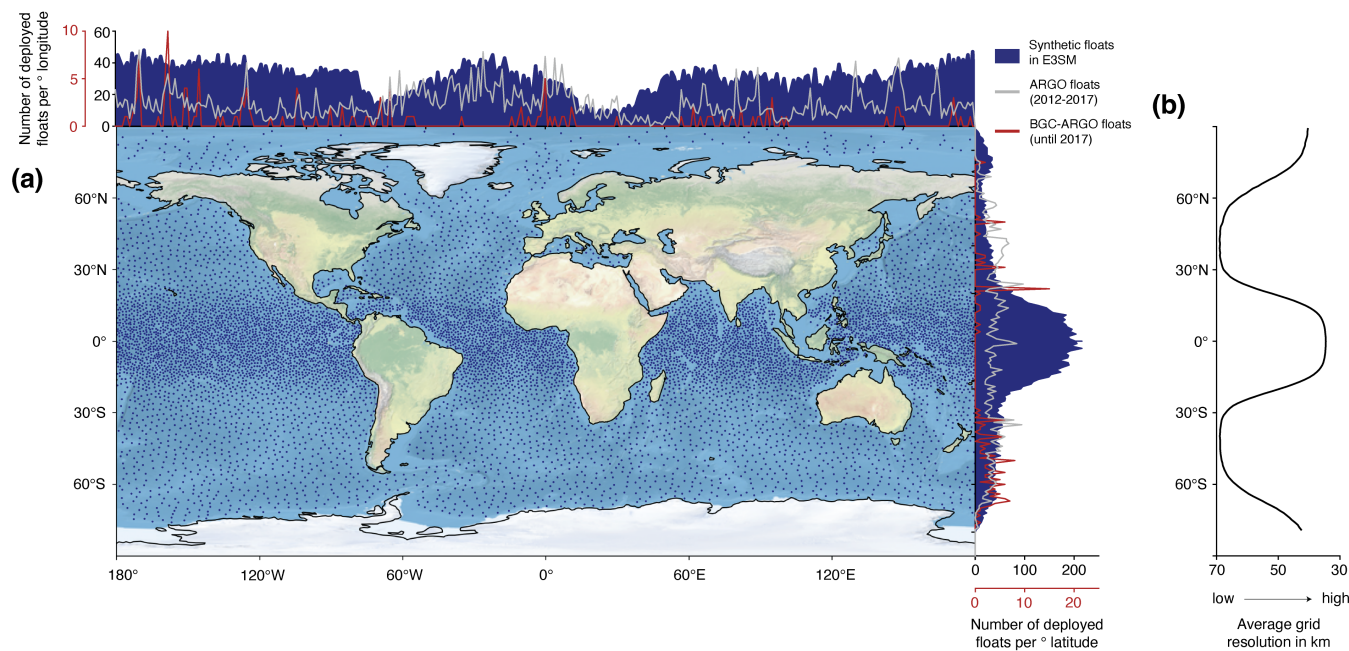
Figure 2. (a) Distribution of deployed synthetic floats in E3SMv2 and core and biogeochemical Argo floats between 2012-2017. Blue dots on the map indicate the initial positions of synthetic floats seeded in the deep ocean (>2000 m). On the sides of the map, the number of deployed synthetic floats per degree longitude (top) and latitude (right) is shown in blue, with the corresponding numbers for Core Argo and biogeochemical Argo floats denoted as the gray and red lines, respectively. Note the different axis scale for the biogeochemical Argo floats (red scale). (b) Zonal average model resolution in km in the E3SMv2 test simulation with synthetic floats analyzed in this paper.

upon surfacing in ice-free waters, the position of synthetic floats in E3SMv2 is known at all times (Fig. 1). In general, any prognostic or diagnostic physical or biogeochemical model variable can be recorded by the synthetic floats, and the sampled

100 variables are bilinearly interpolated to the float's current location. The distribution of virtual particles to be seeded (or floats to be "deployed"), the sampling frequency, and the variables to be recorded are defined by the user prior to the model simulation to best align with a given application or research question.

2.2 Model setup and simulation

For this study, we use the ocean-ice version of E3SM, i.e., MPAS-O and MPAS-Seaice, each with their corresponding biogeo-

105 chemical module (MARBL and MPAS-Seaice zbgc, see section 2.1). In this version, the ecosystem within MARBL consists of three phytoplankton functional types (diatoms, diazotrophs, and a mixed small phytoplankton group) and a single zooplankton functional type. To demonstrate the functionality of the new synthetic float tool, we conduct a six-year model simulation from 2012 to 2017, i.e., overlapping with the SOCCOM period starting in 2014 (Riser et al., 2018; Sarmiento et al., 2023). In our simulation, the model grid has a horizontal resolution that ranges from ~30 km in the tropics and the high latitudes to ~60 km

110 in the subtropics (Fig. 2), and includes 60 z-star levels in the vertical (EC30to60E2r2 mesh; Petersen et al., 2019). The simulation is forced with 3-hourly atmospheric data from the Japanese atmospheric reanalysis version 1.4 (JRA; Tsujino et al., 2018). All model fields are initialized from an existing, unpublished simulation using the same model grid and atmospheric forcing data as the six-year simulation analyzed in this paper (Takano et al., 2023). Model tracers in this existing simulation were initialized in the same manner as in Burrows et al. (2020). The simulation was spun-up from 1750 to 1957 using repeat-year

115 atmospheric forcing derived from the period July 1984 to June 1985, with atmospheric CO_2 concentrations held constant at 284 ppm between 1750 and 1850 and increasing according to historical records thereafter (Meinshausen et al., 2017). Eulerian model output is stored at monthly frequency for all variables; daily output is stored for phytoplankton and zooplankton biomass.

Synthetic biogeochemical floats are "deployed" on January 1, 2012 at every third vertex of the model grid, resulting in a

120 total of 10560 seeded floats. Only floats seeded in the open-ocean away from continental shelves and slopes are retained for the analysis in this study (8739 floats, see Fig. 2a). The resulting density of synthetic floats is up to four times higher than the density of Core Argo floats deployed for the same period in the (sub)tropics (especially in the Pacific and Indian sectors) and comparable to it elsewhere (Fig. 2a). For all latitudes and longitudes, the synthetic float density exceeds that of Deep Argo (not shown) and BGC Argo (red lines; note the different scale; Fig. 2a). The synthetic floats instantaneously sample

125 the whole water column every day at midnight Greenwich Mean Time. In addition to every float's position at sampling, the following model variables are recorded: temperature, salinity, dissolved inorganic carbon (DIC), alkalinity, nitrate, silicic acid, phosphate, oxygen, total phytoplankton carbon biomass, total phytoplankton chlorophyll, zooplankton carbon biomass, O_2 consumption, O_2 production, the zonal and meridional resolved velocity components, sea-ice fraction, mixed-layer depth (as determined with the 0.03 kg m^{-3} density criterion), sea-level pressure, surface partial pressure of CO_2 (pCO_2) in seawater,

130 difference in pCO_2 between the ocean and the atmosphere, atmospheric CO_2 concentration, air-sea CO_2 flux, and air-sea O_2 flux. We calculate pH on each float profile offline using the python routines to model the ocean carbonate system (mocsy v2.0; code was obtained from https://github.com/jamesorr/mocsy on July 19, 2023; Orr and Epitalon, 2015). We use DIC, alkalinity, potential temperature, salinity, silicic acid, phosphate, and sea-level pressure from each float profile as inputs for the mocsy functions.

3 Results & Discussion

3.1 Evaluation of synthetic floats in E3SMv2

We evaluate the synthetic floats in E3SMv2 in two ways: 1) by comparing the synthetic float data to the full Eulerian model output, and 2) by comparing the synthetic float data to Core Argo data (Argo, 2023). Adequate representation of currents at 1000 m, the float parking depth, is required for realistic sampling with the synthetic modeled floats (Fig. 3). In E3SM, current

140 velocities at 1000 m are highest in the Antarctic Circumpolar Current and in the subpolar North Atlantic off the southeast coast of Greenland (locally >8 cm s^{-1}), with velocities of less than 3 cm s^{-1} elsewhere (Fig. 3a). Fig. 3b shows a Lagrangian-based velocity estimate derived from the 10-day synthetic float positions. The spatial patterns and magnitudes in current velocity
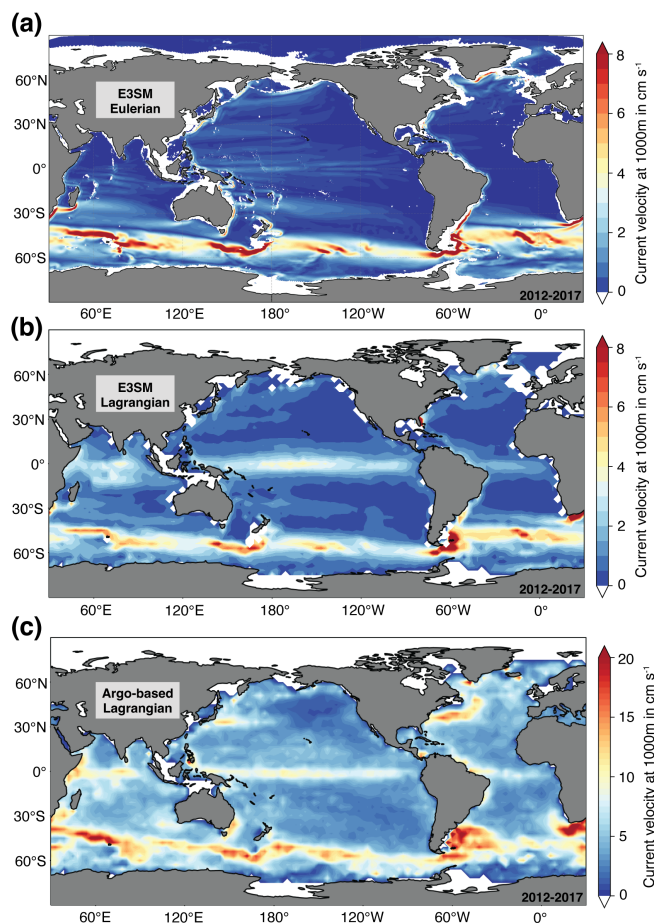
Figure 3. Horizontal current velocity at 1000 m in cm s^{-1} (a) in the full Eulerian E3SMv2 output from 2012 to 2017, (b) derived from 10-day position data of synthetic floats in E3SMv2 and averaged for regular 3° x 3° boxes and (c) derived from 10-day position data of Argo floats and averaged for regular 3° x 3° boxes (Zilberman et al., 2022, 2023). Note the different scales between the panels (a)-(b) and (c).

produced by the Eulerian model output are largely captured by the synthetic floats (cf. Fig. 3a & b). The equatorial regions are the only exception, for which the Lagrangian E3SMv2 estimate suggests higher velocities (up to 4 cm s^{-1}) than the Eulerian estimate (up to 2 cm s^{-1}). This implies substantial variability of current speeds at 1000 m in E3SMv2 at sub-monthly time scales that are not captured by the Eulerian time-averaged output. In comparison to Argo-derived current speeds (Zilberman et al., 2022, 2023), velocities in E3SMv2 at 1000 m are a factor 2-3 too low (compare panels a-c in Fig. 3). This bias in ocean current speeds is a common feature in non-eddying ocean circulation models and is possibly related to how high-frequency dynamical processes are parameterized (e.g. internal mixing or tides; Su et al., 2023), in addition to limitations related to grid resolution. In spite of this bias, most high-velocity features present in the Argo-derived data set are reproduced in E3SMv2 (Fig. 3b & c).
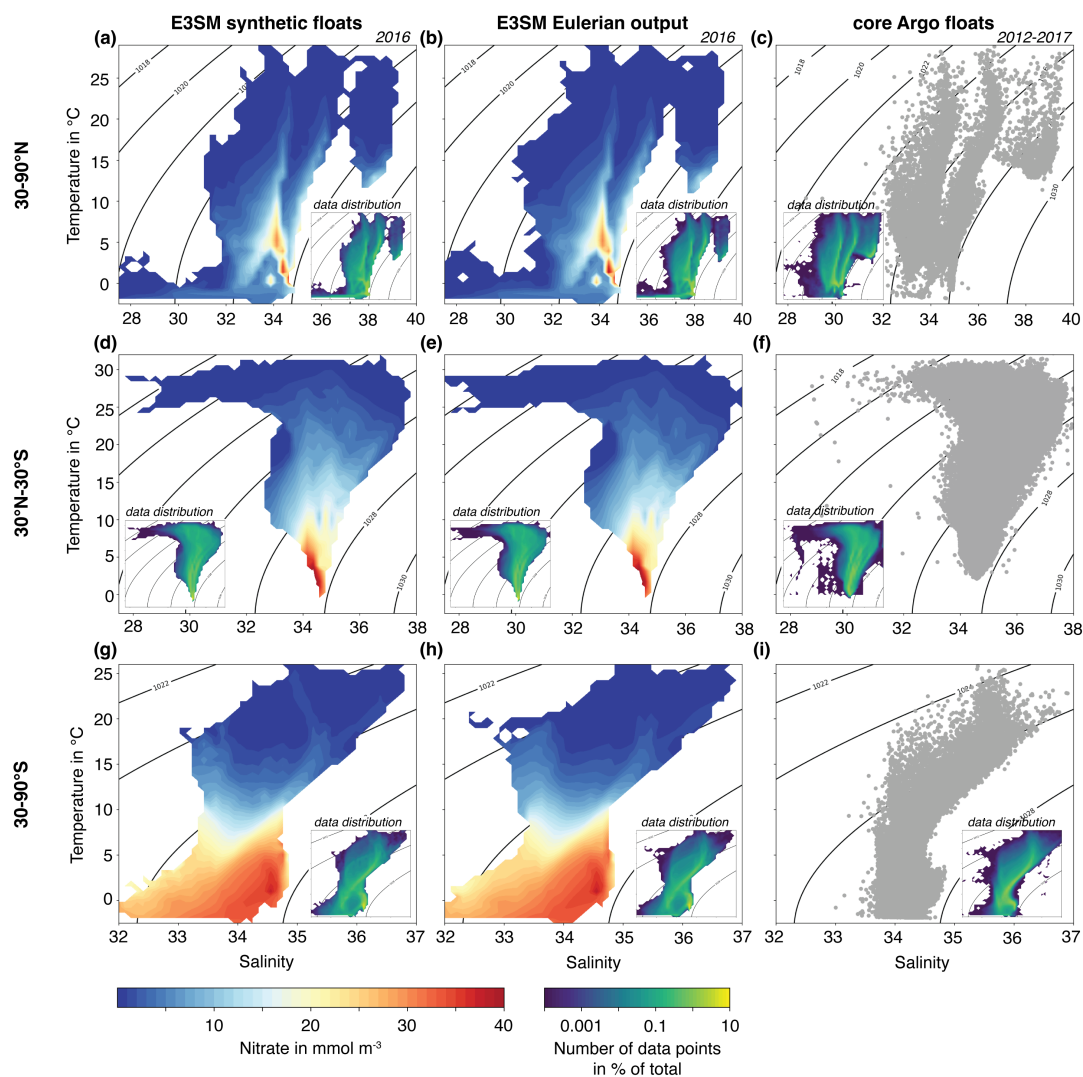
Figure 4. (a)-(c) Temperature-salinity-diagrams north of 30°N of (a) 10-day data from synthetic floats in E3SM, (b) monthly mean Eulerian E3SMv2 output, and (c) Core Argo float data (Argo, 2023). Data points in panels (a)-(b) are for the whole water column in 2016, and data points in panel (c) are a subset of all available points for the top 2000 m and 2012-2017. We note that subsurface interannual variability of temperature and salinity is negligible on the large spatial scale assessed here and that >80% of all E3SMv2 data are in the top 2000 m of the water column, implying comparability of the first two columns with the third. Data points in panels (a)-(b) are colored as a function of nitrate concentrations in mmol m^{-3} and shown as averages within 50 equally-sized bins in temperature and salinity space. The small inlets in panels (a)-(c) show the distribution of data. (d)-(i) Same as (a)-(c) but for (d)-(f) 30°N-30°S and (g)-(i) south of 30°S.

The only exception is the Gulf Stream, which is too shallow in E3SMv2 (not shown), resulting in much lower current speeds at 1000 m in E3SMv2 than in the Argo-based estimate in the northwest Atlantic (<2 cm s^{-1} compared to ~12 cm s^{-1}).
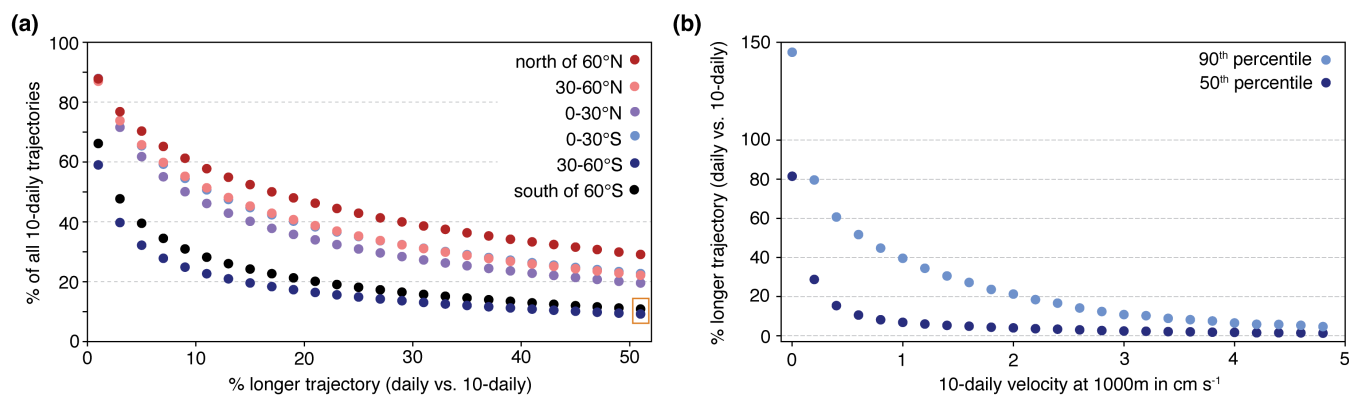
Figure 5. (a) Comparison of the 10-day trajectory length of the synthetic floats for daily and 10-day sampling. Plotted is the fraction of all 10-day trajectories (y axis) displaying a trajectory length at daily sampling that exceeds the length at 10-day sampling by a certain percentage (x axis). Data are grouped into 30° latitudinal bands (see colors). "% longer trajectory" as shown on the x axis is equivalent to "% lower velocity". The orange box refers to the example given in the manuscript text. (b) Difference in 10-day trajectory length of the synthetic floats for daily and 10-day sampling plotted as a function of the average 10-day velocity at 1000 m in cm s^{-1} as derived from the float positions. Shown are the 50$^{\text{th}}$ (dark blue) and 90$^{\text{th}}$ percentile (light blue).

Synthetic floats in our configuration of E3SMv2 are capable of sampling the wide range of global physical and biogeochemical properties that appear in the Eulerian-mean model output, and they sample across a wider range of global temperature and
155   salinity values than Core Argo floats. By comparing the data sets in temperature-salinity space, we evaluate the ability of the synthetic floats in E3SMv2 to correctly sample their environment (Fig. 4). For the three latitudinal bands 30-90°N, 30N-30°S, and 30-90°S, the sampled temperature-salinity-nitrate space is very similar for the 10-day whole-water-column synthetic float output and the monthly mean Eulerian output (compare the first two columns in Fig. 4; note that only data for the year 2016 is shown here). We attribute any differences to the differing temporal resolution of the data. In comparison to all core-Argo data
160   2012-2017 (third column in Fig. 4), the model output samples a larger temperature-salinity space (see e.g., cold, fresh waters for 30-90°N in panels a-c). This difference can likely mostly be attributed to differences in the float distribution (e.g., in contrast to in E3SM, there are very few floats in the Arctic within Argo, see Fig. 2) and to differences in the sampled water depth (see e.g., fewer data points in Core Argo than in E3SMv2 for the latitudinal band 30-90°S at temperatures between 0°C and 5°C and salinities between 34 and 35; these data points lie in the deep ocean >2000 m as indicated by nitrate concentrations
165   exceeding 35 mmol m^{-3}). In summary, the synthetic floats in E3SMv2 reproduce key patterns of variability both of the Eulerian model output and of the Core Argo floats, making these floats a valuable tool for the assessment of spatio-temporal variability in physical and biogeochemical properties from a Lagrangian perspective and for sampling network design.

Geoscientific
Model Development
Discussions

3.2 The impact of sampling frequency on float-derived velocities

Analysis of our modeled synthetic floats reveals that the true distance traveled by the floats over a 10-day period is longer than
170   indicated by their 10-day position differences. We use the synthetic E3SMv2 floats to compare the difference in the 10-day
trajectory length (equivalent to the 10-day averaged velocity) when a) knowing the respective float's position once per day
and b) only knowing the respective float's position on day 1 and day 10 (as for Core Argo floats) of each 10-day period.
Our analysis shows that the mismatch in the trajectory length between daily and 10-day float profiling frequencies can be
substantial (Fig. 5), with 10% of all 10-day trajectories in the Southern Ocean south of 30°S being at least 50% longer when
175   the float position is known every day as opposed to only at the start and end of the trajectory (orange box in Fig. 5a). In other
words, for 10% of all Southern Ocean synthetic float trajectories, velocities derived from the float positions are more than
50% too high if the float positions are only known at the start and end of any 10-day period. In general, the smaller the error
in trajectory length/velocity (x axis in Fig. 5a), the more trajectories are affected. In addition, trajectories in more northerly
latitudinal bands are more severely affected by this error than those in the Southern Ocean. For example, while 60% of all
180   trajectories north of 60°N (25% in the Southern Ocean south of 60°S) display a difference of at least 10% in 10-day trajectory
length/velocity for daily vs. 10-day float profiling frequencies, the number of affected trajectories amounts to 30% (10% in the
Southern Ocean) for a 50% mismatch. These regional differences reflect the general gradient from high velocities at 1000 m in
the Southern Ocean (Antarctic Circumpolar Current) to lower velocities in the northern hemisphere (see Fig. 5b and Fig. 3a).
Since the 10-day trajectory length forms the basis for float-derived velocity estimates (Fig. 3; Ollitrault and Rannou, 2013;
185   Zilberman et al., 2023), our analysis illustrates the bias introduced by the absence of more frequent knowledge on every floats'
position, and demonstrates that this bias can be quite substantial in certain instances.

3.3 Case studies

Here, we present four example mini-studies using output from the synthetic floats in E3SMv2 to illustrate some of the capa-
bilities of this modeling tool. The analysis in each of the following subsections is not meant to be comprehensive, and many
190   more applications are imaginable, some of which will be outlined further.

3.3.1 Case study I: Quantifying physical-biogeochemical variability in Lagrangian and Eulerian analysis frameworks

Our first case study quantifies the seasonal amplitude of physical and biogeochemical properties using monthly-mean Eulerian
model output and daily synthetic floats. This case study is motivated by the need to ascertain knowledge of the present-
day exposure of marine organisms to certain environmental conditions, as this can inform their potential for acclimation and
195   adaptation to future environmental change (Kapsenberg and Cyronak, 2019). While the occurrence of extreme environmental
conditions in e.g., temperature, oxygen, or carbonate chemistry have been quantified in studies using daily-mean Eulerian
model output (Gruber et al., 2021), a similar observation-based assessment is often complicated by scarce spatial and temporal
data availability (see discussion in Gruber et al., 2021). High temporal resolution data can be provided by satellites for the
global surface ocean (although temporal composites are often necessary due to data gaps at a daily scale; Prend et al., 2022c;
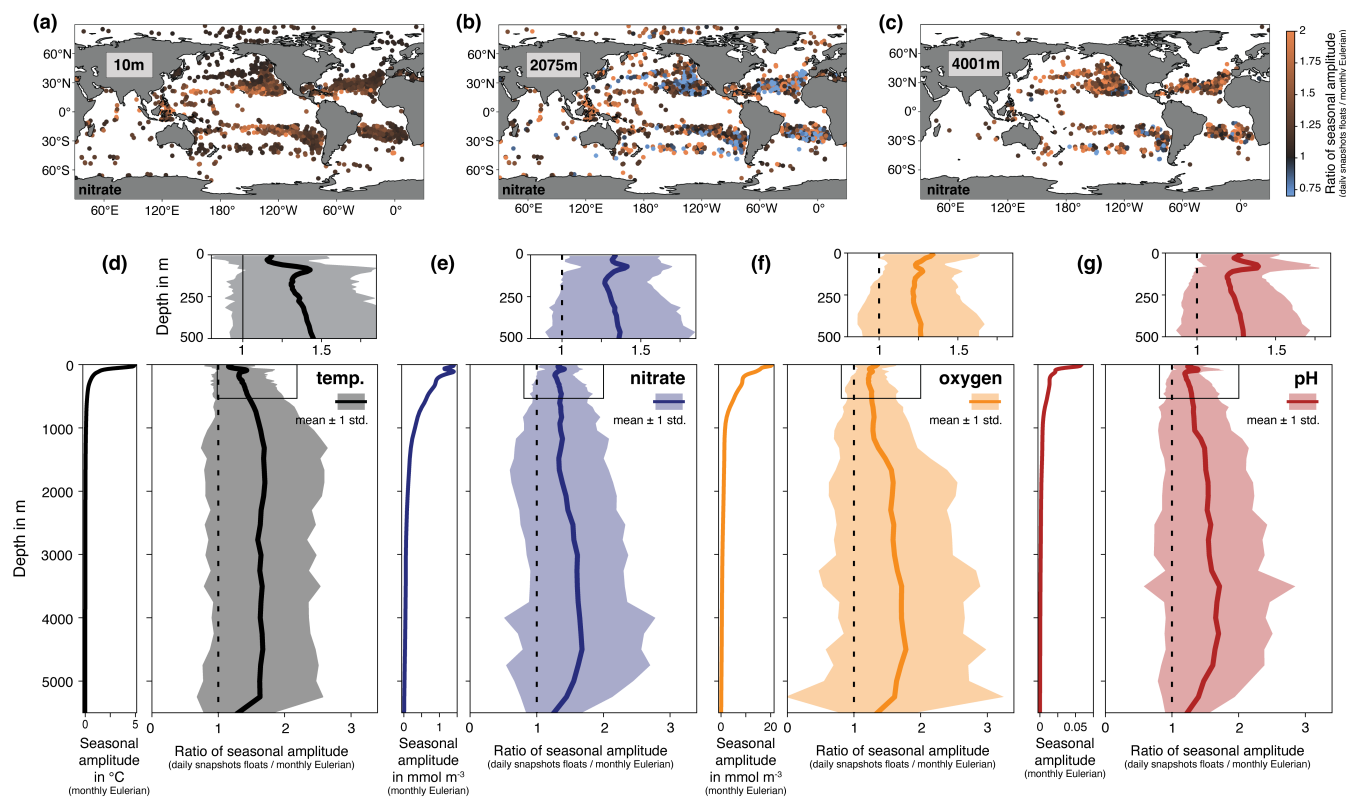
Figure 6. Case study "Quantifying physical-biogeochemical variability": (a)-(c) Maps showing the ratio of the seasonal amplitude of nitrate between the daily snapshots from the synthetic floats and the monthly mean Eulerian output at those model grid cells which are sampled by a float each day over any full calendar year between 2012 and 2017. The seasonal amplitude is defined as the annual maximum minus the annual minimum of the daily snapshots from the synthetic floats and the monthly mean Eulerian output, respectively. Maps are shown at (a) 10 m, (b) 2075 m and (c) 4001 m. (d) Vertical profiles of (left) the seasonal amplitude of temperature in the monthly mean Eulerian output and (right) the ratio of the seasonal amplitude between the daily snapshots from the synthetic floats and the monthly mean Eulerian output. The solid line represents the mean over all available grid cells, and the shading denotes one standard deviation. The top panel zooms into the top 500 m. (e)-(g) Same as (d) but for (e) nitrate concentrations, (f) oxygen concentrations and (g) pH.

Clow et al., 2023), by gliders for the upper ocean in small regions (Chai et al., 2020) and by profiling floats, which offer advances over these aforementioned technologies in terms of the horizontal, vertical and temporal coverage of the global ocean. Nonetheless, the contribution of sub-monthly variability to estimates of the seasonal cycle amplitude in key ecosystem drivers remains unquantified. Here, we quantify the seasonal amplitude (maximum minus minimum over a calendar year) of temperature, nitrate, oxygen, and pH obtained from monthly-mean Eulerian E3SMv2 output and from daily profiling synthetic floats and illustrate their ratios in Fig. 6.

On average, the daily Lagrangian float snapshots capture a higher seasonal amplitude than the monthly mean Eulerian output for all four ecosystem drivers (ratio>1 in Fig. 6) when considering only those model grid cells that are sampled by a synthetic

float each day of the calendar year, i.e., corresponding mostly to areas where velocities at 1000 m are low (compare Fig. 6a-c
to Fig. 3). In the upper 500 m of the water column, where the seasonal amplitude is largest, the Lagrangian-based seasonal

210 amplitude exceeds the Eulerian-based amplitude by >25% for all variables. Temperature, nitrate, and pH have their maximum
discrepancy in the seasonal amplitude near the base of the mixed layer at ∼100 m, indicating that short-term variability in
mixed-layer depth can affect the interpretation of float-based records. However, oxygen displays a maximum discrepancy in
amplitude at the surface (inlets in Fig. 6d-g), likely associated with the time scale for air-sea exchange of oxygen (a few weeks;
Sarmiento and Gruber, 2006). Our findings imply that atmosphere-ocean oxygen disequilibrium manifests more strongly in the

215 daily snapshots. For all four variables, the mean difference in the Lagrangian-based and Eulerian-based seasonal amplitude is
largest below 1000 m (Fig. 6d-g). However, the variability around the mean is large (typically ranging from just below 1 to at
least 2.5, see shading in Fig. 6d-g), and the seasonal amplitude is small at these depths (see left panels in Fig. 6d-g).

In the absence of daily averaged Eulerian model output for temperature, nitrate, oxygen, and pH, our reported discrepancy
in the seasonal amplitude represents an upper limit. Much of the higher-frequency variability is averaged out when computing

220 monthly averages, suggesting that differences in seasonal amplitude would be smaller if float-based snapshots were compared
to daily-averaged instead of monthly-averaged Eulerian model output. While daily-averaged model output is typically used
in studies assessing changes in extreme environmental conditions, our understanding of variability on daily and sub-daily
time scales and its impacts on global ecosystems remains incomplete (Gruber et al., 2021). Based on regional hydrographic,
glider, and mooring observations, we know that substantial short-term variability in ecosystem stressors is caused by, e.g.,

225 the diurnal cycle (Torres et al., 2021), tides (Droste et al., 2022), or ocean weather (Nicholson et al., 2022). As a result, such
variability is possibly a key factor in determining an organism's resilience to environmental change (Helmuth et al., 2014;
Kapsenberg and Cyronak, 2019). Biogeochemical Argo floats will facilitate the characterization of spatio-temporal variability
in the global distribution of ecosystem stressors at unprecedented scales, but it remains unclear to what extent short-term
variability interferes with their 10-daily sampling protocol in the detection of larger-scale patterns or longer-term trends. To

230 close this knowledge gap, future work with synthetic floats in E3SMv2 should include a more detailed assessment of the role
of temporal averaging in determining spatio-temporal variability in ecosystem stressors.

3.3.2 Case study II: Float-based detection of changes in deep-ocean water-mass properties

To obtain basin-scale estimates of variability and trends in the properties of climatically important water masses such as
Antarctic Bottom Water or North Atlantic Deep Water, a denser observing network in both space and time is required than

235 hydrographic observations can provide (Talley et al., 2016; Jayne et al., 2017; Roemmich et al., 2019). To that end, a key goal
of Deep Argo is the detection of changes in deep-ocean heat content, which will facilitate the tracking of the deep ocean's
contribution to steric sea-level rise (Johnson et al., 2015; Roemmich et al., 2019). Further, the addition of oxygen sensors on
Deep Argo will enable the detection of changes in deep-ocean oxygenation, particularly in regions downstream of deep and
bottom water formation in the North Atlantic and Southern Ocean (Hoppema, 2004; Rhein et al., 2017). While any difference

240 in the spatio-temporal variability of deep-ocean temperature and oxygen will directly impact the number of floats required to
capture large-scale changes over time, these differences remain unquantified to date. In this case study, we use different float
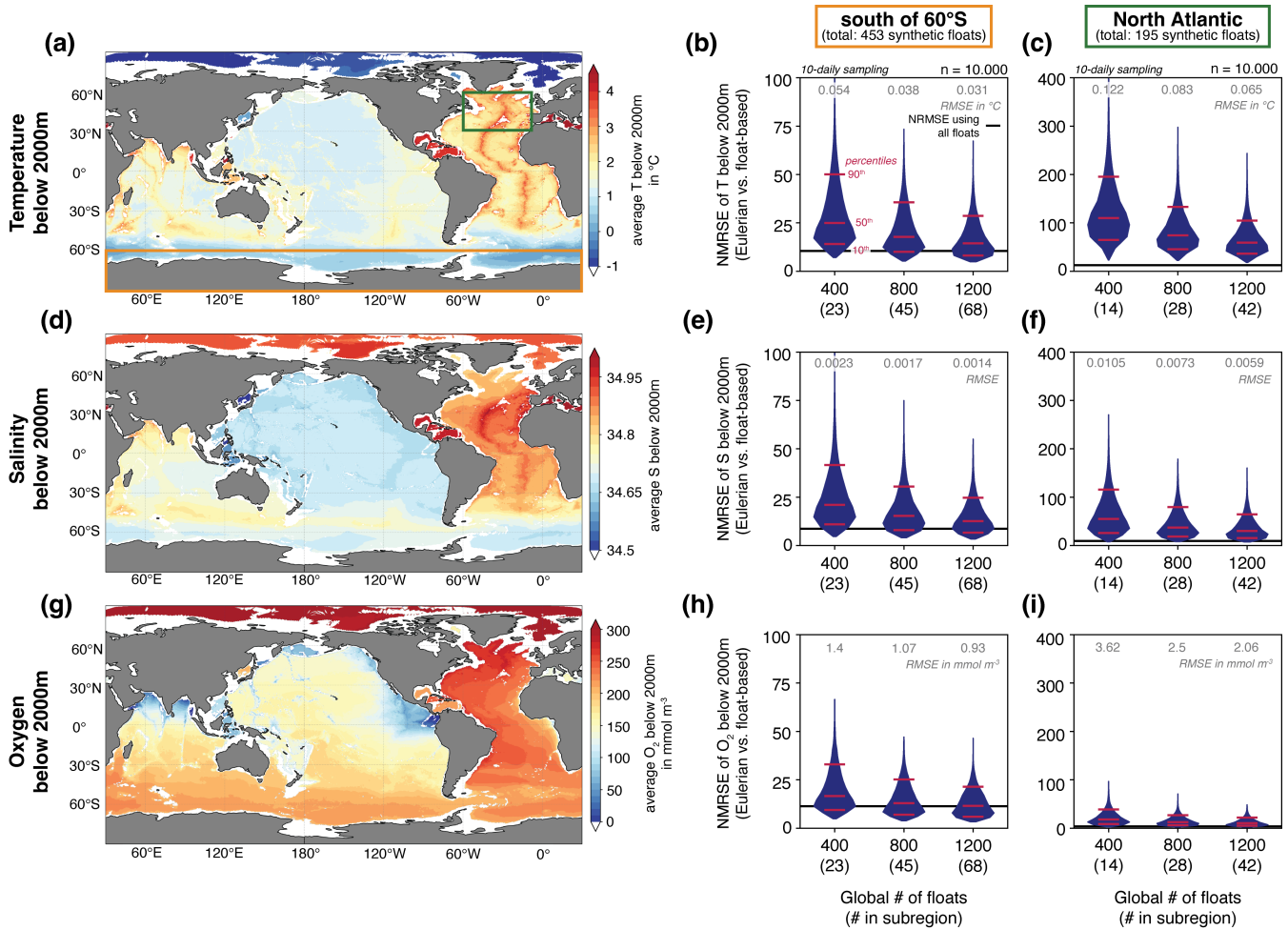
Figure 7. Case study "Deep ocean": (a) Temperature in °C below 2000 m and averaged over 2012-2017. (b) Violin plots of the normalized root mean square error (NRMSE) of the monthly temperature time series for the area south of 60°S (orange box in panel a) between the full Eulerian output and a float-based estimate using 10-day sampling and subsets of all synthetic floats corresponding to different global target float densities (x axis). The subsampling is repeated 10,000 times. The red horizontal lines denote the 90th, 50th, and 10th percentile (from top to bottom). The black horizontal line corresponds to the NRMSE using all synthetic floats in the area. The RMSE is normalized by the area-weighted standard deviation of monthly mean temperature values of all grid cells in the area. The unnormalized RMSE is denoted in grey. (c) Same as (b) but for the North Atlantic between 30-60°N and 10-60°W (green box in panel a). (d)-(f) and (g)-(i) Same as (a)-(c) but for (d)-(f) salinity and (g)-(i) oxygen in mmol m^{-3}, respectively.

densities to quantify the error associated with capturing larger-scale temporal variability in deep-ocean temperature, salinity, and oxygen in the North Atlantic and Southern Ocean.

To facilitate the comparison of errors across variables, we calculate the normalized root mean square error (NRMSE; normalized by one standard deviation of all monthly Eulerian values averaged over the respective subarea) between Eulerian and

245

13

synthetic float model output. In particular, we calculate the NRMSE between the 6-year long monthly time series of the Eulerian model output and the float-derived monthly time series constructed from 10-day sampling for a given float density (see violin plots in Fig. 7 and maps in Fig. 7 for spatial distribution of variables).

Our analysis reveals that (1) the NRMSE decreases when more floats are used in the calculation of the error across all variables and in both regions, (2) the NRMSE is overall higher in the North Atlantic than in the Southern Ocean, and (3) for both regions, the NRMSE is highest for temperature and lowest for oxygen. The first point aligns with that of other studies (e.g., Youngs et al., 2023). The second point can possibly be attributed to more variability in bottom topography in the North Atlantic (note the presence of the mid-Atlantic ridge in Fig. 7a), making a float-based estimate of any water-mass property in this region more dependent on the float distribution than in the Southern Ocean, where the average spatial variability is much smaller (Fig. 7a, d, g). This finding is unaffected by the normalization of the error metric; the root mean square error (RMSE) is approximately two (temperature), four (salinity) and two (oxygen) times larger in the North Atlantic than in the Southern Ocean for all float densities over the 6-year time series (see printed gray numbers in violin plots in Fig. 7).

That the NRMSE is highest for temperature and lowest for oxygen in both regions has implications for the design of a deep-ocean float array. A generally lower NRMSE for biogeochemical (oxygen) than for physical (temperature and salinity) water-mass properties implies that, to capture property changes at comparable accuracy (at least in terms of normalized error metrics), it would be sufficient to equip a subset of all floats with oxygen sensors, thereby reducing the overall cost of a deep-ocean float array. However, we note that based on our analysis, the RMSE for a deep-ocean array consisting of 400 floats, i.e., one third of the global target density (Johnson et al., 2015; Jayne et al., 2017), amounts to 1.4 mmol m^{-3} (Southern Ocean; printed gray numbers in Fig. 7h) and 3.62 mmol m^{-3} (North Atlantic; Fig. 7i), which is approximately six times larger than the range of regionally and monthly averaged Eulerian model output over the 6-year time series (0.24 mmol m^{-3} and 0.60 mmol m^{-3}; not shown). Our analysis thus suggests that such a reduced deep-ocean oxygen array would complicate the float-based detection of trends in deep-ocean ventilation on large spatial scales. Importantly, the RMSE is reduced by 33% and 43% in the Southern Ocean and the North Atlantic, respectively, when equipping a full global network of 1200 floats with oxygen sensors instead of only 400 floats (Fig. 7h-i), enhancing our ability to capture variability in oxygen concentrations over large spatial scales with floats.

3.3.3 Case study III: Float trajectories under Southern Ocean sea-ice cover

Argo floats rely on localization via GPS upon surfacing in ice-free waters (Fig. 1b). Since the risk of damaging instrumentation is high upon contact with sea ice, conventional floats follow an ice-avoidance protocol, which makes them abort their ascent when subsurface temperatures indicate high likelihood of sea ice present at the surface (-1.79°C between 20-50 m; Klatt et al., 2007). In the absence of position data, the trajectory of such a float is then linearly interpolated between the last position before and the first position after the under-ice period. Depending on how long a float cannot be localized, this procedure potentially causes large uncertainties in the estimated trajectory (Chamberlain et al., 2018; Nguyen et al., 2020). In this case study, we use our ability to localize synthetic floats at all times, including under sea-ice cover (Fig. 1a), to quantify the impact of sea-ice presence on float trajectories for different sectors of the Southern Ocean.
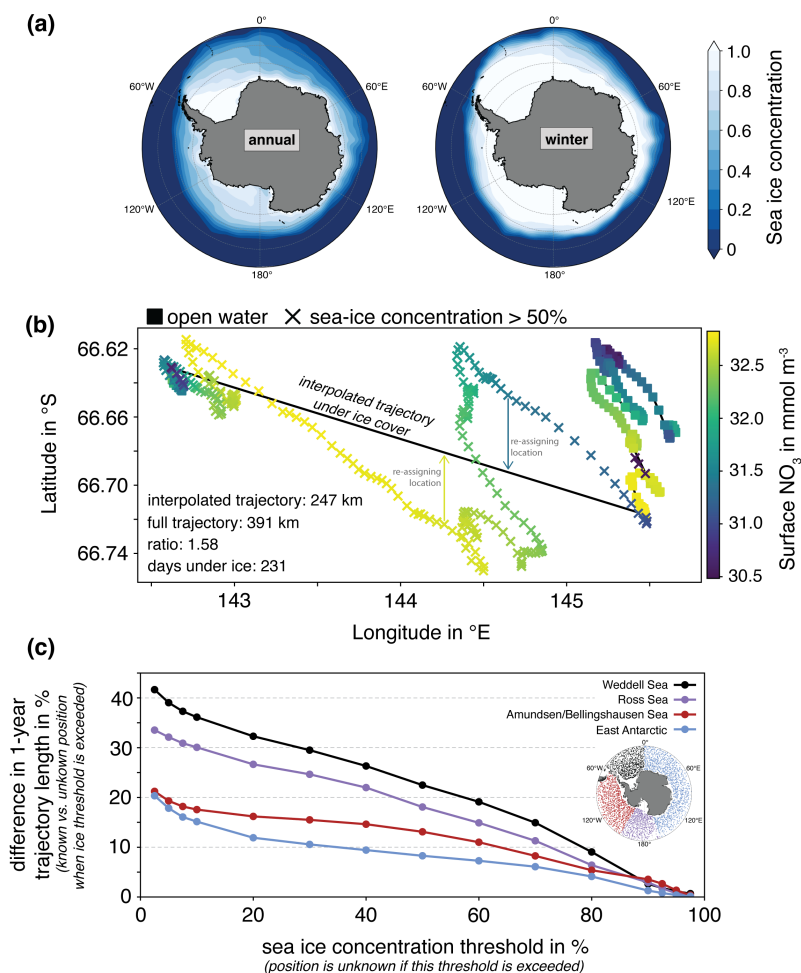
Figure 8. Case study "Sea-ice cover": (a) Annual mean (left) and winter mean (June, July, August; right) sea-ice concentration in E3SMv2 averaged over 2012-2017. (b) Example illustrating the effect of unknown float positions under sea ice on the trajectory length. The daily location of the float is plotted in latitude-longitude space and colored as a function of surface nitrate concentrations (in mmol m^{-3}). For this example, the location is assumed unknown whenever the sea-ice concentration exceeds 50% (crosses) and known elsewhere (squares). When the float location is unknown, the trajectory is obtained by linearly interpolating between the last and first known position (black line). (c) Difference in trajectory length in % for all 1-year trajectories between 2012-2017 with known or unknown daily positions as a function of the sea-ice concentration threshold in %. The position of a float is considered unknown whenever a given sea-ice threshold is exceeded, and data gaps are linearly interpolated. Results are shown separately for the Weddell Sea (black), Ross Sea (purple), Amundsen and Bellingshausen Sea (red), and the east Antarctic (blue). See map for the initial float positions of all considered 1-year trajectories.

280   Linear interpolation of float position data in E3SMv2 under-ice floats introduces biases in under-ice trajectory length and estimated under-ice statistical tracer properties (Fig. 8). In agreement with observations (Eayrs et al., 2019), annual mean sea-ice concentration in E3SMv2 is highest in the southwestern Weddell Sea (Fig. 8a). In winter (June-August), large areas of the

high-latitude Southern Ocean are fully ice covered, implying a lack of exact position data for Core Argo floats. The trajectory of an example synthetic float in E3SMv2 reveals that the mismatch in the 1-year trajectory length can be substantial (58% in the example in Fig. 8b, assuming unknown float positions when sea-ice cover exceeds 50%). As the linear interpolation of position data also re-locates the associated data of physical and biogeochemical water-mass properties (as illustrated with nitrate in Fig. 8b), this procedure adds uncertainty to the float-derived tracer distribution field in the sea-ice zone, and challenges our ability to accurately determine decorrelation length scales of these tracers (as done in e.g., Eveleth et al., 2017).

The difference in under-ice trajectory length depends on the threshold used to identify critical sea-ice conditions and on the location of the float (Fig. 8c). The difference in trajectory length is larger the more often a float encounters critical sea-ice conditions. Thus, the higher average sea-ice concentration in the Weddell Sea than in other sectors explains the largest average differences in trajectory length in this region ($\sim$40% for sea ice concentration thresholds $\leq$10%, see Fig. 8c). Even for a sea-ice concentration threshold of 50%, trajectories are between 10% (East Antarctic) and 25% (Weddell Sea) longer when accounting for all floats' true positions than when linearly interpolating positions under sea ice cover. Given that Core Argo floats aim to avoid any direct contact with sea ice to minimize the risk of damage, their ice-avoidance procedures are very risk-averse (Klatt et al., 2007), which corresponds to a low sea-ice concentration threshold for surfacing in Fig. 8. Our analysis of the synthetic floats in E3SMv2 showcases the possibly large uncertainty in float trajectories when float positions for such conditions are unknown. Alternative approaches to locating floats under sea ice, such as via acoustic tracking (Klatt et al., 2007; Chamberlain et al., 2022) or via contours of potential vorticity, sea level, or density (Chamberlain et al., 2018; Oke et al., 2022) offer the potential to reduce errors in the hydrographic measurements (Nguyen et al., 2020) and the geopositioning of the associated tracer data.

3.3.4 Case study IV: Deriving phytoplankton phenology from biogeochemical floats

Biogeochemical floats can provide new information about surface and subsurface phytoplankton abundance, helping to elucidate their role in global carbon and nutrient cycling. In the Southern Ocean, biogeochemical floats have already provided a more detailed description of phytoplankton phenology (including under sea-ice cover, see e.g., Arteaga et al., 2020; Hague and Vichi, 2021), phytoplankton biomass loss (Moreau et al., 2020), and net community production (e.g., Johnson et al., 2017; Su et al., 2021). Similar advances are expected in other regions (e.g., Cornec et al., 2021) as more biogeochemical floats are deployed globally (see https://www.go-bgc.org/array-status; last access August 3). In this case study, we assess the ability of float networks differing in float density to capture subsurface phytoplankton bloom characteristics in the subtropical Pacific, where subsurface maxima in phytoplankton biomass are commonly observed (Cornec et al., 2021; Yasunaka et al., 2022).

In the subtropical Pacific, E3SMv2 peak bloom magnitude and timing is different in the surface and subsurface ocean (Fig. 9). Peak biomass levels display more year-to-year variability and are up to two times higher at 30 m (black and grey lines in Fig. 9a) than at 60 m (dark and light blue lines). Further, based on 10-day averages, maximum biomass is simulated to occur earlier in the year at 30 m than at 60 m, possibly illustrating differences in environmental factors such as light and nutrient availability in driving phytoplankton dynamics throughout the year (Cornec et al., 2021). Overall, the spatio-temporal
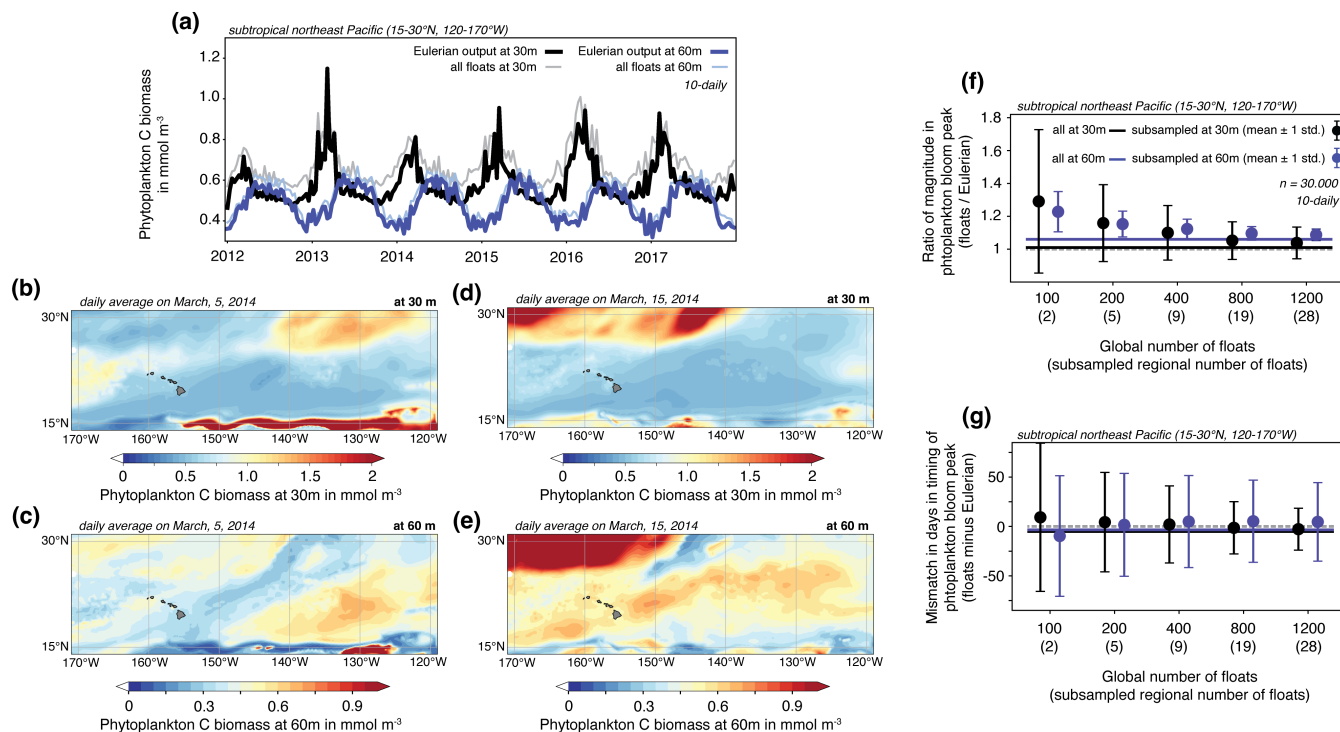
Figure 9. Case study "Phytoplankton phenology": (a) Time series from 2012-2017 of average 10-day total phytoplankton carbon biomass in mmol m^{-3} in the subtropical northeast Pacific between 15-30°N and 120-170°W at 30 m in the full Eulerian output (black) and based on all available floats (light grey). The corresponding time series at 60 m are displayed in dark blue and light blue, respectively. (b)-(e) Example fields of daily averaged total phytoplankton carbon biomass for (b)-(c) March 5, 2014 and (d)-(e) March 15, 2014 in the subtropical northeast Pacific at (b),(d) 30 m and (c),(e) 60 m. (f) Ratio of the annual peak phytoplankton carbon biomass (mean ± one standard deviation) between the 10-day float output subsampled to different global float densities (x axis) and the full Eulerian 10-day output at 30 m (black) and at 60 m (blue) in the subtropical northeast Pacific. The subsampling is repeated 5000 times, and the ratio is thus quantified 30000 times (6 years of data) for each global float density. Horizontal lines denote the mismatch between the full float output and the Eulerian output (197 floats). (g) Same as (f), but for the mismatch in days in the timing of peak phytoplankton carbon biomass.

variability is high at both depth levels (Fig. 9a-e), suggesting that a certain number of floats is needed to adequately capture large-scale bloom dynamics in this region.

The peak bloom magnitude and timing in the subtropical northeast Pacific is dependent on the total number of floats sampling the region (Fig. 9f & g). We conduct random subsampling of the full set of 197 available 6-year synthetic float time series in the area 5,000 times for different float densities. For the vast majority of repetitions in the subsampling exercise, the magnitude of the bloom peak derived from the synthetic floats is larger than the true bloom peak derived from the Eulerian output (ratio >1 in Fig. 9f). Increasing the float density improves the ability of the float network to capture the true magnitude of the bloom peak at both depths (Fig. 9f). While the mismatch at float densities of ≤200 global floats is larger for the bloom at 30 m, the

opposite holds for float densities larger than this. At 60 m, even the full synthetic float network in this region, whose density

325 corresponds to 8400 floats globally, overestimates the magnitude of the bloom peak by ∼5% (blue horizontal line), while it is well captured at 30 m (black horizontal line). On average, the timing of the bloom is reasonably well captured for all float densities (±10 days at most; Fig. 9g). However, the variability around the mean is large for both depths and all float densities (one standard deviation ranges from 21 to 75 days at 30 m and from 40 to 61 days at 60 m across float densities; see whiskers in Fig. 9g). Given that the bloom duration in the ocean typically amounts to a few weeks (see Fig. 9 and e.g., Soppa et al., 2016;

330 Silva et al., 2021), a misrepresentation of the bloom timing of >3 weeks combined with an error in the magnitude of >10% complicate the detection of any long-term trend. In this context, we further note that the analysis of phytoplankton bloom phenology based on float-derived 10-day snapshots of phytoplankton biomass, as done here to mimic the typical real-world sampling frequency, possibly masks substantial biomass variability at sub-10-day time scales (compare panels b and d as well as panels c and e in Fig. 9). While we focus our analysis on the uncertainty in deriving phytoplankton bloom characteristics

335 stemming from the float density in a given focus area, the synthetic float observations should also be used in future work to assess the uncertainty related to the sampling frequency. Our subtropical Pacific-focused case study highlights the spatio-temporal variability of phytoplankton dynamics in this region (McKee et al., 2022) and underlines the difficulty to adequately capture bloom dynamics on large spatial scales from sparse measurements. Future studies should investigate the magnitude of these uncertainties in other ocean regions.

340 ## 4 Limitations and future work

While biogeochemical float capabilities in E3SMv2 are a remarkable new tool for the ocean modeling and observational communities, the synthetic floats sample the model fields somewhat differently than Argo floats sample the real ocean. Argo floats can be laterally displaced by ocean circulation while profiling the upper ocean and while transmitting data at the ocean surface. This study does not account for the possibility of lateral displacement in the synthetic floats during this profiling,

345 adding uncertainty to the comparison of velocity estimates derived from synthetic and Argo floats (Fig. 3). Previous work has quantified the effect of velocity shear to amount to up to 1.2 cm s^{-1} in the tropics, leading to an uncertainty of ±8° in the current direction (Wang et al., 2022). Acknowledging that the absence of this effect in E3SMv2 likely has to be considered for regional applications of the synthetic floats, we assume this shortcoming to be of lesser importance for basin-scale applications presented here. In contrast to Argo floats, all synthetic floats in E3SMv2 sample the water column at the same time each day,

350 i.e., midnight Greenwich Mean Time. This means that the sampling of all variables undergoing strong diurnal fluctuations, e.g., light-sensitive processes such as biological productivity and hence biomass, is skewed towards a particular phase of the respective diurnal cycle, increasing the discrepancy between float-derived estimates and daily-averaged model output. As the seasonal cycle directly impacts the variability in the diurnal cycle over the course of the year, this effect is expected to be less pronounced in tropical regions (see section 3.3.4) than in polar regions.

355 For this study, synthetic floats in E3SMv2 are seeded uniformly as a function of grid resolution (Fig. 2). While a uniform ocean coverage by Argo float is the ultimate goal (Roemmich et al., 2019), achieving this is complicated by the dependence

on ships for float deployment. In addition, in contrast to the unlimited life time of synthetic floats, the typical life time of today's Argo floats is ∼5-7 years (typically lower for BGC Argo than Core Argo; Riser et al., 2018; Roemmich et al., 2019), causing spatial gaps in our observing system if a timely re-deployment of a float is not possible for a given region. While we did not account for random failures of some sensors or entire floats in the 6-year proof-of-concept simulation analyzed for this paper, "imperfect" synthetic float data sets could easily be constructed offline to assess the impact of a temporary or permanent absence of observations in a specific region or for a specific time period. Future work should assess the impact of data scarcity and of using different mapping methods on reconstructed fields of biogeochemical tracers and, e.g., air-sea CO_2 fluxes (Gloege et al., 2021; Hauck et al., 2023; Heimdal et al., 2023). Lastly, we note that on small spatial scales, structural limitations of the E3SMv2 configuration used here due to e.g., parametrizations of internal wave dynamics, could reduce the utility of the synthetic biogoechemical float capabilities as an ideal test bed, as the simulated variability on small spatial scales might differ from the variability experienced by real-world Argo floats.

5 Conclusions

We implement synthetic biogeochemical floats into the Energy Exascale Earth System Model version 2 (E3SMv2). The synthetic floats are advected with ocean circulation at 1000 m online during the model integration, sampling all desired prognostic and diagnostic model variables throughout the whole water column at the frequency prescribed by the end-user. Using E3SMv2 with synthetic floats as a perfect test bed, in which the true distribution of modeled physical, biogeochemical and biological variables is known, we demonstrated the utility of this new tool in different use cases. In particular, we demonstrated the impact of 1) sampling frequency on float-derived velocity estimates, 2) short-term variability on estimates of the seasonal amplitude of ecosystem stressors, 3) sampling density on float-based estimates of temporal changes in deep-ocean temperature, salinity, and oxygen, 4) the presence of sea-ice cover on derived under-ice float trajectories, and 5) sampling density on the characterization of phytoplankton bloom dynamics.

Even though differences exist in how synthetic and Argo floats sample the simulated and real ocean, respectively, the synthetic floats in E3SMv2 can be used in the future to improve our understanding of how Argo floats "see" the ocean, thereby contributing to the interpretation of existing observational records. Computing synthetic floats online eliminates both the need to store high-frequency model output and the uncertainty associated with time-averaging model output to extract synthetic observations offline. Given that Argo floats do not sample time-averaged water-mass properties but provide a snapshot view of the ocean, the online computation produces a more realistic data set of synthetic observations. Expanding the online synthetic observing capabilities in E3SMv2 or other models by further sampling methodologies, e.g., ship-based hydrography, deep-sea moorings, gliders or surface drifters, should be a key focus of future work, with the aim to improve our global ocean observing system.

Code and data availability. The model source code of E3SMv2 including synthetic biogeochemical float capabilities is available at Zenodo: Energy Exascale Earth System Model Program (2023). The synthetic biogeochemical float data set and the corresponding Eulerian model fields from E3SMv2 are deposited in the "PetaLibrary" of the University of Colorado Boulder and can be accessed via Globus (www.globus.org). To find the data, enter "E3SM-BGCArgo" as the name of the collection in the file manager.

Author contributions. C.N., N.S.L., M.M., and A.R.G. conceived the study. C.N. performed the analysis and wrote the manuscript. N.S.L. and A.R.G. acquired the funding. M.M. implemented the synthetic floats into E3SMv2 and performed the test simulation analyzed here. Y.T. performed the Eulerian model simulation serving as the spin up for the test simulation with synthetic floats. All authors gave input on the case studies and commented on the manuscript.

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

References

Allison, L. C., Roberts, C. D., Palmer, M. D., Hermanson, L., Killick, R. E., Rayner, N. A., Smith, D. M., and Andrews, M. B.: Towards quantifying uncertainty in ocean heat content changes using synthetic profiles, Environmental Research Letters, 14, 084 037, https://doi.org/10.1088/1748-9326/ab2b0b, 2019.

Argo: Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) - Snapshot of Argo GDAC of March 10st 2023, SEANOE, https://doi.org/10.17882/42182#100487, 2023.

Arteaga, L. A., Boss, E., Behrenfeld, M. J., Westberry, T. K., and Sarmiento, J. L.: Seasonal modulation of phytoplankton biomass in the Southern Ocean, Nature Communications, 11, 5364, https://doi.org/10.1038/s41467-020-19157-2, 2020.

Behrenfeld, M. J. and Falkowski, P. G.: Photosynthetic rates derived from satellite-based chlorophyll concentration, Limnology and Oceanography, 42, 1–20, https://doi.org/10.4319/lo.1997.42.1.0001, 1997.

Blunden, J. and Arndt, D. S.: State of the Climate in 2016, Bulletin of the American Meteorological Society, 98, Si–S280, https://doi.org/10.1175/2017BAMSStateoftheClimate.1, 2017.

Brady, R. X., Maltrud, M. E., Wolfram, P. J., Drake, H. F., and Lovenduski, N. S.: The Influence of Ocean Topography on the Upwelling of Carbon in the Southern Ocean, Geophysical Research Letters, 48, 1–11, https://doi.org/10.1029/2021GL095088, 2021.

Burrows, S. M. et al.: The DOE E3SM v1.1 Biogeochemistry Configuration: Description and Simulated Ecosystem-Climate Responses to Historical Changes in Forcing, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/https://doi.org/10.1029/2019MS001766, 2020.

Chai, F., Johnson, K. S., Claustre, H., Xing, X., Wang, Y., Boss, E., Riser, S., Fennel, K., Schofield, O., and Sutton, A.: Monitoring ocean biogeochemistry with autonomous platforms, Nature Reviews Earth & Environment, 1, 315–326, https://doi.org/10.1038/s43017-020-0053-y, 2020.

Chamberlain, P., Cornuelle, B., Talley, L. D., Speer, K., Hancock, C., and Riser, S.: Acoustic Float Tracking with the Kalman Smoother, Journal of Atmospheric and Oceanic Technology, 40, 15–35, https://doi.org/10.1175/JTECH-D-21-0063.1, 2022.

Chamberlain, P. M., Talley, L. D., Mazloff, M. R., Riser, S. C., Speer, K., Gray, A. R., and Schwartzman, A.: Observing the Ice-Covered Weddell Gyre With Profiling Floats: Position Uncertainties and Correlation Statistics, Journal of Geophysical Research: Oceans, 123, 8383–8410, https://doi.org/10.1029/2017JC012990, 2018.

Claustre, H., Johnson, K. S., and Takeshita, Y.: Observing the Global Ocean with Biogeochemical-Argo, Annual Review of Marine Science, 12, 23–48, https://doi.org/10.1146/annurev-marine-010419-010956, 2020.

Clow, G. L., Lovenduski, N. S., Levy, M. N., Lindsay, K., and Kay, E.: The utility of simulated ocean chlorophyll observations: a case study with the Chlorophyll Observation Simulator Package (version 1) in CESMv2.2, Geoscientific Model Development Discussion, pp. 1–31, https://doi.org/10.5194/gmd-2023-143, 2023.

Cornec, M., Claustre, H., Mignot, A., Guidi, L., Lacour, L., Poteau, A., D'Ortenzio, F., Gentili, B., and Schmechtig, C.: Deep Chlorophyll Maxima in the Global Ocean: Occurrences, Drivers and Characteristics, Global Biogeochemical Cycles, 35, 1–30, https://doi.org/10.1029/2020GB006759, 2021.

Droste, E. S., Hoppema, M., González-Dávila, M., Santana-Casiano, J. M., Queste, B. Y., Dall'Olmo, G., Venables, H. J., Rohardt, G., Ossebaar, S., Schuller, D., Trace-Kleeberg, S., and Bakker, D. C. E.: The influence of tides on the marine carbonate chemistry of a coastal polynya in the south-eastern Weddell Sea, Ocean Science, 18, 1293–1320, https://doi.org/10.5194/os-18-1293-2022, 2022.

Eayrs, C., Holland, D., Francis, D., Wagner, T., Kumar, R., and Li, X.: Understanding the Seasonal Cycle of Antarctic Sea Ice Extent in the Context of Longer-Term Variability, Reviews of Geophysics, 57, 1037–1064, https://doi.org/10.1029/2018RG000631, 2019.

445 Energy Exascale Earth System Model Program: Energy Exascale Earth System Model (E3SMv2) code with Argo float simulator, https://doi.org/10.5281/zenodo.10094349, Zenodo, 2023.

Eveleth, R., Cassar, N., Doney, S. C., Munro, D. R., and Sweeney, C.: Biological and physical controls on O2/Ar, Ar and pCO_2 variability at the Western Antarctic Peninsula and in the Drake Passage, Deep-Sea Research Part II: Topical Studies in Oceanography, 139, 77–88, https://doi.org/10.1016/j.dsr2.2016.05.002, 2017.

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

Garry, F. K., McDonagh, E. L., Blaker, A. T., Roberts, C. D., Desbruyères, D. G., Frajka-Williams, E., and King, B. A.: Model-Derived Uncertainties in Deep Ocean Temperature Trends Between 1990 and 2010, Journal of Geophysical Research: Oceans, 124, 1155–1169, https://doi.org/10.1029/2018JC014225, 2019.

455 Gasparin, F., Hamon, M., Rémy, E., and Le Traon, P.-Y.: How Deep Argo Will Improve the Deep Ocean in an Ocean Reanalysis, Journal of Climate, 33, 77–94, https://doi.org/10.1175/JCLI-D-19-0208.1, 2020.

Gloege, L., McKinley, G. A., Landschützer, P., Fay, A. R., Frölicher, T. L., Fyfe, J. C., Ilyina, T., Jones, S., Lovenduski, N. S., Rodgers, K. B., Schlunegger, S., and Takano, Y.: Quantifying Errors in Observationally Based Estimates of Ocean Carbon Sink Variability, Global Biogeochemical Cycles, 35, 1–14, https://doi.org/10.1029/2020GB006788, 2021.

460 Golaz, J., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H., Lin, W., Lipscomb, W. H., Ma, P., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy,

465 R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Reeves Eyre, J. E. J., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X., Singh, B., Tang, J., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H., Wang, S., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S., Yang, Y., Yoon, J., Zelinka, M. D., Zender, C. S., Zeng, X., Zhang, C., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, Journal of Advances in Modeling Earth Systems, 11, 2089–2129, https://doi.org/10.1029/2018MS001603, 2019.

470 Golaz, J. C., Van Roekel, L. P., Zheng, X., Roberts, A. F., Wolfe, J. D., Lin, W., Bradley, A. M., Tang, Q., Maltrud, M. E., Forsyth, R. M., Zhang, C., Zhou, T., Zhang, K., Zender, C. S., Wu, M., Wang, H., Turner, A. K., Singh, B., Richter, J. H., Qin, Y., Petersen, M. R., Mametjanov, A., Ma, P. L., Larson, V. E., Krishna, J., Keen, N. D., Jeffery, N., Hunke, E. C., Hannah, W. M., Guba, O., Griffin, B. M., Feng, Y., Engwirda, D., Di Vittorio, A. V., Dang, C., Conlon, L. A. M., Chen, C. C. J., Brunke, M. A., Bisht, G., Benedict, J. J., Asay-Davis, X. S., Zhang, Y., Zhang, M., Zeng, X., Xie, S., Wolfram, P. J., Vo, T., Veneziani, M., Tesfa, T. K., Sreepathi, S., Salinger, A. G.,

475 Reeves Eyre, J. E., Prather, M. J., Mahajan, S., Li, Q., Jones, P. W., Jacob, R. L., Huebler, G. W., Huang, X., Hillman, B. R., Harrop, B. E., Foucar, J. G., Fang, Y., Comeau, D. S., Caldwell, P. M., Bartoletti, T., Balaguru, K., Taylor, M. A., McCoy, R. B., Leung, L. R., and Bader, D. C.: The DOE E3SM Model Version 2: Overview of the Physical Model and Initial Model Evaluation, Journal of Advances in Modeling Earth Systems, 14, 1–51, https://doi.org/10.1029/2022MS003156, 2022.

Gray, A. R., Johnson, K. S., Bushinsky, S. M., Riser, S. C., Russell, J. L., Talley, L. D., Wanninkhof, R., Williams, N. L., and Sarmiento, J. L.:
480   Autonomous Biogeochemical Floats Detect Significant Carbon Dioxide Outgassing in the High-Latitude Southern Ocean, Geophysical
       Research Letters, 45, 9049–9057, https://doi.org/10.1029/2018GL078013, 2018.

Gruber, N., Boyd, P. W., Frölicher, T. L., and Vogt, M.: Biogeochemical extremes and compound events in the ocean, Nature, 600, 395–407,
       https://doi.org/10.1038/s41586-021-03981-7, 2021.

Hague, M. and Vichi, M.: Southern Ocean Biogeochemical Argo detect under-ice phytoplankton growth before sea ice retreat, Biogeo-
485   sciences, 18, 25–38, https://doi.org/10.5194/bg-18-25-2021, 2021.

Hauck, J., Nissen, C., Landschützer, P., Rödenbeck, C., Bushinsky, S., and Olsen, A.: Sparse observations induce large biases in estimates of
       the global ocean CO_2 sink: An ocean model subsampling experiment, Philosophical Transactions of the Royal Society A: Mathematical,
       Physical and Engineering Sciences, 381, https://doi.org/10.1098/rsta.2022.0063, 2023.

Heimdal, T. H., McKinley, G. A., Sutton, A. J., Fay, A. R., and Gloege, L.: Assessing improvements in global ocean
490   pCO2 machine learning reconstructions with Southern Ocean autonomous sampling, Biogeosciences Discussions, pp. 1–35,
       https://doi.org/https://doi.org/10.5194/bg-2023-160, 2023.

Helmuth, B., Russell, B. D., Connell, S. D., Dong, Y., Harley, C. D., Lima, F. P., Sará, G., Williams, G. A., and Mieszkowska,
       N.: Beyond long-term averages: making biological sense of a rapidly changing world, Climate Change Responses, 1, 6,
       https://doi.org/10.1186/s40665-014-0006-0, 2014.

495 Hoppema, M.: Weddell Sea is a globally significant contributor to deep-sea sequestration of natural carbon dioxide, Deep Sea Research Part
       I: Oceanographic Research Papers, 51, 1169–1177, https://doi.org/10.1016/j.dsr.2004.02.011, 2004.

Jayne, S. R., Roemmich, D., Zilberman, N., Riser, S. C., Johnson, K. S., Johnson, G. C., and Piotrowicz, S. R.: The argo program: Present
       and future, Oceanography, 30, 18–28, https://doi.org/10.5670/OCEANOG.2017.213, 2017.

Jeffery, N., Maltrud, M. E., Hunke, E. C., Wang, S., Wolfe, J., Turner, A. K., Burrows, S. M., Shi, X., Lipscomb, W. H., Maslowski,
500   W., and Calvin, K. V.: Investigating controls on sea ice algal production using E3SMv1.1-BGC, Annals of Glaciology, 61, 51–72,
       https://doi.org/10.1017/aog.2020.7, 2020.

Johnson, G. C., Lyman, J. M., and Purkey, S. G.: Informing Deep Argo Array Design Using Argo and Full-Depth Hydrographic Section
       Data, Journal of Atmospheric and Oceanic Technology, 32, 2187–2198, https://doi.org/10.1175/JTECH-D-15-0139.1, 2015.

Johnson, G. C., Hosoda, S., Jayne, S. R., Oke, P. R., Riser, S. C., Roemmich, D., Suga, T., Thierry, V., Wijffels, S. E., and Xu, J.: Argo—Two
505   Decades: Global Oceanography, Revolutionized, Annual Review of Marine Science, 14, 379–403, https://doi.org/10.1146/annurev-
       marine-022521-102008, 2022.

Johnson, K. S., Plant, J. N., Dunne, J. P., Talley, L. D., and Sarmiento, J. L.: Annual nitrate drawdown observed by SOCCOM pro-
       filing floats and the relationship to annual net community production, Journal of Geophysical Research: Oceans, 122, 6668–6683,
       https://doi.org/10.1002/2017JC012839, 2017.

510 Kapsenberg, L. and Cyronak, T.: Ocean acidification refugia in variable environments, Global Change Biology, 25, 3201–3214,
       https://doi.org/10.1111/gcb.14730, 2019.

Kawai, Y. and Wada, A.: Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: A review, Journal of
       Oceanography, 63, 721–744, https://doi.org/10.1007/s10872-007-0063-0, 2007.

King, B., Thierry, V., Zilberman, N., and Walicka, K.: 3rd Deep-Argo Workshop Report, Tech. rep., accessed via https://www.euro-
515   argo.eu/content/download/159856/file/DeepArgoWorkshop2021-Report.pdf on Oct 20, 2023, 2021.

Klatt, O., Boebel, O., and Fahrbach, E.: A Profiling Float's Sense of Ice, Journal of Atmospheric and Oceanic Technology, 24, 1301–1308, https://doi.org/10.1175/JTECH2026.1, 2007.

Long, M. C., Moore, J. K., Lindsay, K., Levy, M., Doney, S. C., Luo, J. Y., Krumhardt, K. M., Letscher, R. T., Grover, M., and Sylvester, Z. T.: Simulations With the Marine Biogeochemistry Library (MARBL), Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2021MS002647, 2021.

Matsumoto, G. I., Johnson, K. S., Riser, S., Talley, L., Wijffels, S., and Hotinski, R.: The Global Ocean Biogeochemistry (GO-BGC) Array of Profiling Floats to Observe Changing Ocean Chemistry and Biology, Marine Technology Society Journal, 56, 122–123, https://doi.org/10.4031/MTSJ.56.3.25, 2022.

McKee, D. C., Doney, S. C., Della Penna, A., Boss, E. S., Gaube, P., Behrenfeld, M. J., and Glover, D. M.: Lagrangian and Eulerian time and length scales of mesoscale ocean chlorophyll from Bio-Argo floats and satellites, Biogeosciences, 19, 5927–5952, https://doi.org/10.5194/bg-19-5927-2022, 2022.

Meinshausen, M., Vogel, E., Nauels, A., Lorbacher, K., Meinshausen, N., Etheridge, D. M., Fraser, P. J., Montzka, S. A., Rayner, P. J., Trudinger, C. M., Krummel, P. B., Beyerle, U., Canadell, J. G., Daniel, J. S., Enting, I. G., Law, R. M., Lunder, C. R., O'Doherty, S., Prinn, R. G., Reimann, S., Rubino, M., Velders, G. J. M., Vollmer, M. K., Wang, R. H. J., and Weiss, R.: Historical greenhouse gas concentrations for climate modelling (CMIP6), Geoscientific Model Development, 10, 2057–2116, https://doi.org/10.5194/gmd-10-2057-2017, 2017.

Moore, J. K., Doney, S. C., and Lindsay, K.: Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Global Biogeochemical Cycles, 18, GB4028, https://doi.org/10.1029/2004GB002220, 2004.

Moore, J. K., Lindsay, K., Doney, S. C., Long, M. C., and Misumi, K.: Marine Ecosystem Dynamics and Biogeochemical Cycling in the Community Earth System Model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 Scenarios, Journal of Climate, 26, 9291–9312, https://doi.org/10.1175/JCLI-D-12-00566.1, 2013.

Moreau, S., Boyd, P. W., and Strutton, P. G.: Remote assessment of the fate of phytoplankton in the Southern Ocean sea-ice zone, Nature Communications, 11, 3108, https://doi.org/10.1038/s41467-020-16931-0, 2020.

Nguyen, A. T., Heimbach, P., Garg, V. V., Ocaña, V., Lee, C., and Rainville, L.: Impact of Synthetic Arctic Argo-Type Floats in a Coupled Ocean–Sea Ice State Estimation Framework, Journal of Atmospheric and Oceanic Technology, 37, 1477–1495, https://doi.org/10.1175/JTECH-D-19-0159.1, 2020.

Nicholson, S.-A., Whitt, D. B., Fer, I., du Plessis, M. D., Lebéhot, A. D., Swart, S., Sutton, A. J., and Monteiro, P. M. S.: Storms drive outgassing of CO_2 in the subpolar Southern Ocean, Nature Communications, 13, 158, https://doi.org/10.1038/s41467-021-27780-w, 2022.

Oke, P. R., Rykova, T., Pilo, G. S., and Lovell, J. L.: Estimating Argo Float Trajectories Under Ice, Earth and Space Science, 9, 1–16, https://doi.org/10.1029/2022EA002312, 2022.

Ollitrault, M. and Rannou, J.-P.: ANDRO: An Argo-Based Deep Displacement Dataset, Journal of Atmospheric and Oceanic Technology, 30, 759–788, https://doi.org/10.1175/JTECH-D-12-00073.1, 2013.

Orr, J. C. and Epitalon, J.-M.: Improved routines to model the ocean carbonate system: mocsy 2.0, Geoscientific Model Development, 8, 485–499, https://doi.org/10.5194/gmd-8-485-2015, 2015.

Petersen, M. R., Asay-Davis, X. S., Berres, A. S., Chen, Q., Feige, N., Hoffman, M. J., Jacobsen, D. W., Jones, P. W., Maltrud, M. E., Price, S. F., Ringler, T. D., Streletz, G. J., Turner, A. K., Van Roekel, L. P., Veneziani, M., Wolfe, J. D., Wolfram, P. J., and Woodring, J. L.: An Evaluation of the Ocean and Sea Ice Climate of E3SM Using MPAS and Interannual CORE-II Forcing, Journal of Advances in Modeling Earth Systems, p. 2018MS001373, https://doi.org/10.1029/2018MS001373, 2019.

Prend, C. J., Gray, A. R., Talley, L. D., Gille, S. T., Haumann, F. A., Johnson, K. S., Riser, S. C., Rosso, I., Sauvé, J., and
555   Sarmiento, J. L.: Indo-Pacific Sector Dominates Southern Ocean Carbon Outgassing, Global Biogeochemical Cycles, 36, 1–22,
       https://doi.org/10.1029/2021GB007226, 2022a.

Prend, C. J., Hunt, J. M., Mazloff, M. R., Gille, S. T., and Talley, L. D.: Controls on the Boundary Between Thermally and Non-Thermally
       Driven pCO$_2$ Regimes in the South Pacific, Geophysical Research Letters, 49, 1–11, https://doi.org/10.1029/2021GL095797, 2022b.

Prend, C. J., Keerthi, M. G., Lévy, M., Aumont, O., Gille, S. T., and Talley, L. D.: Sub-Seasonal Forcing Drives Year-To-Year Variations of
560   Southern Ocean Primary Productivity, Global Biogeochemical Cycles, 36, 1–15, https://doi.org/10.1029/2022GB007329, 2022c.

Rhein, M., Steinfeldt, R., Kieke, D., Stendardo, I., and Yashayaev, I.: Ventilation variability of Labrador Sea Water and its impact on oxygen
       and anthropogenic carbon: a review, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,
       375, 20160 321, https://doi.org/10.1098/rsta.2016.0321, 2017.

Ringler, T., Petersen, M., Higdon, R. L., Jacobsen, D., Jones, P. W., and Maltrud, M.: A multi-resolution approach to global ocean modeling,
565   Ocean Modelling, 69, 211–232, https://doi.org/10.1016/j.ocemod.2013.04.010, 2013.

Riser, S. C., Swift, D., and Drucker, R.: Profiling Floats in SOCCOM: Technical Capabilities for Studying the Southern Ocean, Journal of
       Geophysical Research: Oceans, 123, 4055–4073, https://doi.org/10.1002/2017JC013419, 2018.

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., Oke, P., Owens, W. B., Pouliquen, S., Purkey, S., Scanderbeg,
       M., Suga, T., Wijffels, S., Zilberman, N., Bakker, D., Baringer, M., Belbeoch, M., Bittig, H. C., Boss, E., Calil, P., Carse, F., Carval, T.,
570   Chai, F., Conchubhair, D. Ó., D'Ortenzio, F., Dall'Olmo, G., Desbruyeres, D., Fennel, K., Fer, I., Ferrari, R., Forget, G., Freeland, H.,
       Fujiki, T., Gehlen, M., Greenan, B., Hallberg, R., Hibiya, T., Hosoda, S., Jayne, S., Jochum, M., Johnson, G. C., Kang, K., Kolodziejczyk,
       N., Körtzinger, A., Traon, P.-Y. L., Lenn, Y.-D., Maze, G., Mork, K. A., Morris, T., Nagai, T., Nash, J., Garabato, A. N., Olsen, A., Pattabhi,
       R. R., Prakash, S., Riser, S., Schmechtig, C., Schmid, C., Shroyer, E., Sterl, A., Sutton, P., Talley, L., Tanhua, T., Thierry, V., Thomalla, S.,
       Toole, J., Troisi, A., Trull, T. W., Turton, J., Velez-Belchi, P. J., Walczowski, W., Wang, H., Wanninkhof, R., Waterhouse, A. F., Waterman,
575   S., Watson, A., Wilson, C., Wong, A. P. S., Xu, J., and Yasuda, I.: On the Future of Argo: A Global, Full-Depth, Multi-Disciplinary Array,
       Frontiers in Marine Science, 6, 1–28, https://doi.org/10.3389/fmars.2019.00439, 2019.

Sarmiento, J. L. and Gruber, N.: Ocean Biogeochemical Dynamics, Princeton University Press, 1st edn., 2006.

Sarmiento, J. L., Johnson, K. S., Arteaga, L. A., Bushinsky, S. M., Cullen, H. M., Gray, A. R., Hotinski, R. M., Maurer, T. L., Mazloff, M. R.,
       Riser, S. C., Russell, J. L., Schofield, O. M., and Talley, L. D.: The Southern Ocean Carbon and Climate Observations and Modeling
580   (SOCCOM) project: A review, Progress in Oceanography, 7, 103 130, https://doi.org/10.1016/j.pocean.2023.103130, 2023.

Schofield, O., Fassbender, A., Hood, M., Hill, K., and Johnson, K.: A global ocean biogeochemical observatory becomes a reality, Eos, 103,
       https://doi.org/10.1029/2022EO220149, 2022.

Silva, E., Counillon, F., Brajard, J., Korosov, A., Pettersson, L. H., Samuelsen, A., and Keenlyside, N.: Twenty-One Years
       of Phytoplankton Bloom Phenology in the Barents, Norwegian, and North Seas, Frontiers in Marine Science, 8, 1–16,
585   https://doi.org/10.3389/fmars.2021.746327, 2021.

SOCCOM: "Southern Ocean Carbon and Climate Observations and Modeling" Project, https://soccom.princeton.edu, accessed: 2023-08-31,
       2023.

Soppa, M., Völker, C., and Bracher, A.: Diatom Phenology in the Southern Ocean: Mean Patterns, Trends and the Role of Climate Oscilla-
       tions, Remote Sensing, 8, 420, https://doi.org/10.3390/rs8050420, 2016.

590 Su, F., Fan, R., Yan, F., Meadows, M., Lyne, V., Hu, P., Song, X., Zhang, T., Liu, Z., Zhou, C., Pei, T., Yang, X., Du, Y., Wei, Z., Wang, F., Qi, Y., and Chai, F.: Widespread global disparities between modelled and observed mid-depth ocean currents, Nature Communications, 14, 2089, https://doi.org/10.1038/s41467-023-37841-x, 2023.

Su, J., Strutton, P. G., and Schallenberg, C.: The subsurface biological structure of Southern Ocean eddies revealed by BGC-Argo floats, Journal of Marine Systems, 220, 103 569, https://doi.org/10.1016/j.jmarsys.2021.103569, 2021.

595 Takano, Y., Maltrud, M., Sinha, A., Jeffery, N., Smith, K., Conlon, L., Wolfe, J., and Petersen, M.: Global Ocean Carbon Cycle Simulations with the 2 E3SM version 2 (E3SMv2), https://doi.org/10.5281/zenodo.10093369, Zenodo, 2023.

Talley, L., Feely, R., Sloyan, B., Wanninkhof, R., Baringer, M., Bullister, J., Carlson, C., Doney, S., Fine, R., Firing, E., Gruber, N., Hansell, D., Ishii, M., Johnson, G., Katsumata, K., Key, R., Kramp, M., Langdon, C., Macdonald, A., Mathis, J., McDonagh, E., Mecking, S., Millero, F., Mordy, C., Nakano, T., Sabine, C., Smethie, W., Swift, J., Tanhua, T., Thurnherr, A., Warner, M., and Zhang, J.-Z.: Changes

600 in Ocean Heat, Carbon Content, and Ventilation: A Review of the First Decade of GO-SHIP Global Repeat Hydrography, Annual Review of Marine Science, 8, 185–215, https://doi.org/10.1146/annurev-marine-052915-100829, 2016.

Torres, O., Kwiatkowski, L., Sutton, A. J., Dorey, N., and Orr, J. C.: Characterizing Mean and Extreme Diurnal Variability of Ocean CO_2 System Variables Across Marine Environments, Geophysical Research Letters, 48, 1–12, https://doi.org/10.1029/2020GL090228, 2021.

Tsujino, H., Urakawa, S., Nakano, H., Small, R. J., Kim, W. M., Yeager, S. G., Danabasoglu, G., Suzuki, T., Bamber, J. L., Bentsen, M.,

605 Böning, C. W., Bozec, A., Chassignet, E. P., Curchitser, E., Boeira Dias, F., Durack, P. J., Griffies, S. M., Harada, Y., Ilicak, M., Josey, S. A., Kobayashi, C., Kobayashi, S., Komuro, Y., Large, W. G., Le Sommer, J., Marsland, S. J., Masina, S., Scheinert, M., Tomita, H., Valdivieso, M., and Yamazaki, D.: JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do), Ocean Modelling, 130, 79–139, https://doi.org/10.1016/j.ocemod.2018.07.002, 2018.

Turner, A. K., Lipscomb, W. H., Hunke, E. C., Jacobsen, D. W., Jeffery, N., Engwirda, D., Ringler, T. D., and Wolfe, J. D.:

610 MPAS-Seaice (v1.0.0): sea-ice dynamics on unstructured Voronoi meshes, Geoscientific Model Development, 15, 3721–3751, https://doi.org/10.5194/gmd-15-3721-2022, 2022.

Wang, T., Du, Y., and Wang, M.: Overlooked Current Estimation Biases Arising from the Lagrangian Argo Trajectory Derivation Method, Journal of Physical Oceanography, 52, 3–19, https://doi.org/10.1175/JPO-D-20-0287.1, 2022.

Wolfram, P. J., Ringler, T. D., Maltrud, M. E., Jacobsen, D. W., and Petersen, M. R.: Diagnosing Isopycnal Diffusivity in an Eddying,

615 Idealized Midlatitude Ocean Basin via Lagrangian, in Situ, Global, High-Performance Particle Tracking (LIGHT), Journal of Physical Oceanography, 45, 2114–2133, https://doi.org/10.1175/JPO-D-14-0260.1, 2015.

Wong, A. P. S., Wijffels, S. E., Riser, S. C., Pouliquen, S., Hosoda, S., Roemmich, D., Gilson, J., Johnson, G. C., Martini, K., Murphy, D. J., Scanderbeg, M., Bhaskar, T. V. S. U., Buck, J. J. H., Merceur, F., Carval, T., Maze, G., Cabanes, C., André, X., Poffa, N., Yashayaev, I., Barker, P. M., Guinehut, S., Belbéoch, M., Ignaszewski, M., Baringer, M. O., Schmid, C., Lyman, J. M., McTaggart, K. E., Purkey, S. G.,

620 Zilberman, N., Alkire, M. B., Swift, D., Owens, W. B., Jayne, S. R., Hersh, C., Robbins, P., West-Mack, D., Bahr, F., Yoshida, S., Sutton, P. J. H., Cancouët, R., Coatanoan, C., Dobbler, D., Juan, A. G., Gourrion, J., Kolodziejczyk, N., Bernard, V., Bourlès, B., Claustre, H., D'Ortenzio, F., Le Reste, S., Le Traon, P.-Y., Rannou, J.-P., Saout-Grit, C., Speich, S., Thierry, V., Verbrugge, N., Angel-Benavides, I. M., Klein, B., Notarstefano, G., Poulain, P.-M., Vélez-Belchí, P., Suga, T., Ando, K., Iwasaska, N., Kobayashi, T., Masuda, S., Oka, E., Sato, K., Nakamura, T., Sato, K., Takatsuki, Y., Yoshida, T., Cowley, R., Lovell, J. L., Oke, P. R., van Wijk, E. M., Carse, F., Donnelly, M.,

625 Gould, W. J., Gowers, K., King, B. A., Loch, S. G., Mowat, M., Turton, J., Rama Rao, E. P., Ravichandran, M., Freeland, H. J., Gaboury, I., Gilbert, D., Greenan, B. J. W., Ouellet, M., Ross, T., Tran, A., Dong, M., Liu, Z., Xu, J., Kang, K., Jo, H., Kim, S.-D., and Park, H.-M.:

Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations From a Global Array of Profiling Floats, Frontiers in Marine Science, 7, 1–23, https://doi.org/10.3389/fmars.2020.00700, 2020.

Yasunaka, S., Ono, T., Sasaoka, K., and Sato, K.: Global distribution and variability of subsurface chlorophyll a concentrations, Ocean Science, 18, 255–268, https://doi.org/10.5194/os-18-255-2022, 2022.

630

Youngs, M. K., Freilich, M. A., and Lovenduski, N. S.: Air-Sea CO_2 Fluxes Localized by Topography in a Southern Ocean Channel, Geophysical Research Letters, 50, 1–9, https://doi.org/10.1029/2023GL104802, 2023.

Zilberman, N. V., Scanderbeg, M., Gray, A. R., and Oke, P. R.: Scripps Argo trajectory-based velocity product 2001-01 to 2020-12, Scripps Argo Trajectory-Based Velocity Product. UC San Diego Library Digital Collections, https://doi.org/10.6075/J0KD1Z35, 2022.

635

Zilberman, N. V., Scanderbeg, M., Gray, A. R., and Oke, P. R.: Scripps Argo Trajectory-Based Velocity Product: Global Estimates of Absolute Velocity Derived from Core, Biogeochemical, and Deep Argo Float Trajectories at Parking Depth, Journal of Atmospheric and Oceanic Technology, 40, 361–374, https://doi.org/10.1175/JTECH-D-22-0065.1, 2023.