



Reconstructing ocean carbon storage with CMIP6 Earth system models and synthetic Argo observations

Katherine E. Turner¹, Doug M. Smith², Anna Katavouta^{1,3}, and Richard G. Williams¹

¹Department of Earth, Ocean, and Ecological Sciences, School of Environmental Sciences University of Liverpool, Liverpool, United Kingdom

²UK Met Office, Exeter, United Kingdom

³National Oceanography Centre, Liverpool, United Kingdom

Correspondence: Katherine E. Turner (k.e.turner2@liverpool.ac.uk)

Received: 8 August 2022 – Discussion started: 18 August 2022

Revised: 9 March 2023 – Accepted: 17 March 2023 – Published: 28 April 2023

Abstract. The ocean carbon store plays a vital role in setting the carbon response to emissions and variability in the carbon cycle. However, due to the ocean's strong regional and temporal variability, sparse carbon observations limit our understanding of historical carbon changes. Ocean temperature and salinity profiles are more widespread and rapidly expanding due to autonomous programmes, and so we explore how temperature and salinity profiles can provide information to reconstruct ocean carbon inventories with ensemble optimal interpolation. Here, ensemble optimal interpolation is used to reconstruct ocean carbon using synthetic Argo temperature and salinity observations, with examples for both the top 100 m and top 2000 m carbon inventories. When considering reconstructions of the top 100 m carbon inventory, coherent relationships between upper-ocean carbon, temperature, salinity, and atmospheric CO₂ result in optimal solutions that reflect the controls of undersaturation, solubility, and alkalinity. Out-of-sample reconstructions of the top 100 m show that, in most regions, the trend in ocean carbon and over 60 % of detrended variability can be reconstructed using local temperature and salinity measurements, with only small changes when considering synthetic profiles consistent with irregular Argo sampling. Extending the method to reconstruct the upper 2000 m reveals that model uncertainties at depth limit the reconstruction skill. The impact of these uncertainties on reconstructing the carbon inventory over the upper 2000 m is small, and full reconstructions with historical Argo locations show that the method can reconstruct regional inter-annual and decadal variability. Hence, optimal interpolation based on model relationships combined with

hydrographic measurements can provide valuable information about global ocean carbon inventory changes.

1 Introduction

The global ocean plays a important role in the carbon cycle, being both a major reservoir of carbon and a substantial sink of anthropogenic carbon transferred from geological storage to the coupled atmosphere–ocean–terrestrial system. The ocean is estimated to have taken up around 26 % of anthropogenic emissions from fossil fuels and land use changes since 1850 (Khaliwala et al., 2013; DeVries, 2014; Terhaar et al., 2022; Friedlingstein et al., 2022). In addition to the long-term uptake of anthropogenic carbon, ocean carbon uptake exhibits inter-annual and decadal variability on global and regional scales (Landschützer et al., 2016; McKinley et al., 2017; Gruber et al., 2019b; McKinley et al., 2020). This variability impacts the ability to detect trends in observations of the partial pressure of CO₂ in seawater (McKinley et al., 2016) and, through oceanic transport, can lead to regionally enhanced acidification (Burger et al., 2020; Hauri et al., 2021). The ocean carbon inventory is thus an important integral measure of climate change, and up-to-date estimates of its behaviour are vital for understanding its evolution and impacts on the climate system.

In order to characterise the ocean carbon system with its strong regional and temporal variability, extensive spatial and temporal coverage in observations is required. While there are millions of surface ocean *p*CO₂ measurements to provide

a surface view of the ocean carbon system and air–sea carbon fluxes (Bakker et al., 2016), complementary observations of ocean interior dissolved inorganic carbon (DIC) are limited by the logistics of ship-based bottle measurements. Campaigns for ocean interior DIC observations use repeat transects to provide high-quality observations for many regions (Sloyan et al., 2019). The transects resolve spatial variations in carbon but are limited in their ability to resolve temporal variability as they are repeated on decadal timescales. Alternatively, ocean DIC time series such as those found at the Bermuda Atlantic Time Series and the Hawaii Ocean Time Series can resolve seasonal, inter-annual, and decadal variability but are limited in how well they represent variability on larger spatial scales (Bates et al., 2014). The Bio-Argo programme has allowed for more autonomous sampling of ocean interior carbonate system variables such as pH (Clastre et al., 2020), from which interior DIC can be estimated; currently, however, autonomous measurement methods for interior DIC remain in development.

By itself, the sparsity of interior observations hinders the ability to produce a coherent global picture of recent ocean carbon changes from a storage perspective. A mapping technique is necessary to expand points of observations into more coherent spatial patterns of change and behaviour. Mapping techniques employ statistics to propagate information from observations to unobserved regions. Time-invariant climatologies of ocean carbon have been created using mapping procedures on data from repeat transects and other ship-based observations (Lauvset et al., 2021). Non-linear machine-learning procedures have been able to reconstruct spatial patterns in inter-annual and decadal variability for surface $p\text{CO}_2$ (Landschützer et al., 2016, 2019; Gloege et al., 2022) as well as the seasonal cycle for DIC (Keppler et al., 2020), but the non-linearities can create substantial biases for regions with sparse observations (Bushinsky et al., 2019). Data-assimilative methods that use observations to constrain model physics in forward experiments have also been expanded to include biogeochemistry (Verdy and Mazloff, 2017; Carroll et al., 2020). Multiple linear-regression models have also been used to reconstruct changes in the anthropogenic carbon pool (Clement and Gruber, 2018; Gruber et al., 2019a).

For any mapping or data-assimilative technique, it is imperative to use accurate statistics to avoid the erroneous propagation of information from observations. These statistical relationships can be parameterised or calculated from observations; however, parameterised statistics fail to reflect regional differences in the ocean, and for poorly observed variables, averaging over large length scales or timescales can lead to overly coarse covariance fields. State-of-the-art climate models provide complete pseudo data in both space and time and therefore may be used to calculate these statistics, though model biases may lead to errors in the covariances. Nevertheless, fully coupled climate models have been used to reconstruct ocean heat content trends and variability from

observations, with clear improvements in recent years due to the expansion of the Argo programme (Smith and Murphy, 2007; Smith et al., 2015; Cheng and Zhu, 2016).

Ocean carbonate chemistry is controlled by both physical and biogeochemical processes. There are well-understood first-order principles that relate to ocean temperature and salinity to upper-ocean carbon; using these relationships, the observations used in previous heat content mappings may be exploited in a similar manner for carbon mappings. Increases in temperature reduce the uptake of CO_2 through gas solubility (Weiss, 1974). Increases in salinity allow for further CO_2 uptake by increasing alkalinity (Williams and Follows, 2011). Additionally, temperature and salinity provide constraints on ocean circulation, which alters the background vertical gradients of both heat and carbon (Thomas et al., 2018; Williams et al., 2021). Anthropogenic carbon uptake at the high latitudes is constrained by salinity and stratification, which can be taken to be proxies for water mass formation (Terhaar et al., 2020, 2021; Bourgeois et al., 2022). The relationships between temperature, salinity, and carbon are regionally dependent as ocean dynamics and biology can set different drivers of CO_2 uptake (Lauderdale et al., 2016). If these relationships can be exploited, the increase in ocean observations from the Argo programme may provide valuable information that can help reconstruct ocean carbon fields alongside temperature and salinity fields. While the ocean carbonate system can also be approximated using other observations such as pH and salinity, for this study we focus on the potential benefit of observations present within the Argo dataset.

In this study, we apply an ensemble optimal interpolation approach to reconstruct modelled upper-ocean dissolved inorganic carbon from synthetic temperature and salinity observations. Covariance fields are constructed using an ensemble of six Earth system models from the Climate Model Intercomparison Project Phase 6 (CMIP6). In this proof-of-concept study we show the potential skill available in using model covariance fields and Argo-style synthetic measurements to reconstruct carbon content between 0 and 2000 m. Synthetic reconstructions of modelled ocean carbon are created by using pseudo observations of temperature and salinity, similar to the ocean heat content synthetic reconstructions in Allison et al. (2019). The errors within these reconstructions can then be compared with those from the climatological fields to see where and how the method best works in the model world.

The work in this study is set out as follows. Section 2 introduces the ensemble optimal interpolation scheme for ocean DIC using an ensemble of six CMIP6 models and the experiments used to test the reconstruction skill. The method is first assessed by reconstructing DIC inventory changes over the top 100 m and then over the top 2000 m. Section 3.1 presents the ensemble correlation fields between DIC over the top 100 m and temperature and salinity at the same location, with a discussion as to how these correlations re-

late to physical controls on the ocean carbon response. In Sect. 3.3 and 3.4 the reconstruction potential of different temperature and salinity sampling distributions is assessed for the upper 100 m, ranging from perfect spatial coverage to coverage more representative of Argo observations. The method is then extended to consider the full 2000 m profiled by Argo floats in Sect. 4, and the potential to reconstruct recent global carbon changes is illustrated with a reconstruction using time-varying synthetic Argo observations in Sect. 5. Lastly, in Sect. 6 we discuss the potential and shortcomings of this method, both in terms of the set-up with temperature and salinity and how the method may be expanded with other oceanographic observations.

2 The ensemble optimal interpolation method for ocean carbon

Optimal interpolation is a non-dynamical mapping approach that uses weights to propagate information from observations to regions without observations (Daley, 1991; Smith and Murphy, 2007). The optimal interpolation method involves creating an analysis A at locations i and times t from the sum of a background state B_i and a weighted sum of the difference between the observed and background values, also known as the observation increments. The observations can be the same property as the final analysis or can be of other properties; i.e. temperature and salinity observations can be used to reconstruct DIC if there is a physical link between them. The inclusion of other physically relevant data produces a *multivariate analysis*. For observations O at locations k within a sampling region K , the optimal interpolation method can be represented as

$$A_i = B_i + \sum_{k \in K} w_k (O_k - B_k). \quad (1)$$

The crux of the optimal interpolation problem is thus finding a suitable solution for the weights w_k . The optimal weights are those that minimise the expected analysis error at each grid point i , calculated as the root-mean-squared error (RMSE) between the analysis and the truth T_i :

$$\text{RMSE}(A_i, T_i) = \sqrt{\frac{1}{N} \sum_t (A_i(t) - T_i(t))^2}. \quad (2)$$

In this work we explore how the relationships between DIC and observed ocean variables such as temperature and salinity can be used to reconstruct upper-ocean carbon inventories down to a depth of 2000 m. The use of optimal interpolation to reconstruct ocean carbon from available carbon measurements is limited by the poor temporal and spatial coverage of existing DIC observations. However, temperature and salinity observations are more plentiful, particularly in the Argo era, so we explore the extent to which these observations could be used to reconstruct DIC. To avoid the

problems associated with sparse input data, we take a multivariate analysis approach to reconstruct carbon from extensive synthetic observations of ocean temperature T and salinity S consistent with observations from the Argo programme (Wong et al., 2020) as well as annual-average atmospheric CO_2 concentrations. On annual and longer timescales, atmospheric CO_2 is well-mixed, and so the use of annual atmospheric CO_2 measurements can allow the analysis to capture longer-term DIC changes from changes in the global carbon budget. With background fields for DIC, T , and S taken to be their global climatologies (i.e. $\text{DIC}(i, t) = \overline{\text{DIC}}(i) + \text{DIC}'(i, t)$, where $\overline{\text{DIC}}(i)$ is the time-average DIC concentration at location i over the period 1955–2014), the optimal interpolation scheme is formulated to calculate the residual DIC' from T' , S' within a region K :

$$\text{DIC}'(i, t) = \sum_{k \in K} (w_{C,k} p\text{CO}_2'(t) + w_{T,k} T'(k, t) + w_{S,k} S'(k, t)), \quad (3)$$

where $w_{C,k}$ is the local weighting for the annual global-averaged atmospheric $p\text{CO}_2$ and $w_{T,k}$ and $w_{S,k}$ are the weightings for observed ocean temperatures and salinities within the sampling region K . These optimal weights describe how information is propagated from atmospheric $p\text{CO}_2$ and hydrographic observations to the ocean carbon system, taking into account the interdependencies between input variables.

2.1 Choices for calculating background error covariances

The optimal weights w_k in Eqs. (1) and (3) are determined by the covariances between the background errors $B_i - T_i$. The covariance fields describe how information should be propagated from areas with observations to those without; the optimal solution will consider how much new information observations provide to the reconstruction and will depend on local and larger-scale relationships within the climate system. Any ensemble optimal interpolation method requires decisions on how the background error covariances are calculated and which covariances are included to solve for the weights w_k . In the following we describe how we have made these decisions using a CMIP6 multi-model ensemble and with various assumptions as to which observations are used for the multivariate DIC analyses.

2.1.1 Background covariances from a CMIP6 ensemble

Climate model outputs can be used to provide covariance fields for optimal interpolation solutions. By providing complete pseudo data in both space and time, climate models avoid some of the errors that arise from the coarseness of observational or parameterised covariance fields. Background covariance fields from global climate models have been used to reconstruct ocean temperatures and salinities from obser-

vations and to initialise decadal forecasts (Smith and Murphy, 2007; Smith et al., 2015; Cheng and Zhu, 2016). However, the model ensemble background covariance field will still contain errors that will need to be evaluated using sensitivity testing before the fields can be used with real-world observations.

To construct the background covariance fields, we obtained ocean potential temperature, ocean salinity, and ocean DIC model output from six CMIP6 Earth system models with a nominal horizontal resolution around 1° (Table 1). The output was taken from the historical experiments and covers the period of year 1955 to year 2014. This period was chosen as it has consistent behaviour in atmospheric CO_2 concentrations and is long enough to allow some multi-decadal variability to be captured in the covariance fields. For each model, five realisations were used to capture the models' internal variability.

For all the variables, annual averages were calculated from monthly mean outputs, and outputs were bi-linearly regridded from their native horizontal grids to a $1^\circ \times 1^\circ$ grid using the Python package `xESMF` (<https://doi.org/10.5281/zenodo.1134365>). The drifts in temperature, salinity, and DIC were calculated and removed by subtracting linear trends at each ocean grid cell in the piControl runs. Oceanic variables were further integrated vertically in layers that span the ocean surface to 2000 m depth. The integrated layers were chosen to be 0–100, 100–500, 500–1000, and 1000–2000 m.

2.1.2 Spatial limits on background covariances

The sampling region K for temperature and salinity observations in Eq. (3) can vary from co-located observations to global observations; the choice of sampling region requires balancing of the extra information provided by additional observations with spurious propagation of information through errors in the covariance fields. In this study we use two idealised sampling methods to explore the limits of this first-order reconstruction of ocean carbon.

1. The first reconstruction method assumes full global coverage of temperature and salinity observations. For these synthetic reconstructions, the ocean inputs to reconstruct ocean DIC are co-located model temperature and salinity anomalies as well as globally uniform atmospheric $p\text{CO}_2$ anomalies. The resulting system has three input parameters to solve for ocean carbon at each grid cell. With this method, we explore how the local relationships between upper-ocean DIC, temperature, and salinity can be used to reconstruct carbon with perfect hydrographic knowledge. Since observations are complete in this test, the presence of errors can be attributed to poorly known relationships or nonlinear dynamics within the model ensemble.
2. Synthetic observations are made irregularly, consistent with the distribution of Argo profiles (Wong et al.,

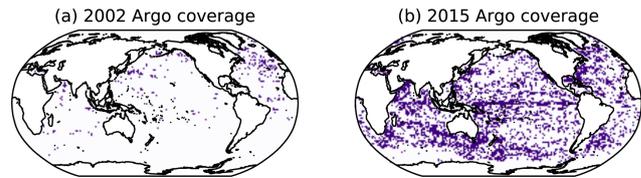


Figure 1. Argo profile distributions consistent with sampling for (a) year 2002 and (b) year 2015. Coloured regions indicate bins that have at least 6 months of observations in the given year.

2020). While coverage of ocean temperature and salinity observations is significantly higher than those for ocean carbon, sampling remains irregular in space and time, so real-world reconstruction methods must account for this irregularity. We take Argo data from years 2002–2015 and bin the profiles onto the $1^\circ \times 1^\circ$ horizontal grid. Any grid cell that has Argo observations for at least 6 months within a given year (running from January to December) is taken to be sufficiently observed, and the modelled annual-average profile there is used as an observation for the reconstruction. For year 2002, this leads to a distribution of synthetic profiles that is heavily concentrated in the Northern Hemisphere and the North Atlantic in particular (Fig. 1a). By year 2015, the scheme allows for sampling of most of the global ocean outside the Southern Ocean and the Arctic Ocean (Fig. 1b). Using this distribution of available profiles, we create reconstructions using time-varying profile locations and test how profile density impacts the reconstruction by setting the distribution as constant in time.

As the Argo profiles have incomplete global coverage, we conduct additional sensitivity tests to see how nearby Argo observations can be used to construct near-global reconstructions. For these reconstructions, the reconstruction at any given point is made to be a linear combination of observations within a certain radius. Radii of 1, 2, or 5° are chosen: for a radius of 5° and current Argo coverage like that for year 2015, most regions of the ocean have observations within the search radius, allowing for nearly globally complete DIC reconstructions.

2.2 Synthetic reconstructions and tests

Evaluating how well an analysis reproduces variability requires a comparison with a field truth. As the ocean carbon field is not known to a high accuracy, synthetic reconstructions of models' ocean carbon fields are created using various distributions of modelled temperature and salinity (Allison et al., 2019). For these tests, the synthetic observations come from the models and therefore contain no errors outside of small errors possible from sampling and regridding. As the synthetic observations can be assumed to have negligible error, the weights w_k in (2) can be solved for through a

Table 1. CMIP6 Earth system models and realisations used for reconstruction.

Model (reference)	Realisations
ACCESS-ESM-1.5 (Ziehn et al., 2020)	r1i1p1f1, r2i1p1f1, r4i1p1f1, r5i1p1f1, r6i1p1f1
CanESM5 (Swart et al., 2019)	r10i1p1f1, r11i1p1f1, r12i1p1f1, r13i1p1f1, r14i1p1f1
CESM2 (Danabasoglu et al., 2020)	r1i1p1f1, r2i1p1f1, r3i1p1f1, r4i1p1f1, r5i1p1f1
IPSL-CM6A-LR (Boucher et al., 2020)	r1i1p1f1, r2i1p1f1, r3i1p1f1, r4i1p1f1, r32i1p1f1
MPI-ESM1.2-LR (Mauritsen et al., 2019)	r1i1p1f1, r2i1p1f1, r3i1p1f1, r4i1p1f1, r5i1p1f1
UKESM1 (Sellar et al., 2019)	r1i1p1f2, r2i1p1f2, r3i1p1f2, r4i1p1f2, r8i1p1f2

least-squares algorithm without having to consider observational errors (Smith and Murphy, 2007).

The models used in the synthetic reconstructions are the same models used to calculate the ensemble covariance fields. The inclusion of a model in the ensemble can over-fit the results and produce spuriously accurate reconstructions; therefore, for a synthetic reconstruction of, for instance, UK Earth System Model (UKESM1) DIC inventories, we eliminate all the UKESM1 realisations from the ensemble. The covariances and optimal weights are then calculated from the remaining ensemble members. Then, these optimal weights are used with the UKESM1 temperature and salinity profiles to create a DIC analysis.

To compare the reconstructions, the improvement in the RMSE is calculated relative to the RMSE of the first-guess background using

$$\varepsilon_i(A) = \frac{\text{RMSE}(B_i, T_i) - \text{RMSE}(A_i, T_i)}{\text{RMSE}(B_i, T_i)}. \quad (4)$$

Here the conventions follow those in Eq. (1), where $\text{RMSE}(B_i, T_i)$ is the RMSE between the modelled truth T and the background climatology field B (equivalent to the standard deviation of T), and $\text{RMSE}(A_i, T_i)$ is the RMSE between the modelled truth and the multivariate analysis A . Each RMSE is calculated for the same historical period used in the creation of the model ensemble; i.e. t is taken from output within the modelled years 1955 to 2014. The maximum value of 1 indicates a perfect reconstruction A , whereas values below 0 indicate that the errors are larger for the analysis than they would be if the solution were to be taken as the climatological first guess.

For the following analysis, synthetic reconstructions of DIC are first made over the upper 100 m and then over the upper 2000 m, using both global and irregular temperature and salinity observations. Over the top 100 m, the ocean carbon distribution is expected to reflect the controls of solubility and alkalinity on ocean carbon due to the dominance of air–sea gas exchange on the carbon inventory. Over the upper 2000 m, ocean circulation and ventilation are expected to play a larger role in determining the reconstructed carbon fields.

3 Reconstruction of ocean DIC within the top 100 m

Our reconstruction approach using the ensemble of six CMIP6 model runs is now applied to carbon inventory changes in the top 100 m. We begin by exploring the covariance fields between the system input and output variables through correlations. As the covariance fields remain the same for both synthetic and real-world analyses, it is important to test whether these fields reflect the physical relationships between DIC, $p\text{CO}_2$, T , and S .

3.1 Model correlation fields between $p\text{CO}_2$, temperature, salinity, and DIC

The least-squares solution for the weights for temperature, salinity, and atmospheric CO_2 is a function of the covariances between the input variables and DIC. Thus, the structure of the covariance fields provides insight into how temperature and salinity can be used to reconstruct carbon. For simplicity, we illustrate these relationships through correlation fields taken for the entire model ensemble, which normalise the relationships using the variances of each input variable.

3.1.1 Correlations with atmospheric $p\text{CO}_2$

The correlation fields between atmospheric $p\text{CO}_2$ and DIC, temperature, or salinity reflect the impact of emissions on the ocean mixed layer. Atmospheric $p\text{CO}_2$ is strongly positively correlated with mixed-layer DIC and reaches values near 1 in the mid-latitude ocean (Fig. 2a). This strong correlation reflects how the ocean takes up carbon under higher atmospheric $p\text{CO}_2$ due to gas disequilibrium. The correlation between atmospheric $p\text{CO}_2$ and integrated temperature changes is widely positive, consistent with the long-term ocean uptake of both heat and carbon during carbon emissions (Fig. 2b). Lastly, the correlation between atmospheric $p\text{CO}_2$ and integrated salinity changes is close to zero in most regions, although smaller regions such as the Gulf Stream show stronger correlations.

3.1.2 Correlations between DIC and temperature

Whereas atmospheric $p\text{CO}_2$ has little inter-annual variability, mixed-layer temperature, salinity, and DIC will have variability on inter-annual and decadal timescales alongside

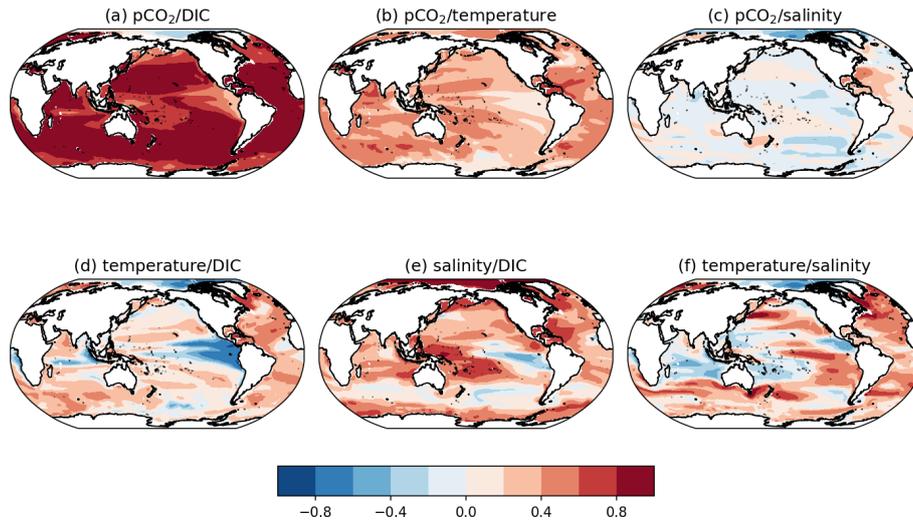


Figure 2. Correlations for the entire CMIP6 historical ensemble between changes in (a) atmospheric $p\text{CO}_2$ and upper-ocean (0–100 m integrated) DIC, (b) atmospheric $p\text{CO}_2$ and upper-ocean temperature, (c) atmospheric $p\text{CO}_2$ and upper-ocean salinity, (d) upper-ocean temperature and DIC, (e) upper-ocean salinity and DIC, and (f) upper-ocean temperature and salinity.

long-term trends. The correlation fields between mixed-layer DIC, temperature, and salinity reflect the combined response to external forcing from atmospheric $p\text{CO}_2$ and internal variability.

Mixed-layer temperature and DIC are positively correlated in most of the ocean, with strong negative correlations in the eastern equatorial Pacific and Indian oceans (Fig. 2d). This positive correlation contradicts the negative correlation expected from the solubility control of temperature on DIC.

To separate the correlation into terms relating to the external forcing and ocean variability, we can decompose temperature and DIC into terms proportional to atmospheric $p\text{CO}_2$ and an anomaly term:

$$T'(x, y, t) = \alpha(x, y) p\text{CO}_2(t) + T_a(x, y, t), \quad (5)$$

$$\text{DIC}'(x, y, t) = \gamma(x, y) p\text{CO}_2(t) + \text{DIC}_a(x, y, t). \quad (6)$$

The terms α, γ describe the respective spatial patterns of changes in T', DIC' with changes in $p\text{CO}_2$. The anomalies of T', DIC' after a regression against $p\text{CO}_2$ is removed are denoted as T_a and DIC_a .

With the decomposition in Eqs. (5) and (6), the correlation between DIC' and T' can be approximated as the sum of a $p\text{CO}_2$ component and a non- $p\text{CO}_2$ component, where $\sigma_{\text{DIC}}, \sigma_T$ are the standard deviations of DIC' and T' :

$$\rho(\text{DIC}', T') \approx \underbrace{\alpha\gamma \frac{\text{var}(p\text{CO}_2)}{\sigma_{\text{DIC}}\sigma_T}}_{p\text{CO}_2 \text{ term}} + \underbrace{\frac{\text{cov}(\text{DIC}_a, T_a)}{\sigma_{\text{DIC}}\sigma_T}}_{\text{non-}p\text{CO}_2 \text{ term}}. \quad (7)$$

This decomposition is accurate to first order. A more detailed derivation of this breakdown and a discussion of the approximation can be found in Appendix A.

The $p\text{CO}_2$ component of the correlation field between temperature and DIC is positive almost everywhere (Fig. 3a).

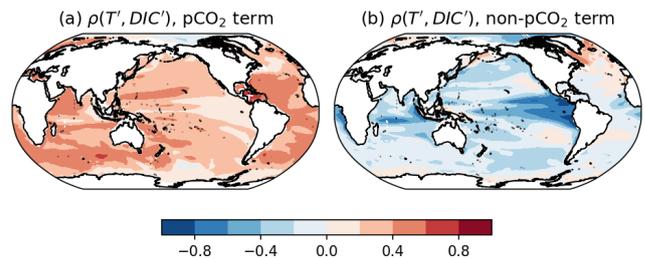


Figure 3. Breakdown of the correlation between upper-ocean DIC and temperature, $\rho(\text{DIC}', T')$, from Fig. 2d: (a) a term proportional to the variance of atmospheric $p\text{CO}_2$ and (b) a term consisting of the covariance of the residuals calculated after removing a linear fit against atmospheric $p\text{CO}_2$. For a discussion of the breakdown of the correlation into terms and of the approximation, see Appendix A.

This relationship is consistent with the combined oceanic uptake of heat and carbon from CO_2 emissions. Conversely, the anomaly term for the correlation field between temperature and DIC is broadly negative, which corresponds to the solubility control for DIC (Fig. 3b). Thus, the heterogeneity found in the overall correlation between DIC and temperature in Fig. 2 can be understood as the sum of an emissions-driven undersaturated response that correlates with but is not driven by warming and a temperature-driven solubility response.

3.1.3 Correlations between salinity and DIC or temperature

Unlike DIC, temperature, and atmospheric $p\text{CO}_2$, mixed-layer salinity changes have weak trends. Decomposing the

correlations between salinity and temperature or DIC shows little role for the $p\text{CO}_2$ term in Eq. (7) (Appendix A). Therefore, we continue by exploring the full correlation fields with mixed-layer salinity.

The correlations between upper-ocean salinity and upper-ocean DIC are positive in most regions, reflecting the alkalinity control on carbon solubility (Fig. 2e). Correlations are weak and slightly negative in the Southern Ocean and eastern equatorial Pacific, respectively. These correlations may be imprints of dynamical changes: for the Southern Ocean, the salinity in the frontal zone impacts the strength of mode and intermediate water formation (Terhaar et al., 2022), while for the equatorial Pacific enhanced wind-driven upwelling transports fresher, carbon-rich waters to the surface (Williams et al., 2021). Lastly, correlations between upper-ocean temperature and salinity are moderate and show strong regional variability (Fig. 2f).

Overall, the correlation fields show first-principle controls relating to the increase in ocean heat and carbon under emissions as well as solubility and alkalinity controls of temperature and salinity anomalies on ocean DIC.

3.2 Ensemble optimal weights for $p\text{CO}_2$, temperature, and salinity

Translating covariance and correlation fields to optimal mapping parameters involves both the relationships between the input variables and the output variables as well as the relationships between input variables. Therefore, optimal solutions are non-trivial; for instance, regions with similar correlations between DIC and temperature and salinity (such as the equatorial Pacific) may have coefficients with different signs depending on how the system fits the linear model to the data. We continue by comparing the least-square-coefficient fit to atmospheric $p\text{CO}_2$, integrated ocean temperature, and integrated ocean salinity to the correlations in Sect. 3.1 and the first-principle controls on carbon storage.

Increases in atmospheric $p\text{CO}_2$ are translated to increases in upper-ocean carbon (Fig. 4a). The weight magnitudes reach local maxima in the subtropics, in accordance with regional variability in the Revelle buffer factor, which describes the ratio between increases in DIC and increases in atmospheric $p\text{CO}_2$ (Williams and Follows, 2011) and thermocline ventilation. Temperature coefficients are negative almost everywhere (Fig. 4b), while salinity coefficients are positive everywhere (Fig. 4c). While the correlations between upper-ocean DIC, temperature, and salinity show regional variability, the ensemble ultimately reveals a consistent forced control of atmospheric CO_2 , solubility control by temperature, and alkalinity control by salinity.

Thus, the least-squares solution for DIC as a combination of atmospheric $p\text{CO}_2$, temperature, and salinity changes is able to capture the controls of the Revelle buffer factor, solubility, and alkalinity on upper-ocean carbon. The information from ocean temperature and salinity observations is comple-

mentary: regions with relatively high-/low-temperature coefficients correspond to low-/high-salinity coefficients. The optimal weight solutions indicate that the information provided by these observations is consistent with our hypothesised first-order controls.

3.3 Reconstruction potential using co-located observations

Sensitivity tests are now conducted to estimate the ability to reconstruct ocean carbon and to explore how the ensemble composition impacts the reconstruction. In total, six reconstructions were created, in which one model was removed from the ensemble, the covariance fields between DIC' , T' , S' , and atmospheric $p\text{CO}_2$ are recalculated using the remaining models, and then DIC' from the excluded model is reconstructed using the new covariance fields and its own T' and S' observations (Sect. 2). The coefficients for atmospheric CO_2 , temperature, and salinity are similar across the sensitivity set-ups with each model removed (Figs. S1–S3 in the Supplement), so we continue by comparing the RMSE improvements ε from Eq. (4) for each sensitivity test.

The ensemble minimum, average, and maximum relative RMSE improvements are calculated for six reconstructions, each created by eliminating a model from the CMIP6 ensemble. The minimum and maximum error improvements $\min(\varepsilon)$ and $\max(\varepsilon)$ reflect the makeup of the ensemble. A negative minimum error improvement at any point indicates where one of the models in the ensemble has covariances that are substantially different than the others; therefore, including the model within the ensemble adds uncertainty to the reconstruction and pushes the reconstruction towards the first-guess climatology field. A high maximum error improvement indicates that the solution weights for the sensitivity results are similar. This similarity can arise from strong physical constraints on the upper-ocean carbon system but may also arise because the ensemble members are spuriously similar in their architecture or representations of climate processes.

The RMSE improvement of the reconstruction relative to the climatological first guess is now considered for both the full upper-ocean carbon response and the detrended carbon response (Fig. 5). Over multiple decades, we expect the carbon inventory response to be dominated by an upwards trend due to continuing carbon emissions; thus, the improvement in the detrended response provides insight into how inter-annual and decadal variability in ocean carbon storage is reproduced. Areas where the RMSE increases relative to the climatological first guess are noted in red. For this analysis we focus on the open-ocean RMSE reductions, as the CMIP6 models have different coastlines after being regridded, leading to small areas of RMSE increases near the land–sea boundary.

Outside of some coastal and Arctic regions, each sensitivity test reduces the upper-ocean RMSE for upper-ocean carbon (Fig. 5a). This improvement is not due solely to the

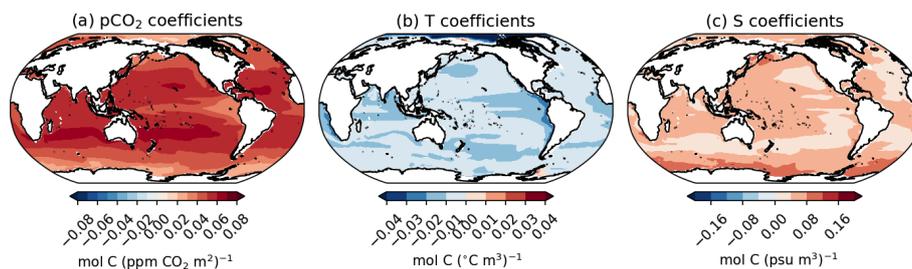


Figure 4. Ensemble coefficients for reconstructing upper-ocean (0–100 m integrated) DIC with co-located observations. (a) Coefficients for annual-average atmospheric CO_2 concentrations ($\text{mol C (ppm CO}_2 \text{ m}^2)^{-1}$), (b) coefficients for upper-ocean integrated temperature ($\text{mol C (}^\circ\text{C m}^3)^{-1}$), and (c) coefficients for upper-ocean integrated salinity ($\text{mol C (psu m}^3)^{-1}$).

reproduction of the long-term increase in ocean carbon, as most regions show improvements in detrended ocean carbon as well; exceptions to this improvement can be found at the high latitudes and in the subtropical regions (Fig. 5d). The average reconstruction reduces the RMSE by between 60 % and 90 % on average relative to the climatology first-guess RMSE (Fig. 5b) and reduces the detrended RMSE by 30 % to 80 % (Fig. 5e). These relative improvements are equivalent to the method capturing over 60 % of the detrended variability in most regions of the ocean. The high (close to 1) relative RMSE reductions found in the ensemble maximum statistics (Fig. 5c and e) suggest that there are models which are highly similar to one another.

For reconstructions of global DIC inventories using co-located temperature and salinity observations, the sensitivity experiments show an average RMSE reduction of 93 %. When considering detrended DIC inventories, the sensitivity experiments reduce the RMSE by 68 % on average.

There are noticeable regional variations throughout all of the improvement statistics. The western low-latitude Pacific Ocean and subtropical Indian Ocean show consistent local maxima in all of the improvement statistics. These regions of maximum relative RMSE reduction are characterised by their strong correlation between salinity and DIC, suggesting that the most constrained responses within the CMIP6 ensemble may be related to the control of alkalinity on DIC. Conversely, the small regions that show potential degradation in the reconstructions due to the errors in the covariances are characterised by weak correlations against atmospheric $p\text{CO}_2$ changes (for the full carbon signal reconstruction, Fig. 2a–c) and weakly positive covariances between T_a and DIC_a (for the detrended carbon system, Fig. 3b).

Thus, when considering point-wise observations of temperature and salinity alongside global-average CO_2 concentrations, our sensitivity experiments indicate that a substantial amount of upper-ocean carbon variability can be reconstructed. These estimates can provide an upper bound on the reconstruction potential as global coverage of temperature and salinity observations is theoretically ideal but difficult to accomplish, even with widespread autonomous observing tools. We therefore continue by exploring how irregular ob-

servations of temperature and salinity impact reconstructions of upper-ocean carbon.

3.4 Reconstructing carbon using irregular Argo observations

The point-wise reconstructions of upper-ocean carbon from atmospheric $p\text{CO}_2$, temperature, and salinity reveal that the physical controls of solubility and alkalinity can explain a substantial amount of inter-annual variability in the upper-ocean carbon system. Although temperature and salinity profiles have a larger global coverage than ocean carbon profiles, coverage remains irregular and incomplete. In this section we explore how the irregular coverage in Argo profiles impacts carbon reconstructions by conducting further synthetic reconstructions with observations consistent with year 2015 Argo coverage (Fig. 1b).

When considering irregular sampling, reconstructions with global coverage use background covariances to propagate information from Argo profile locations to the rest of the ocean. There is the potential for information from multiple Argo sites being used to reconstruct ocean carbon at a given location. If the covariance fields are correct, additional information should improve the reconstruction and increase the relative error reduction; however, any errors in the covariance fields can propagate and decrease the relative error reduction. The radius of influence for an observation should ideally balance the extra information and errors it provides in the reconstruction.

With the smaller radii of 1 and 2°, the method creates an incomplete reconstruction even under current-day Argo profile distributions. With a larger radius of 5°, most of the global ocean has nearby temperature and salinity profiles that can be used for the reconstruction (see the Fig. 6 coverage). Overall, the benefit of increased coverage outweighs the increases in reconstruction errors (Fig. S4). Thus, in this section we focus on the 5° radius results for simplicity.

With the 5° cut-off radius for observations, the carbon reconstruction at each point will use between 0 and 81 (= 1 local observation + 80 observations within the cut-off radius) temperature and salinity inputs. Including the globally aver-

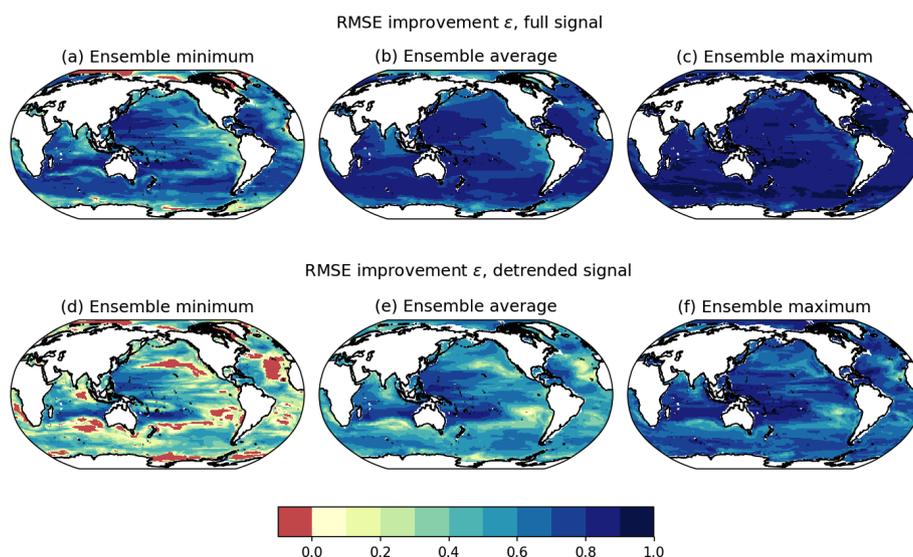


Figure 5. Ensemble statistics for the relative RMSE reductions ε using co-located temperature and salinity over the upper 100 m: (a) ensemble minimum, (b) ensemble average, and (c) ensemble maximum. Ensemble statistics for the relative RMSE reductions but considering only the detrended carbon signal: (d) ensemble minimum, (e) ensemble average, and (f) ensemble maximum. Red areas indicate regions where the sensitivity tests show a RMSE increase relative to the assumption of climatology.

aged atmospheric CO_2 inputs, the DIC at each point can be reconstructed from at least 1 and up to 82 input variables (up to 81 temperature and salinity profiles + globally averaged atmospheric $p\text{CO}_2$).

Errors in spatial covariances result in some regions with lower RMSE improvements compared to reconstructions using co-located observations (Fig. 6a, d). Across the sensitivity tests and outside of the Southern Ocean and small regions in the North Atlantic, North Pacific, and equatorial Pacific, carbon can be reconstructed with a relative RMSE reduction of at least 50%. When considering the detrended carbon signal, most regions show an average RMSE reduction of at least 40%. Notably, the ensemble minimum RMSE improvement for the detrended signal shows substantial regions where the sensitivity test that excludes MPI-ESM1.2-LR produces a poor reconstruction. MPI-ESM1.2-LR may be viewed as an outlier model that provides additional uncertainties not covered by the other models.

For most of the upper ocean the RMSE improvements are similar to those found in the reconstruction with co-located temperature and salinity measurements, and the regional variation in the RMSE improvements is similar regardless of whether co-located or removed observations are used. The additional information provided by nearby measurements counteracts the error propagation inherent in the scheme, suggesting that the use of irregular Argo profiles provides sufficient information to reproduce carbon variability in the mixed layer.

4 Reconstruction of ocean DIC within the top 2000 m

Our reconstruction approach is now applied to carbon inventories within the upper 2000 m. Our assessment of the CMIP6 models indicates that the relationships between carbon, temperature, and salinity may be used to reconstruct ocean carbon variability within the near-surface ocean. Anthropogenic carbon uptake and decadal variability are concentrated in the top 1000 m (DeVries et al., 2017; Gruber et al., 2019a). As Argo profiles can provide temperature and salinity information down to 2000 m, the synthetic reconstructions are extended to cover this vertical extent. For these interior reconstructions, the carbon inventories are reconstructed as a set of layers: 100–500, 500–1000, and 1000–2000 m. For the interior layers, we conduct the same analysis as was used for the top 100 m reconstructions by first examining the uncertainties when reconstructing carbon using co-located temperature and salinity measurements and then extending them to reconstructions using irregular observations consistent with Argo profile locations.

4.1 Sensitivity of optimal weights to depth level

The optimal coefficients for carbon as a function of co-located temperature and salinity, along with atmospheric $p\text{CO}_2$ concentrations, display distinct depth sensitivity (Fig. 7). The optimal coefficients for atmospheric $p\text{CO}_2$ remain positive for most region–depth combinations. For the 100–500 m layer, $p\text{CO}_2$ coefficients reach their local maxima in the subtropical gyres, similarly to the optimal solution found for the upper 100 m layer (Fig. 4a). Below 500 m, the optimal weights reflect regions of strong ventilation in the

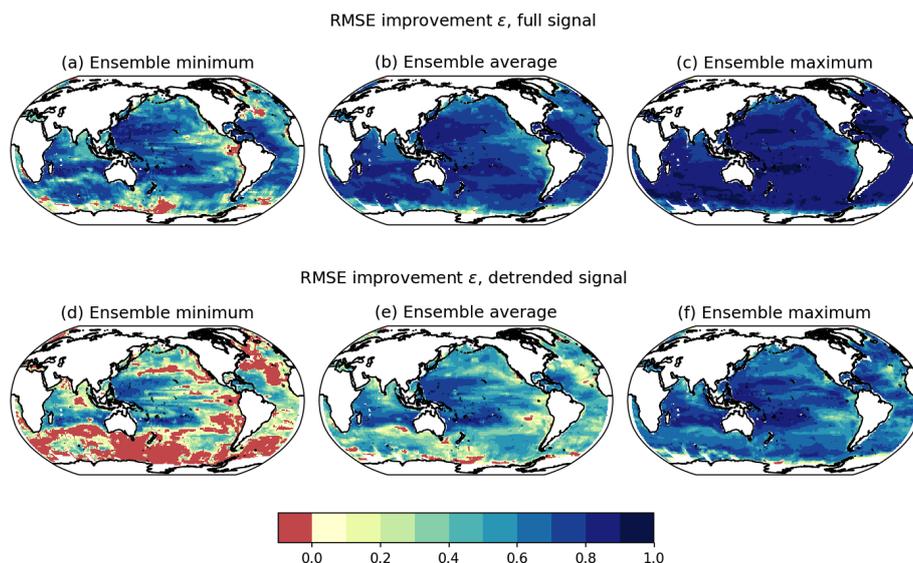


Figure 6. As for Fig. 5 but with RMSE improvements over the upper 100 m for the scheme that uses Argo 2015-type temperature and salinity observations within 5° of the reconstruction location.

North Atlantic and Southern oceans; below 1000 m, the coefficients outside these regions are near zero.

The optimal coefficients for interior temperature changes remain negative for most regions above 1000 m, with a depth-dependent structure (Fig. 7d–e). For the 1000–2000 m layer, the optimal temperature coefficients exhibit a zonal asymmetry in the Pacific basin, where the coefficients are negative in the western Pacific and positive in the eastern Pacific. Salinity coefficients operate similarly to temperature coefficients: for the 100–500 and 500–1000 m solutions, salinity coefficients are largely positive (Fig. 7g–h). Below 1000 m, salinity coefficients exhibit the same zonal structure seen in the temperature coefficients. The optimal coefficients at this depth are spatially correlated, which suggests that the information provided by temperature and salinities at this depth is less complementary than at shallower levels.

4.2 Carbon inventory reconstructions over 0–2000 m using co-located observations

The skill of the ensemble optimal interpolation reconstructions is now assessed for carbon inventory changes in the upper 2000 m, following the prior assessment over the upper 100 m.

For both the full carbon signal and the detrended carbon signal, the reconstruction skill decreases with depth. For carbon within the top 100 m of the water column, the average RMSE reduction is at least 60 % in most regions, and for globally integrated DIC reconstructions the average RMSE reduction is over 90 % (Fig. 5b). This RMSE reduction for the full DIC signal drops to 30 %–80 % for the 100–500 m layer (Fig. 8b).

Below 500 m, the sensitivity tests show large regions where the reconstruction increases the RMSE relative to a climatological first guess: for 500–1000 m the eastern North Pacific and Southern Ocean show increased errors, whereas below 1000 m most of the Pacific and Indian basins show increased errors (Fig. 8c, d). The regions in which the reconstruction maintains skill with depth are the well-ventilated regions that show maximum optimal coefficients for $p\text{CO}_2$ (Fig. 7c).

As with the reconstructions for the top 100 m, reconstructions at depth also show a lower RMSE reduction for the detrended DIC signal (Fig. 7f–h) and 0 %–60 %. For carbon between 100 and 500 m, the reconstruction reduces the detrended RMSE by between 0 % and 60 % in most regions. Below 500 m, the regions that exhibit increased errors in the detrended DIC signal are similar to those with larger errors in the full DIC signal.

While the sensitivity tests show poor reconstruction skill for large areas below 1000 m, carbon changes below this depth horizon are generally small, and large negative relative errors are found in regions that have little variability (see Fig. S5). Thus, the skill found in the 0–100 and 100–500 m layers mitigates the impact of covariance errors in these deeper layers when considering full column reconstructions.

5 Potential of the method to provide global and regional carbon time series

To illustrate the potential for reconstructing upper-ocean carbon using temperature and salinity measurements, we reconstruct output from the Norwegian Earth System Model

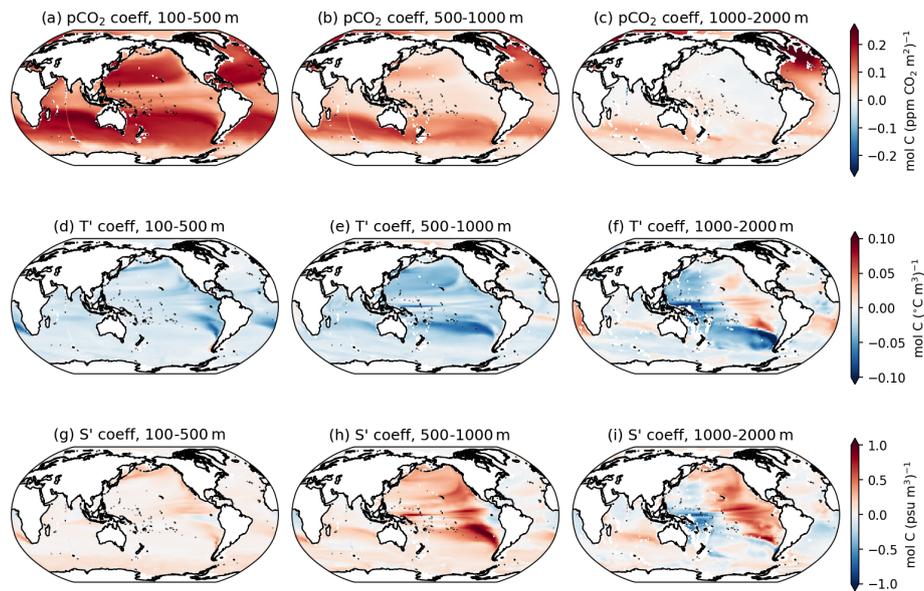


Figure 7. Optimal coefficients for interior DIC as a function of $p\text{CO}_2$ ($\text{mol C m}^{-2} \text{ppm}^{-1}$, **a–c**), temperature ($\text{mol C m}^{-2} \text{°C}^{-1}$, **d–f**), and salinity ($\text{mol C m}^{-2} \text{psu}^{-1}$, **g–i**). Columns show coefficients for interior ocean layers: 100–500 m (**a, d, g**), 500–1000 m (**b, e, h**), and 1000–2000 m (**c, f, i**).

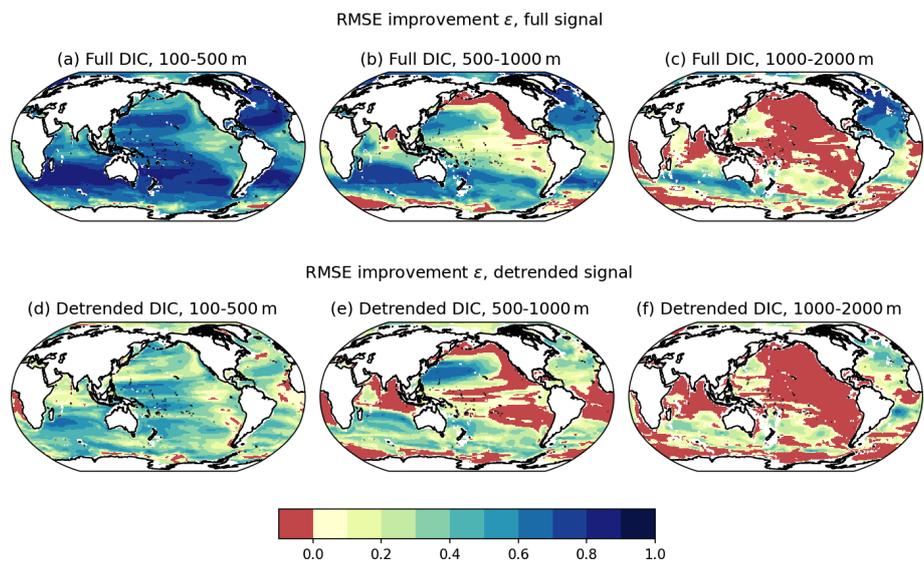


Figure 8. Average RMSE reduction ε , relative to a climatology, for sensitivity tests with out-of-sample models: panels (**a–c**) use average ε for the full DIC' signal and panels (**d–f**) use average ε for the detrended signal. The average RMSE reduction is calculated for individual depth levels: 100–500 m (**a, d**), 500–1000 m (**b, e**), and 1000–2000 m (**c, f**).

(NorESM2, Seland et al., 2020) using the full ensemble covariance fields and temperature and salinity measurements at locations similar to Argo observation locations. This reconstruction is an additional out-of-sample reconstruction, similar to those made for the error-reduction statistics in Sect. 3.3 and 3.4, but uses covariance fields constructed from the entire CMIP6 model ensemble. We elect to use NorESM2 as an independent check as it has three available realisations and

uses isopycnal depth coordinates, whereas all the models in the CMIP6 ensemble are z -level models.

Within this reconstruction, the cut-off radii for our localisation are depth-specific. The cut-off radii decrease with depth: 5° for 0–100 m, 2° for 100–500 m, 1° for 500–1000, and only using co-located observations for 1000–2000 m. These cut-off radii were chosen to limit the influence of errors in the covariance fields on the final reconstruction.

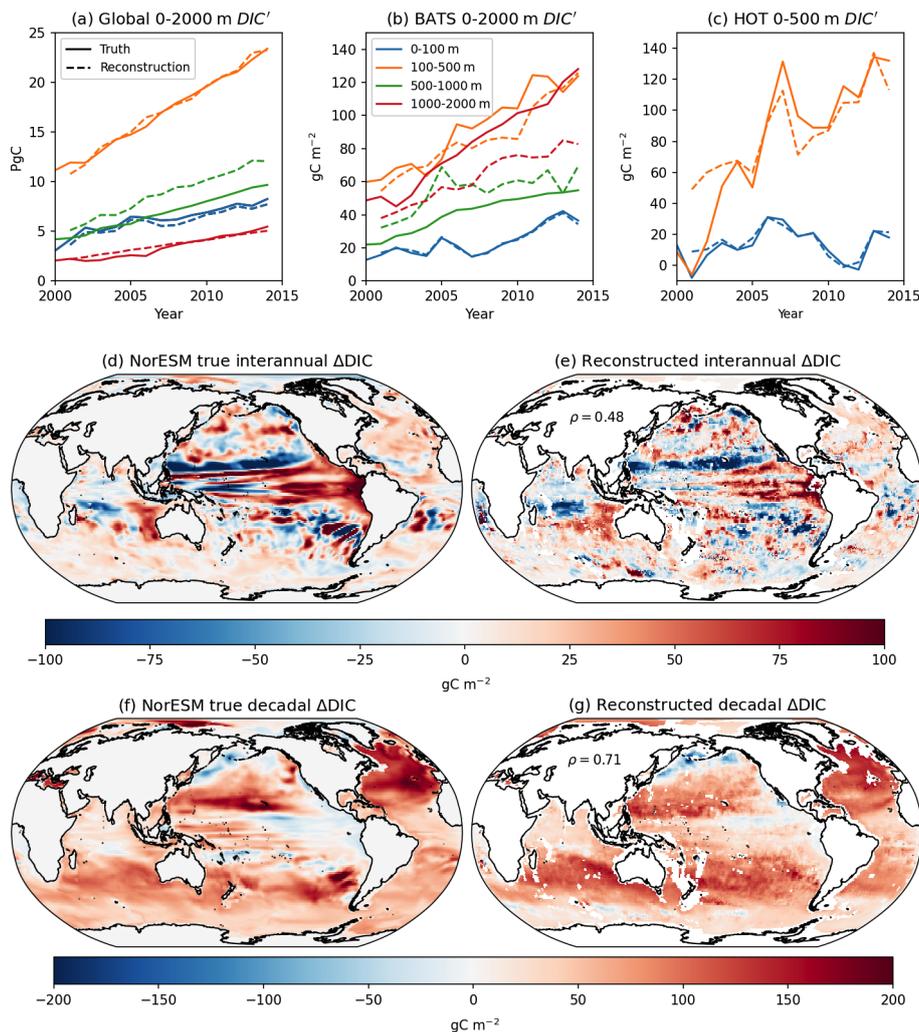


Figure 9. Reconstructions of NorESM 0–2000 m carbon inventories. **(a)** Globally integrated NorESM carbon inventories (solid lines) and reconstructed carbon inventories using temporally varying Argo coverage (dashed lines; Pg C). Colours indicate depths: 0–100 m (blue), 100–500 m (orange), 500–1000 m (green), and 1000–2000 m (red). **(b)** As for panel **(a)** but with vertically integrated DIC concentrations at BATS (gC m^{-2}). **(c)** As for panel **(b)** but with vertically integrated DIC concentrations at HOT to 500 m (gC m^{-2}). Snapshots of inter-annual changes in 0–2000 m DIC (gC m^{-2}): **(d)** NorESM truth and **(e)** reconstruction. Inter-annual changes are taken as the difference between year 2013 and year 2014 carbon fields. Snapshot of near-decadal changes in 0–2000 m DIC between periods 2001–2006 and 2009–2014 (gC m^{-2}): **(f)** NorESM truth and **(g)** reconstruction.

The global upper-ocean carbon inventory is dominated by a positive, near-linear trend across all depth levels (solid lines, Fig. 9a). Most carbon changes occur within the top 500 m (blue and orange lines); at these levels the reconstruction captures the long-term global behaviour. Below 500 m, the errors in the covariance fields result in biases in the reconstruction. Between 500 and 1000 m, the reconstruction overestimates DIC accumulation, whereas below 1000 m the reconstruction underestimates DIC accumulation. We note that, particularly below 1000 m, the enhanced coverage of the Argo programme over time results in a less accurate reconstruction due to the errors in the ensemble covariance fields.

To show how the reconstructions operate on more regional scales, we provide a reconstructed modelled carbon inventory at the Bermuda Atlantic Time Series (BATS, $31^{\circ}50' \text{ N } 64^{\circ}10' \text{ W}$) and the Hawaii Ocean Time Series (HOT, $22^{\circ}45' \text{ N } 158^{\circ} \text{ W}$) (Fig. 9b, c). The multi-decadal time series at BATS will be a useful validation tool for future reconstructions. With the localisation procedure, BATS has observations for a reconstruction for the top 2000 m; at HOT a lack of observations restricts the analysis to the upper 500 m.

At BATS, there is a long-term positive trend in DIC in all the layers (Fig. 9b). Reconstructed DIC captures the trend for carbon within the 0–100 and 100–500 m layers but underestimates DIC in the 1000–2000 m layer. DIC also exhibits inter-

annual variability within the top 500 m. While both the trend and variability are captured well in the top 100 m, between 100 and 500 m, the inter-annual variability is dampened.

At HOT, DIC in the upper 100 m is characterised by inter-annual variability, which the reconstruction captures. Between 100 and 500 m, DIC has a strong positive trend as well as inter-annual variability. The reconstruction captures both the trend and most of the inter-annual variability at this depth level.

Lastly, for a global view of the reconstruction, we provide maps of the NorESM DIC changes and their reconstruction analogues. Reconstructed inter-annual carbon inventory changes, shown here to be the top 2000 m DIC changes between year 2013 and year 2014, broadly match those in the model (Fig. 9c–d). The reconstruction captures the increase in carbon in the equatorial Pacific, the decrease in carbon in the subtropical North Pacific, and much of the smaller-scale structure. Globally, the reconstruction has a spatial pattern correlation of $\rho = 0.48$ for these inter-annual changes, due partially to the choices for the cut-off radii.

Longer-term carbon changes in the Argo period are marked by widespread increases, particularly in the Northern Hemisphere, and a DIC decrease in the Pacific (Fig. 9e–f). The reconstruction captures the near-global increase in DIC and the decreases in DIC in the North Pacific and equatorial Pacific. The finer structures within the decadal DIC, such as the regions with the strongest DIC increases, are not fully captured by the reconstruction. The pattern correlation of $\rho = 0.71$ reflects the heightened role of atmospheric $p\text{CO}_2$ changes in setting carbon inventories on decadal timescales.

6 Discussion and conclusions

In this study a new method for reconstructing upper-ocean carbon using observations of ocean temperatures and salinities is presented. While the ocean plays a large role in determining the partitioning of carbon in the Earth system, sparse observations inhibit a full characterisation of ocean interior carbon. Using synthetic profiles and creating mapped reconstructions of model truths, we have explored how widespread synthetic observations of temperature and salinity representative of autonomous sampling programmes can provide global information about ocean carbon. Through reconstructions of CMIP6 model carbon fields, near-surface carbon can be reconstructed using synthetic Argo observations due to consistent controls of solubility, alkalinity, and undersaturation on the carbonate system. The method retains skill for reconstructions down to 1000 m. Uncertainties in the model covariance fields below 1000 m reduce the skill with which irregular hydrographic measurements can be used to reconstruct carbon changes.

6.1 Near-surface carbon reconstructions

When considering the carbon system in the near-surface ocean, the correlation fields between atmospheric $p\text{CO}_2$, ocean temperature, ocean salinity, and ocean carbon reflect first-order controls (Fig. 2). Increases in atmospheric $p\text{CO}_2$ are correlated with increases in ocean heat and carbon due to the chemical and thermal disequilibria created by emissions. Increases in ocean salinity are broadly correlated with increases in ocean carbon due to the impact of alkalinity on solubility. Regional variations in the correlations between temperature and ocean carbon can be decomposed as a sum of a response related to added carbon from emissions and a residual; the residual correlation reflects the impact of temperature on CO_2 solubility in seawater. Within all these correlation fields there are regional variations. The structure in the correlations between temperature and DIC and between salinity and DIC attain their strongest values in different regions, and the cross-correlations between temperature and DIC are generally more moderate, indicating that the information provided by these observations is complementary.

The strength and structure in the covariance fields result in optimal weights for $p\text{CO}_2$, temperature, and salinity that reflect global first-order controls of solubility and alkalinity (Fig. 4). For this linear model, using co-located temperature and salinity observations along with globally averaged atmospheric $p\text{CO}_2$ concentrations reduces the RMSE of the reconstructed DIC content by over 75 % on average relative to a reconstruction based purely on a time-invariant climatology (Fig. 5). In addition to replicating the long-term trend, the reconstructions are able to reproduce local inter-annual and decadal variability, capturing over 60 % of the detrended upper-ocean carbon signal. When considering the detrended carbon inventory, the sensitivity tests show an average relative RMSE reduction of 50 % in most regions. The lower skill in reconstructing inter-annual and decadal variability is expected as the carbon system is non-linear and controlled by other aspects such as circulation changes and ocean biology, which are not explicitly included in the variables used to calculate the covariance fields.

We have also explored whether the use of irregular temperature and salinity observations impacts the capacity to reconstruct ocean carbon. In theory, observations do not need to be co-located with the desired reconstruction, as the ensemble optimal interpolation scheme can include spatial covariances; however, there is the possibility of extra errors in the covariance fields from poorly represented spatial variability. In our sensitivity tests, we find that using irregular observations consistent with current Argo coverage can be combined with the CMIP6 covariance fields and provide similar levels of skill to the reconstruction. When taking the average RMSE reduction across our sensitivity tests, there is a reduction of over 60 % in most regions when considering the full carbon signal and a reduction of over 40 % in most regions when considering the detrended carbon signal (Fig. 6b,

d). However, regions such as the Southern Ocean and the North Atlantic show substantial error increases relative to the climatological first guess in some of the sensitivity experiments (Fig. 6a, d). These increases in errors are found particularly when considering the detrended signal (i.e. the signal from inter-annual and decadal variability). The regions that show the largest error increases are those with high-mesoscale ocean eddy activity and are mainly in the sensitivity test that reconstructs ocean carbon within MPI-ESM1.2-LR. MPI-ESM1.2-LR has a lower nominal resolution than the other models in our ensemble, which suggests that the models' representations of mesoscale ocean dynamics and their correlation length scales are important factors for any real-world carbon reconstruction.

6.2 Reconstructions of interior carbon

The Argo programme measures temperature and salinity to 2000 m, which potentially allows for the majority of ocean carbon changes to be captured by our reconstruction. However, within the interior, uncertainties in ocean circulation, ventilation pathways, and regenerated carbon pools can limit the skill of our reconstruction. We find that the optimal solutions for our interior carbon reconstructions are highly depth-dependent. Below 1000 m, the use of only temperature and salinity as oceanographic variables may be insufficient as the optimal coefficients become highly correlated.

Sensitivity experiments suggest that the heightened uncertainties in the CMIP6 models reduce the available skill in reconstructing carbon from only temperature and salinity observations at depths greater than 1000 m. This increase in errors below 1000 m arises from the ensemble construction, particularly the inclusion of the UKESM1. When the UKESM1 is excluded, the remaining models are more easily able to replicate each other's carbon fields. Before creating a real-world carbon reconstruction, further investigation is necessary to determine whether the uncertainties provided by outlier models such as MPI-ESM1.2-LR and UKESM1 are physically based or an artefact of model architecture or spin-up procedure.

When considering global reconstructions within the upper 2000 m, we find that the top 500 m is important for capturing both the trend and variability in ocean carbon. Within the top 500 m the method is able to reproduce the trend and variability with skill, and synthetic tests with the NorESM indicate that patterns of inter-annual and decadal DIC changes can be captured.

6.3 Flexibility of ensemble optimal interpolation

While we have opted to use only temperature and salinity as our ocean observations for this scheme, the ensemble optimal interpolation method is flexible and can take additional oceanographic variables such as oxygen, pH, nutrients, or chlorophyll. With increased coverage of these variables from

campaigns such as the Southern Ocean Carbon and Climate Observations and Modeling programme (SOCCOM, Johnson et al., 2017) and Bio-Argo (Claustre et al., 2020), there is the potential to include these biologically affected variables. For the high-latitude oceans where there are few Argo profiles, sea ice observations could lend additional information on upper-ocean carbon through the impact of sea ice on air–sea gas exchange.

With each new variable it is possible to quantify the amount of added information and test the ensemble optimal interpolation method at various stages by examining the covariance fields and uncertainties and conducting sensitivity tests in a similar way to that done in this work. As more complex reconstruction schemes such as the machine-learning algorithms of Landschützer et al. (2016) use these biogeochemical observations as well, including them in the ensemble optimal interpolation scheme could enable a direct comparison between linear and non-linear mapping methods and can help to quantify the merits of linear versus non-linear assumptions when reconstructing ocean interior carbon.

6.4 Usage of CMIP6 covariance fields

The ensemble optimal interpolation scheme relies on some important assumptions. Firstly, the covariance fields assume that the processes relating ocean carbon to other variables are stationary. We have used historical CMIP6 experiments to calculate the covariances fields, so these covariances should be able to represent ocean carbon behaviour under current-day carbon emission forcing; however, under low or negative emissions the relationships between atmospheric $p\text{CO}_2$, ocean temperature, and ocean carbon changes will likely change due to hysteresis from continued heat uptake. Additionally, we have chosen to focus on annual-average and depth-integrated carbon content in order to focus on the physical carbon response. It is also possible to consider seasonal or monthly variability by calculating covariance fields for each month or season, as is done for surface $p\text{CO}_2$ and heat content (Smith and Murphy, 2007; Jones et al., 2015). In the current format, months or seasons that are poorly observed will have solutions that return towards climatological inventories. However, the system can be set up in such a way that lagged covariance fields propagate temperature and salinity information from observed to unobserved months as well as locations. The lack of wintertime carbon observations has created significant biases in machine-learning products, as enhanced winter outgassing in polar waters is not present in the training data (Bushinsky et al., 2019). As the responses to gaps in data are sensitive to the reconstruction method, the ensemble optimal interpolation reconstructions with seasonal data could provide added insight as to how information on the carbon system is best propagated to unobserved regions and seasons.

The use of a model ensemble rather than a single model allows for quantification of some of the uncertainties within the

covariance fields. Our model ensemble has been constructed to consider both inter-model and intra-model uncertainty. We have been able to explore the sensitivity of this method to the ensemble makeup through out-of-sample reconstructions and assessment of the minimum, mean, and maximum RMSE improvements. These statistics are model-specific; for outlier models such as MPI-ESM1.2-LR for the upper 100 m and the UKESM1 for the 1000–2000 m layer, the poor reconstructions reduce the average skill. Including outlier models should dampen the solution towards a climatology in regions where there are physical uncertainties. The ensemble makeup itself is limited only by the number of CMIP6 model runs available, and so for future work we will include additional models with fewer realisations.

The improvements in the reconstructions indicate that both the models have regions with well-defined and strong correlations that lead to high improvements as well as regions where the models are relatively uncertain in their relationships (Figs. 5, 6). While these sensitivity experiments and synthetic reconstructions allow for useful insight into the capacity of our reconstruction method, some level of uncertainty remains from the use of coarse-resolution CMIP6 models. Ocean models underestimate decadal variability in anthropogenic carbon (DeVries et al., 2019) and CMIP6 models exhibit biases in the Revelle factor (Terhaar et al., 2022). More iterative ensemble optimal interpolation techniques can be used to reduce the errors from the model ensemble by recalculating the covariance fields after observations have been included in the analysis (Smith and Murphy, 2007; Smith et al., 2015; Allison et al., 2019).

6.5 Moving from synthetic to real-world reconstructions of ocean DIC

Our tests using synthetic reconstructions have suggested that a large amount of ocean carbon variability can be reconstructed using only hydrographic observations. As such, this work provides a theoretical basis for applying linear reconstruction methods to real-world Argo profile data to create an independent estimate of the recent ocean carbon inventory. The transition from synthetic reconstructions to real-world reconstructions involves careful understanding of observational errors.

Unlike model output, which can be considered to be “perfect” outside of rounding and regridding errors, observational data have multiple sources of uncertainty, such as sensor errors, the representativeness of area- and time-averaged fields from Lagrangian observations, and errors in the background climatology fields. While there is rigorous quality control for Argo observations, the representation errors and climatology errors may contribute a substantial amount of uncertainty to the real-world reconstruction. Reconstructions of monthly salinity and temperature have used parameterised observation error covariance fields (Smith and Murphy, 2007). Fu-

ture work towards a real-world reconstruction will include a biogeochemical extension of this parameterisation.

In conclusion, we have presented a method that draws upon first-principle relationships between temperature, salinity, and carbon to reconstruct ocean carbon changes from Argo-style observations. We find that, for what can be considered the mixed layer, the CMIP6 models show consistent first-principle constraints of solubility, alkalinity, and under-saturation, leading to accurate reproductions of trends and variability in the model world. When considering current-day distributions of Argo profiles, reconstructions that use multiple nearby observations are able to still reconstruct the trend and variability in carbon content. Below the mixed layer, the optimal coefficients become more spatially heterogeneous, and out-of-sample tests indicate lower skill in reconstructing carbon changes. However, because carbon uptake occurs at the atmosphere–ocean interface and the response within the top 500 m dominates the upper-ocean response, reconstructions of the upper 2000 m using Argo-type observations can recreate patterns of inter-annual and decadal variability. The results of our work provide a strong theoretical basis for using state-of-the-art autonomous hydrographic observations to supplement sparse interior carbon coverage. This proof-of-concept work shows that there is strong potential in using these measurements to create a new, independent estimate that can be used alongside reconstruction methods to better understand ocean carbon variability and its controls.

Appendix A: Breakdown of covariance fields into $p\text{CO}_2$ and non- $p\text{CO}_2$ terms

It is possible to decompose the changes in ocean temperature, salinity, and DIC from their time-averaged states (T' , S' , and DIC' from Sect. 2) into terms proportional to atmospheric $p\text{CO}_2$ and a residual:

$$T'(x, y, t) = \alpha(x, y)p\text{CO}_2'(t) + T_a(x, y, t), \quad (\text{A1})$$

$$S'(x, y, t) = \beta(x, y)p\text{CO}_2'(t) + S_a(x, y, t), \quad (\text{A2})$$

$$\text{DIC}'(x, y, t) = \gamma(x, y)p\text{CO}_2'(t) + \text{DIC}_a(x, y, t), \quad (\text{A3})$$

where α, β, γ are the least-squares coefficients against changes in globally uniform atmospheric $p\text{CO}_2$ changes and T_a, S_a, DIC_a indicate the residuals for these decompositions.

While the decomposition into $p\text{CO}_2$ and non- $p\text{CO}_2$ terms is not inherently physically based, this breakdown may help us understand how the complex structure in the correlation fields in Fig. 2 arises. Using covariance identities, the covariance between, for instance, T' and S' is thus the sum of the covariances between the addends:

$$\begin{aligned} \text{cov}(T', S') &= \text{cov}(\alpha p\text{CO}_2(t), \beta p\text{CO}_2(t)) \\ &\quad + \text{cov}(\alpha p\text{CO}_2(t), S_a) \\ &\quad + \text{cov}(T_a, \beta p\text{CO}_2(t)) + \text{cov}(T_a, S_a). \end{aligned} \quad (\text{A4})$$

The first term is a function of the variance of $p\text{CO}_2$, scaled by the product $\alpha\beta$. If the term proportional to atmospheric

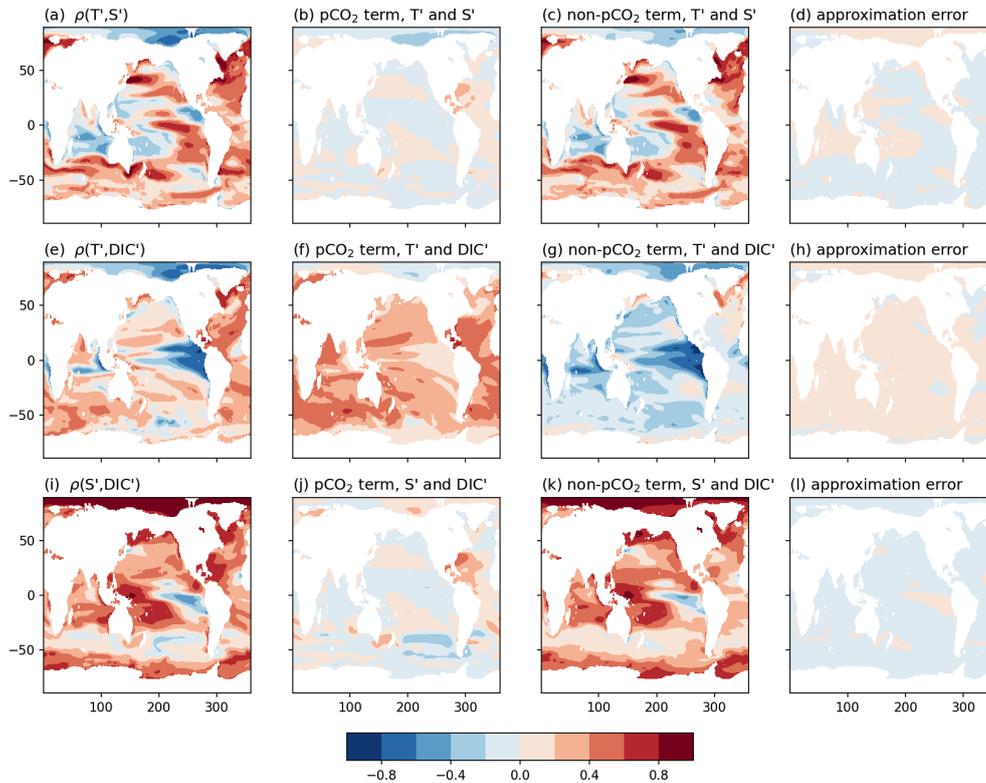


Figure A1. Correlation fields between non- $p\text{CO}_2$ variables T' , S' , and DIC' , with their terms from Eq. (A5) and approximation errors to measure the level of mismatch between Eqs. (A4) and (A5).

$p\text{CO}_2$ were orthogonal to the residual term for each of these decompositions, the covariances $\text{cov}(p\text{CO}_2, X_a) = 0$ for any variable X . An ordinary least-squares solution does not lead to orthogonal components, but the covariances between atmospheric $p\text{CO}_2$ and DIC_a , T_a , or S_a are small (shown as the approximation errors in Fig. A1d, h, and l) and can therefore be ignored to first order. Thus, for these variables, we ignore these terms as second-order components and can thus approximate the covariance as

$$\text{cov}(T', S') \approx \alpha\beta\text{var}(p\text{CO}_2) + \text{cov}(T_a, S_a). \quad (\text{A5})$$

With this decomposition of the covariance fields into $p\text{CO}_2$ and non- $p\text{CO}_2$ terms, we have the added benefit of approximating the covariance fields between $p\text{CO}_2$ and one of T' , S' , or DIC' through a global scaling of the coefficient terms α , β , or γ , respectively.

Scaling the decomposition by the product of the standard deviations allows us to explore how the correlation fields arise in Fig. 2. We compare the original correlation fields from Fig. 2 with their decompositions according to Eq. (A4) in Fig. A1. The approximation errors show how well the approximation in Eq. (A5) reflects the full correlation field.

For the correlations between T' and S' and S' and DIC' , the signal is dominated by the term associated with $\text{cov}(T_a, S_a)$ and $\text{cov}(S_a, \text{DIC}_a)$, respectively. We therefore understand these covariance and correlation fields to be set by

the variability not forced by carbon emissions. Conversely, the correlation field between T' and DIC' appears to arise almost equally from the $p\text{CO}_2$ scaling and the $\text{cov}(T_a, \text{DIC}_a)$ terms. Thus we provide both terms in Sect. 3 to gain insight as to how the covariance fields reflect both the combined response to added carbon in the Earth system as well as the response to climate variability.

Code availability. The code to de-drift, regrid, concatenate, and analyse CMIP6 model output for synthetic DIC reconstructions is available in a Zenodo repository (<https://doi.org/10.5281/zenodo.7782759>, Turn, 2023).

Data availability. The CMIP6 model outputs used in this study are publicly available at <https://esgf-node.llnl.gov/search/cmip6/> (ESGF, 2022).

Supplement. The supplement related to this article is available online at: <https://doi.org/10.5194/bg-20-1671-2023-supplement>.

Author contributions. All the authors contributed towards the conceptualisation of the work. KET created the method and analysis with input from RGW, AK, and DMS. KET wrote the manuscript

draft, and all the authors reviewed and edited the manuscript. All the authors signed off on the manuscript for submission.

Competing interests. The authors declare that they have no competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Acknowledgements. The authors thank two anonymous referees for their constructive feedback.

Financial support. This research has been supported by the Department for Business, Energy and Industrial Strategy, UK Government (Met Office Hadley Centre Climate Programme), the Department for Environment, Food and Rural Affairs, UK Government Met Office Hadley Centre Climate Programme, the Natural Environment Research Council (grant no. NE/T007788/1), and the Leverhulme Trust via the Leverhulme Research Centre for Functional Materials Design.

Review statement. This paper was edited by Peter Landschützer and reviewed by two anonymous referees.

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, *Environ. Res. Lett.*, 14, 084037, <https://doi.org/10.1088/1748-9326/ab2b0b>, 2019.
- Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S.: A multi-decade record of high-quality $f\text{CO}_2$ data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT), *Earth Syst. Sci. Data*, 8, 383–413, <https://doi.org/10.5194/essd-8-383-2016>, 2016.
- Bates, N. R., Astor, Y. M., Church, M. J., Currie, K., Dore, J. E., González-Dávila, M., Lorenzoni, L., Muller-Karger, F., Olafsson, J., and Santana-Casiano, J. M.: A Time-Series View of Changing Ocean Chemistry Due to Ocean Uptake of Anthropogenic CO_2 and Ocean Acidification, *Oceanography*, 27, 126–141, <https://doi.org/10.5670/oceanog.2014.16>, 2014.
- Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., Bekki, S., Bonnet, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Caubel, A., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., D'Andrea, F., Davini, P., Lavergne, C., Denvil, S., Deshayes, J., Devilliers, M., Ducharne, A., Dufresne, J., Dupont, E., Éthé, C., Fairhead, L., Falletti, L., Flavoni, S., Foujols, M., Gardoll, S., Gastineau, G., Ghattas, J., Grandpeix, J., Guenet, B., Guez, E., L., Guilyardi, E., Guimberteau, M., Hauglustaine, D., Hourdin, F., Idelkadi, A., Jous-saume, S., Kageyama, M., Khodri, M., Krinner, G., Lebas, N., Levassasseur, G., Lévy, C., Li, L., Lott, F., Lurton, T., Luys-saert, S., Madec, G., Madeleine, J., Maignan, F., Marchand, M., Marti, O., Mellul, L., Meurdesoif, Y., Mignot, J., Musat, I., Otlé, C., Peylin, P., Planton, Y., Polcher, J., Rio, C., Rochetin, N., Rousset, C., Sepulchre, P., Sima, A., Swingedouw, D., Thiéblemont, R., Traore, A. K., Vancoppenolle, M., Vial, J., Vialard, J., Viovy, N., and Vuichard, N.: Presentation and Evaluation of the IPSL-CM6A-LR Climate Model, *J. Adv. Model. Earth Sy.*, 12, e2019MS002010, <https://doi.org/10.1029/2019MS002010>, 2020.
- Bourgeois, T., Goris, N., Schwinger, J., and Tjiputra, J. F.: Stratification constrains future heat and carbon uptake in the Southern Ocean between 30°S and 55°S , *Nat. Commun.*, 13, 1–8, 2022.
- Burger, F. A., John, J. G., and Frölicher, T. L.: Increase in ocean acidity variability and extremes under increasing atmospheric CO_2 , *Biogeosciences*, 17, 4633–4662, <https://doi.org/10.5194/bg-17-4633-2020>, 2020.
- Bushinsky, S. M., Landschützer, P., Rödenbeck, C., Gray, A. R., Baker, D., Mazloff, M. R., Resplandy, L., Johnson, K. S., and Sarmiento, J. L.: Reassessing Southern Ocean Air-Sea CO_2 Flux Estimates With the Addition of Biogeochemical Float Observations, *Global Biogeochem. Cy.*, 33, 1370–1388, <https://doi.org/10.1029/2019GB006176>, 2019.
- Carroll, D., Menemenlis, D., Adkins, J. F., Bowman, K. W., Brix, H., Dutkiewicz, S., Fenty, I., Gierach, M. M., Hill, C., Jahn, O., Landschützer, P., Lauderdale, J. M., Liu, J., Manizza, M., Naviaux, J. D., Rödenbeck, C., Schimel, D. S., Van der Stocken, T., and Zhang, H.: The ECCO-Darwin Data-Assimilative Global Ocean Biogeochemistry Model: Estimates of Seasonal to Multidecadal Surface Ocean pCO_2 and Air-Sea CO_2 Flux, *J. Adv. Model. Earth Sy.*, 12, e2019MS001888, <https://doi.org/10.1029/2019MS001888>, 2020.
- Cheng, L. and Zhu, J.: Benefits of CMIP5 multimodel ensemble in reconstructing historical ocean subsurface temperature variations, *J. Climate*, 29, 5393–5416, 2016.
- Claustre, H., Johnson, K. S., and Takeshita, Y.: Observing the global ocean with biogeochemical-Argo, *Annu. Rev. Mar. Sci.*, 23–48, 2020.
- Clement, D. and Gruber, N.: The eMLR(C*) Method to Determine Decadal Changes in the Global Ocean Storage of

- Anthropogenic CO₂, *Global Biogeochem. Cy.*, 32, 654–679, <https://doi.org/10.1002/2017GB005819>, 2018.
- Daley, R.: *Atmospheric data analysis*, Cambridge University Press, Cambridge, ISBN 9780521458252, 1991.
- Danabasoglu, G., Lamarque, J., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., Emmons, L. K., Fasullo, J., Garcia, R., Gettelman, A., Hannay, C., Holland, M. M., Large, W. G., Lauritzen, P. H., Lawrence, D. M., Lenaerts, J. T. M., Lindsay, K., Lipscomb, W. H., Mills, M. J., Neale, R., Oleson, K. W., Otto-Bliesner, B., Phillips, A. S., Sacks, W., Tilmes, S., Kampenhout, L., Vertenstein, M., Bertini, A., Dennis, J., Deser, C., Fischer, C., Fox-Kemper, B., Kay, J. E., Kinnison, D., Kushner, P. J., Larson, V. E., Long, M. C., Mickelson, S., Moore, J. K., Nienhouse, E., Polvani, L., Rasch, P. J., and Strand, W. G.: The Community Earth System Model Version 2 (CESM2), *J. Adv. Model. Earth Sy.*, 12, <https://doi.org/10.1029/2019MS001916>, 2020.
- DeVries, T.: The oceanic anthropogenic CO₂ sink: Storage, air-sea fluxes, and transports over the industrial era, *Global Biogeochem. Cy.*, 28, 631–647, <https://doi.org/10.1002/2013GB004739>, 2014.
- DeVries, T., Holzer, M., and Primeau, F.: Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning, *Nature*, 542, 215–218, 2017.
- DeVries, T., Le Quéré, C., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., Landschützer, P., Lenton, A., Lima, I. D., Nowicki, M., Schwinger, J., and Séférian, R.: Decadal trends in the ocean carbon sink, *P. Natl. Acad. Sci. USA*, 116, 11646–11651, 2019.
- ESGF: WCRP Coupled Model Intercomparison Project (Phase 6), CMIP [data set], <https://esgf-node.llnl.gov/search/cmip6/> (last access: 26 April 2023), 2022.
- Friedlingstein, P., Jones, M. W., O’Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., Le Quéré, C., Peters, G. P., Peters, W., Pongratz, J., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Anthoni, P., Bates, N. R., Becker, M., Belloouin, N., Bopp, L., Chau, T. T. T., Chevallier, F., Chini, L. P., Cronin, M., Currie, K. I., Decharme, B., Djeutchouang, L. M., Dou, X., Evans, W., Feely, R. A., Feng, L., Gasser, T., Gilfillan, D., Gkritzalis, T., Grassi, G., Gregor, L., Gruber, N., Gürses, Ö., Harris, I., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Luijkx, I. T., Jain, A., Jones, S. D., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lienert, S., Liu, J., Marland, G., McGuire, P. C., Melton, J. R., Munro, D. R., Nabel, J. E. M. S., Nakaoka, S.-I., Niwa, Y., Ono, T., Pierrot, D., Poulter, B., Rehder, G., Resplandy, L., Robertson, E., Rödenbeck, C., Rosan, T. M., Schwinger, J., Schwingshackl, C., Séférian, R., Sutton, A. J., Sweeney, C., Tanhua, T., Tans, P. P., Tian, H., Tilbrook, B., Tubiello, F., van der Werf, G. R., Vuichard, N., Wada, C., Wanninkhof, R., Watson, A. J., Willis, D., Wiltshire, A. J., Yuan, W., Yue, C., Yue, X., Zaehle, S., and Zeng, J.: *Global Carbon Budget 2021*, *Earth Syst. Sci. Data*, 14, 1917–2005, <https://doi.org/10.5194/essd-14-1917-2022>, 2022.
- Gloege, L., Yan, M., Zheng, T., and McKinley, G. A.: Improved Quantification of Ocean Carbon Uptake by Using Machine Learning to Merge Global Models and pCO₂ Data, *J. Adv. Model. Earth Sy.*, 14, e2021MS002620, <https://doi.org/10.1029/2021MS002620>, 2022.
- Gruber, N., Clement, D., Carter, B. R., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Key, R. M., Kozyr, A., Lauvset, S. K., Monaco, C. L., Mathis, J. T., Murata, A., Olsen, A., Perez, F. F., Sabine, C. L., Tanhua, T., and Wanninkhof, R.: The oceanic sink for anthropogenic CO₂ from 1994 to 2007, *Science*, 363, 1193–1199, <https://doi.org/10.1126/science.aau5153>, 2019a.
- Gruber, N., Landschützer, P., and Lovenduski, N. S.: The variable Southern Ocean carbon sink, *Annu. Rev. Mar. Sci.*, 11, 159–186, <https://doi.org/10.1146/annurev-marine-121916-063407>, 2019b.
- Hauri, C., McDonnell, A. M., Stuecker, M. F., Danielson, S. L., Hedstrom, K., Irving, B., Schultz, C., and Doney, S. C.: Modulation of ocean acidification by decadal climate variability in the Gulf of Alaska, *Commun. Earth Environ.*, 2, 1–7, 2021.
- Johnson, K. S., 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 SOCOM profiling float array, *J. Geophys. Res.-Oceans*, 122, 6416–6436, 2017.
- Jones, S. D., Le Quéré, C., Rödenbeck, C., Manning, A. C., and Olsen, A.: A statistical gap-filling method to interpolate global monthly surface ocean carbon dioxide data, *J. Adv. Model. Earth Sy.*, 7, 1554–1575, <https://doi.org/10.1002/2014MS000416>, 2015.
- Keppler, L., Landschützer, P., Gruber, N., Lauvset, S. K., and Stemmler, I.: Seasonal Carbon Dynamics in the Near-Global Ocean, *Global Biogeochem. Cy.*, 34, e2020GB006571, <https://doi.org/10.1029/2020GB006571>, 2020.
- Khatriwala, S., Tanhua, T., Mikaloff Fletcher, S., Gerber, M., Doney, S. C., Graven, H. D., Gruber, N., McKinley, G. A., Murata, A., Ríos, A. F., and Sabine, C. L.: Global ocean storage of anthropogenic carbon, *Biogeosciences*, 10, 2169–2191, <https://doi.org/10.5194/bg-10-2169-2013>, 2013.
- Landschützer, P., Gruber, N., and Bakker, D. C. E.: Decadal variations and trends of the global ocean carbon sink, *Global Biogeochem. Cy.*, 30, 1396–1417, <https://doi.org/10.1002/2015GB005359>, 2016.
- Landschützer, P., Ilyina, T., and Lovenduski, N. S.: Detecting Regional Modes of Variability in Observation-Based Surface Ocean pCO₂, *Geophys. Res. Lett.*, 46, 2670–2679, <https://doi.org/10.1029/2018GL081756>, 2019.
- Lauderdale, J. M., Dutkiewicz, S., Williams, R. G., and Follows, M. J.: Quantifying the drivers of ocean-atmosphere CO₂ fluxes, *Global Biogeochem. Cy.*, 30, 983–999, <https://doi.org/10.1002/2016GB005400>, 2016.
- Lauvset, S. K., Lange, N., Tanhua, T., Bittig, H. C., Olsen, A., Kozyr, A., Álvarez, M., Becker, S., Brown, P. J., Carter, B. R., Cotrim da Cunha, L., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Jeansson, E., Jutterström, S., Jones, S. D., Karlsen, M. K., Lo Monaco, C., Michaelis, P., Murata, A., Pérez, F. F., Pfeil, B., Schirnick, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Velo, A., Wanninkhof, R., Woosley, R. J., and Key, R. M.: An updated version of the global interior ocean biogeochemical data product, *GLODAPv2.2021*, *Earth Syst. Sci. Data*, 13, 5565–5589, <https://doi.org/10.5194/essd-13-5565-2021>, 2021.
- Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez-de-la-Cuesta, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornbluh, L., Marotzke, J., Matei, D., Meraner, K.,

- Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T. J., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Storch, J., Tian, F., Voigt, A., Vrese, P., Wieners, K., Wilkenskjaeld, S., Winkler, A., and Roeckner, E.: Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM1.2) and Its Response to Increasing CO₂, *J. Adv. Model. Earth Sy.*, 11, 998–1038, <https://doi.org/10.1029/2018MS001400>, 2019.
- McKinley, G. A., Pilcher, D. J., Fay, A. R., Lindsay, K., Long, M. C., and Lovenduski, N. S.: Timescales for detection of trends in the ocean carbon sink, *Nature*, 530, 469–472, 2016.
- McKinley, G. A., Fay, A. R., Lovenduski, N. S., and Pilcher, D. J.: Natural Variability and Anthropogenic Trends in the Ocean Carbon Sink, *Annu. Rev. Mar. Sci.*, 9, 125–150, <https://doi.org/10.1146/annurev-marine-010816-060529>, 2017.
- McKinley, G. A., Fay, A. R., Eddebbar, Y. A., Gloege, L., and Lovenduski, N. S.: External Forcing Explains Recent Decadal Variability of the Ocean Carbon Sink, *AGU Advances*, 1, e2019AV000149, <https://doi.org/10.1029/2019AV000149>, 2020.
- Seland, Ø., Bentsen, M., Olivie, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., Debernard, J. B., Gupta, A. K., He, Y.-C., Kirkevåg, A., Schwinger, J., Tjiputra, J., Aas, K. S., Bethke, I., Fan, Y., Griesfeller, J., Grini, A., Guo, C., Ilicak, M., Karset, I. H. H., Landgren, O., Liakka, J., Moseid, K. O., Nummelin, A., Spensberger, C., Tang, H., Zhang, Z., Heinze, C., Iversen, T., and Schulz, M.: Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations, *Geosci. Model Dev.*, 13, 6165–6200, <https://doi.org/10.5194/gmd-13-6165-2020>, 2020.
- Sellar, A. A., Jones, C. G., Mulcahy, J. P., Tang, Y., Yool, A., Wiltshire, A., O'Connor, F. M., Stringer, M., Hill, R., Palmieri, J., Woodward, S., Mora, L., Kuhlbrodt, T., Rumbold, S. T., Kelley, D. I., Ellis, R., Johnson, C. E., Walton, J., Abraham, N. L., Andrews, M. B., Andrews, T., Archibald, A. T., Berthou, S., Burke, E., Blockley, E., Carslaw, K., Dalvi, M., Edwards, J., Folberth, G. A., Gedney, N., Griffiths, P. T., Harper, A. B., Hendry, M. A., Hewitt, A. J., Johnson, B., Jones, A., Jones, C. D., Keeble, J., Liddicoat, S., Morgenstern, O., Parker, R. J., Predoi, V., Robertson, E., Siahann, A., Smith, R. S., Swaminathan, R., Woodhouse, M. T., Zeng, G., and Zerroukat, M.: UKESM1: Description and Evaluation of the U.K. Earth System Model, *J. Adv. Model. Earth Sy.*, 11, 4513–4558, <https://doi.org/10.1029/2019MS001739>, 2019.
- Sloyan, B. M., Wanninkhof, R., Kramp, M., Johnson, G. C., Talley, L. D., Tanhua, T., McDonagh, E., Cusack, C., O'Rourke, E., McGovern, E., Katsumata, K., Diggs, S., Hummon, J., Ishii, M., Azetsu-Scott, K., Boss, E., Anson, I., Perez, F. F., Mercier, H., Williams, M. J. M., Anderson, L., Lee, J. H., Murata, A., Kouketsu, S., Jeansson, E., Hoppema, M., and Campos, E.: The Global Ocean Ship-Based Hydrographic Investigations Program (GO-SHIP): A Platform for Integrated Multidisciplinary Ocean Science, *Front. Mar. Sci.*, 6, 445, <https://doi.org/10.3389/fmars.2019.00445>, 2019.
- Smith, D. M. and Murphy, J. M.: An objective ocean temperature and salinity analysis using covariances from a global climate model, *J. Geophys. Res.-Oceans*, 112, C02022, <https://doi.org/10.1029/2005JC003172>, 2007.
- Smith, D. M., Allan, R. P., Coward, A. C., Eade, R., Hyder, P., Liu, C., Loeb, N. G., Palmer, M. D., Roberts, C. D., and Scaife, A. A.: Earth's energy imbalance since 1960 in observations and CMIP5 models, *Geophys. Res. Lett.*, 42, 1205–1213, <https://doi.org/10.1002/2014GL062669>, 2015.
- Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., Anstey, J., Arora, V., Christian, J. R., Hanna, S., Jiao, Y., Lee, W. G., Majaess, F., Saenko, O. A., Seiler, C., Seinen, C., Shao, A., Sigmond, M., Solheim, L., von Salzen, K., Yang, D., and Winter, B.: The Canadian Earth System Model version 5 (CanESM5.0.3), *Geosci. Model Dev.*, 12, 4823–4873, <https://doi.org/10.5194/gmd-12-4823-2019>, 2019.
- Terhaar, J., Kwiatkowski, L., and Bopp, L.: Emergent constraint on Arctic Ocean acidification in the twenty-first century, *Nature*, 582, 379–383, 2020.
- Terhaar, J., Frölicher, T. L., and Joos, F.: Southern Ocean anthropogenic carbon sink constrained by sea surface salinity, *Sci. Adv.*, 7, eabd5964, <https://doi.org/10.1126/sciadv.abd5964>, 2021.
- Terhaar, J., Frölicher, T. L., and Joos, F.: Observation-constrained estimates of the global ocean carbon sink from Earth system models, *Biogeosciences*, 19, 4431–4457, <https://doi.org/10.5194/bg-19-4431-2022>.
- Thomas, J., Waugh, D., and Gnanadesikan, A.: Relationship between Ocean Carbon and Heat Multidecadal Variability, *J. Climate*, 31, 1467–1482, <https://doi.org/10.1175/JCLI-D-17-0134.1>, 2018.
- Turner, K.: Synthetic ocean DIC ENOI (v2.0.0), Zenodo [code], <https://doi.org/10.5281/zenodo.7782759>, 2023.
- Verdy, A. and Mazloff, M. R.: A data assimilating model for estimating Southern Ocean biogeochemistry, *J. Geophys. Res.-Oceans*, 122, 6968–6988, <https://doi.org/10.1002/2016JC012650>, 2017.
- Weiss, R.: Carbon dioxide in water and seawater: the solubility of a non-ideal gas, *Mar. Chem.*, 2, 203–215, 1974.
- Williams, R. G. and Follows, M. J.: *Ocean Dynamics and the Carbon Cycle: Principles and Mechanisms*, Cambridge University Press, 2011.
- Williams, R. G., Katavouta, A., and Roussenov, V.: Regional Asymmetries in Ocean Heat and Carbon Storage due to Dynamic Redistribution in Climate Model Projections, *J. Climate*, 34, 3907–3925, <https://doi.org/10.1175/JCLI-D-20-0519.1>, 2021.
- 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., Mercœur, 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., 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., Gouillon, 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., Iwasaka,

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., 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, *Front. Mar. Sci.*, 7, 700, <https://doi.org/10.3389/fmars.2020.00700>, 2020.

Ziehn, T., Chamberlain, M. A., Law, R. M., Lenton, A., Bodman, R. W., Dix, M., Stevens, L., Wang, Y.-P., and Srbinovsky, J.: The Australian Earth System Model: ACCESS-ESM1.5, *Journal of Southern Hemisphere Earth Systems Science*, 70, 193–214, <https://doi.org/10.1071/ES19035>, 2020.