## Title
Alternate Histories: Synthetic Large Ensembles of Sea-Air CO2 Flux



## Journal
Global Biogeochemical Cycles, 36(6)

## ISSN
0886-6236

## Authors
Olivarez, Holly C
Lovenduski, Nicole S
Brady, Riley X
et al.

## Publication Date
2022-06-01

## DOI
10.1029/2021gb007174

Peer reviewed

Global Biogeochemical Cycles

RESEARCH ARTICLE
10.1029/2021GB007174

Special Section:
Understanding carbon-climate feedbacks

# Alternate Histories: Synthetic Large Ensembles of Sea-Air $CO_2$ Flux

**Holly C. Olivarez[1,2], Nicole S. Lovenduski[2,3], Riley X. Brady[2,3], Amanda R. Fay[4], Marion Gehlen[5], Luke Gregor[6], Peter Landschützer[7], Galen A. McKinley[4], Karen A. McKinnon[8], and David R. Munro[9]**

[1]Department of Environmental Studies, University of Colorado, Boulder, CO, USA, [2]Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, USA, [3]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder, CO, USA, [4]Columbia University and Lamont-Doherty Earth Observatory, Palisades, NY, USA, [5]Institut Pierre-Simon Laplace, Laboratoire des Sciences du Climat et de l'Environnement, Gif-sur-Yvette, France, [6]Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland, [7]Max Planck Institute for Meteorology, Hamburg, Germany, [8]Department of Statistics and Institute of the Environment and Sustainability, University of California, Los Angeles, CA, USA, [9]National Oceanic and Atmospheric Administration, Global Monitoring Laboratory and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

**Key Points:**

- We construct synthetic large ensembles of observed sea-air carbon dioxide flux using a statistical emulation technique
- The synthetic large ensembles illustrate an important role for internal variability in the temporal evolution of carbon dioxide flux
- We find a wide range of possible decadal trends in carbon dioxide flux over 1990–1999 and 2000–2009 driven by internal variability

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

H. C. Olivarez,
holly.olivarez@colorado.edu

**Citation:**

Olivarez, H. C., Lovenduski, N. S., Brady, R. X., Fay, A. R., Gehlen, M., Gregor, L., et al. (2022). Alternate histories: Synthetic large ensembles of sea-air $CO_2$ flux. *Global Biogeochemical Cycles*, *36*, e2021GB007174. https://doi.org/10.1029/2021GB007174

Received 27 AUG 2021
Accepted 20 MAY 2022



**Abstract** We use a statistical emulation technique to construct synthetic ensembles of global and regional sea-air carbon dioxide ($CO_2$) flux from four observation-based products over 1985–2014. Much like ensembles of Earth system models that are constructed by perturbing their initial conditions, our synthetic ensemble members exhibit different phasing of internal variability and a common externally forced signal. Our synthetic ensembles illustrate an important role for internal variability in the temporal evolution of global and regional $CO_2$ flux and produce a wide range of possible trends over 1990–1999 and 2000–2009. We assume a specific externally forced signal and calculate the rank of the observed trends within the distribution of statistically modeled synthetic trends during these periods. Over the decade 1990–1999, three of four observation-based products exhibit small negative trends in globally integrated sea-air $CO_2$ flux (i.e., enhanced ocean $CO_2$ absorption with time) that are within one standard deviation of the mean in their respective synthetic ensembles. Over the decade 2000–2009, however, three products show large negative trends in globally integrated sea-air $CO_2$ flux that have a low rate of occurrence in their synthetic ensembles. The largest positive trends in global and Southern Ocean flux over 1990–1999 and the largest negative trends over 2000–2009 fall nearly two standard deviations away from the mean in their ensembles. Our approach provides a new perspective on the important role of internal variability in sea-air $CO_2$ flux trends, and furthers understanding of the role of internal and external processes in driving observed sea-air $CO_2$ flux variability.

## 1. Introduction

The ocean plays a key role in the climate system, absorbing ~25% of the annual carbon dioxide ($CO_2$) emissions from anthropogenic activities over the years 1960–2020 (Friedlingstein et al., 2021). While this sea-air $CO_2$ flux slows the rate of anthropogenic climate change (Le Quéré et al., 2018), it also enhances ocean acidification and can thus influence marine organisms, ecosystems, and the societies that depend on those ecosystems (Doney et al., 2020). Earth system models suggest that ocean carbon absorption will continue through the end of the century (Canadell et al., 2021; Ciais & Sabine, 2013), though the magnitude of the globally integrated sea-air $CO_2$ flux will largely depend on our emissions trajectory (Canadell et al., 2021; Friedlingstein et al., 2021; Lovenduski et al., 2016; Ridge & McKinley, 2021).

Global sea-air $CO_2$ exchange is not steady with time, but rather exhibits temporal variability. Studies using estimates of sea-air $CO_2$ flux from sparse measurements of the surface ocean partial pressure of $CO_2$ ($pCO_2$) (Bakker et al., 2016) suggest that this $CO_2$ flux variability is particularly pronounced on decadal timescales. These studies report a period of stagnation in global ocean carbon absorption over the decade 1990–1999 (DeVries et al., 2019; Landschützer et al., 2016; Le Quéré et al., 2009; Rödenbeck et al., 2015), followed by intensification of ocean carbon absorption over the decade 2000–2009 (DeVries et al., 2019; Fay & McKinley, 2013; Landschützer et al., 2016; Rödenbeck et al., 2015). These observed decadal trends in sea-air $CO_2$ flux are superimposed on a background characterized by high interannual variability on global and regional scales (Landschützer et al., 2019), and this challenges our ability to quantify the magnitude of the decadal trends and to attribute them

to particular drivers (for example, Fay & McKinley, 2013; Le Quéré et al., 2007; Metzl et al., 2010; Schuster & Watson, 2007; Schuster et al., 2009; Thomas et al., 2008; Watson et al., 2009). While some studies link the decadal sea-air $CO_2$ flux trends to modes of *internal climate variability*, such as the Southern Annular Mode or the El Niño-Southern Oscillation (ENSO; Landschützer et al., 2015, 2019), others cite external forcing from volcanic eruptions and changes in the atmospheric $CO_2$ growth rate as the driving factor behind these trends (McKinley et al., 2020). It is critical that we quantify and understand the drivers of these decadal trends in sea-air $CO_2$ flux for future predictions of the global carbon cycle that are reported in documents such as the Intergovernmental Panel on Climate Change (IPCC) reports.

Large initial condition ensembles of Earth system models are a relatively new tool that can be used to quantify the roles of internal climate variability and external forcing in long-term trends of Earth system variables. These large ensemble experiments are conducted with a single Earth system model wherein each ensemble member is subject to perturbations in initial conditions, but all ensemble members are subject to identical external forcing. This procedure produces an ensemble where each member portrays modes of internal climate variability with unique phasing and amplitude, and where the average across all ensemble members captures the response of the Earth system to external forcing (Deser et al., 2020). McKinley et al. (2016) and McKinley et al. (2017) used the Community Earth System Model Version 1 Large Ensemble (CESM1-LE; Kay et al., 2015) to illustrate how internal variability can cloud our ability to quantify and interpret sea-air $CO_2$ flux trends on decadal and longer timescales. Their analysis demonstrates that decadal trends in sea-air $CO_2$ flux from a single CESM1-LE ensemble member are strongly affected by internal climate variability (McKinley et al., 2017). Since the historical record of sea-air $CO_2$ flux variations is akin to a single ensemble member in this large ensemble framework, the magnitude of decadal trends in the historical record is likely heavily influenced by internal variability. However, the magnitude of sea-air $CO_2$ flux variability in CESM1-LE and other Earth system models may not match that of the real world (Anav et al., 2013; Dong et al., 2016), and this necessitates our development of a large ensemble that is based on real-world observations.

Here, we use a statistical emulation method to place the observed estimates of sea-air $CO_2$ flux into a large ensemble framework by constructing synthetic ensembles of sea-air $CO_2$ flux from observation-based products. Much like a large ensemble of an Earth system model, each synthetic ensemble member experiences a different phasing of internal climate variability, but an identical externally forced signal. We develop synthetic ensembles of sea-air $CO_2$ flux for four observation-based products and remark on the importance of internal climate variability for the interpretation of decadal trends in the observational record.

Section 2 reviews the observations, observation-based products, and models used for this study. Section 3 describes the method of synthetic ensemble construction using real-world observations from the Drake Passage as an example. Section 4 is a presentation of our Results of synthetic ensembles from four observation-based $CO_2$ flux products. We test our approach and mention the assumptions made for this study in Section 5. Conclusions and Discussion are in Section 6.

## 2. Observations and Models

Our study utilizes a collection of interpolated observations and output from Earth system models to develop, analyze, and test our synthetic ensemble of observed sea-air $CO_2$ fluxes. We illustrate our statistical methodology for the reader using sea-air $CO_2$ fluxes derived from surface ocean $pCO_2$ ($pCO_2^{oc}$) observations collected in the Drake Passage Time-series program. We then develop synthetic ensembles for four global, observation-based sea-air $CO_2$ flux products, for which we use ensemble mean estimates of sea-air $CO_2$ flux from Earth system models contributing to the sixth Coupled Model Intercomparison Project (CMIP6). Finally, we use output from the CESM1-LE and an upper-ocean box model representation of external forcing to test our statistical methodology. In this section, we describe each of these datasets in turn.

### 2.1. Drake Passage Sea-Air $CO_2$ Flux Estimates

We use a single time-series of annual mean sea-air $CO_2$ flux derived from underway estimates of $pCO_2^{oc}$ collected as part of the Drake Passage Time-series program over 2004–2018 (Figure 1a; Fay et al., 2018; Munro, Lovenduski, Stephens, et al., 2015; Munro, Lovenduski, Takahashi, et al., 2015). Sea-air $CO_2$ flux calculated as in Sweeney et al. (2007). NOAA Greenhouse Gas Marine Boundary Layer Reference used for atmospheric $xCO_2$

(NOAA Global Monitoring Laboratory, web). Each annual mean estimate of sea-air $CO_2$ flux is calculated from monthly means of all underway $pCO_2^{oc}$ observations within a region in the center of the Drake Passage (i.e., from 58 to 60°S and 61.5 to 65.5°W) where monthly Cross-Calibrated Multi-Platform version 2 (CCMPv2) winds were used to estimate sea-air $CO_2$ flux (Atlas et al., 2011). Observations were collected in eight to eleven different months of each year within this region, from approximately twenty Southern Ocean crossings per year.

### 2.2. Observation-Based Products: Sea-Air $CO_2$ Flux

Our synthetic ensembles of sea-air $CO_2$ flux are derived from observation-based estimates that rely upon sparse $pCO_2^{oc}$ measurements collected in the Surface Ocean $CO_2$ Atlas (SOCAT; Bakker et al., 2016) and the Lamont-Doherty Earth Observatory database (LDEO; Takahashi et al., 2018). These observation-based products use a range of statistical and machine learning approaches to gap-fill $pCO_2^{oc}$ where and when measurements are not available.

In this study, we use annual mean sea-air $CO_2$ flux estimates from four observation-based products as produced by the SeaFlux product (Fay et al., 2021; Gregor & Fay, 2021; Table 1): The Council for Scientific and Industrial Research-Machine Learning ensemble (CSIR-ML6; Gregor et al., 2019), the Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN; Landschützer et al., 2013, 2014, 2015, 2016), the Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS; Rödenbeck et al., 2014), and the Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN; Denvil-Sommer et al., 2019). Fluxes are calculated monthly and then averaged to annual values for analysis. The SeaFlux product corrects for the spatial coverage differences in each observation-based product $pCO_2$ coverage by filling missing areas in each with a scaled climatology product which extends to coastal and high latitude regions (Landschützer et al., 2020). $CO_2$ flux is then estimated from each product's area-filled $pCO_2$ using common atmosphere, ice, and solubility, wind speed inputs, and a quadratic flux parameterization (Fay et al., 2021 for details). We use only the common observation-based product period of 1985–2014 for analysis.

This method of flux calculation pre-processing ensures that differences in decadal trends or interannual variance in our synthetic ensembles are due solely to differences in the $pCO_2$ products themselves rather than from statistical artifacts or flux calculation parameter choices. The spread between these products' fluxes is due to the difference in machine learning or regression techniques used in each to interpolate full coverage from the available $pCO_2$ observations.

### 2.3. Community Earth System Model Version 1 Large Ensemble

We evaluate our statistical methodology using output from CESM1-LE. CESM Version 1 is a fully coupled climate model that simulates Earth's climate system (Hurrell et al., 2013). The model is comprised of four component models that synchronously simulate Earth's land, atmosphere, ocean, and sea ice, with one central coupler component that exchanges fluxes and boundary conditions between the individual components (Hurrell et al., 2013). The ocean component model of CESM1 is the Parallel Ocean Program model with nominal 1° resolution and 60 vertical levels (Danabasoglu et al., 2012) coupled to the Biogeochemical Elemental Cycling model for ocean biogeochemistry, including full carbonate chemistry thermodynamics and sea-air $CO_2$ fluxes (Moore et al., 2004; Moore & Braucher, 2008; Moore & Doney, 2007). We analyze 34 ensemble members of CESM1-LE that span 1920–2005 and are forced with historical greenhouse gas and aerosol concentrations developed for the 5th Coupled Model Intercomparison Project (CMIP5; Kay et al., 2015; Taylor et al., 2012). Random phasing of internal climate modes is accomplished in CESM1-LE via round-off-level differences in the 1 January 1920 air temperatures (Kay et al., 2015).

### 2.4. Earth System Models From CMIP6

We take advantage of newly available output from three CMIP6 Earth system models with active ocean biogeochemistry that submitted multiple historical (1850–2014) ensemble members derived from initial conditions perturbations to the CMIP6 archive: the Canadian Earth System Model Version 5 (CanESM5; Swart et al., 2019), the Institut Pierre-Simon Laplace Coupled Model 6 (IPSL-CM6; Boucher et al., 2020), and the Community Earth System Model Version 2 (CESM2; Danabasoglu et al., 2020). We analyze 25 ensemble members of CanESM5, 31 ensemble members of IPSL-CM6, and 11 ensemble members of CESM2. These simulation output were

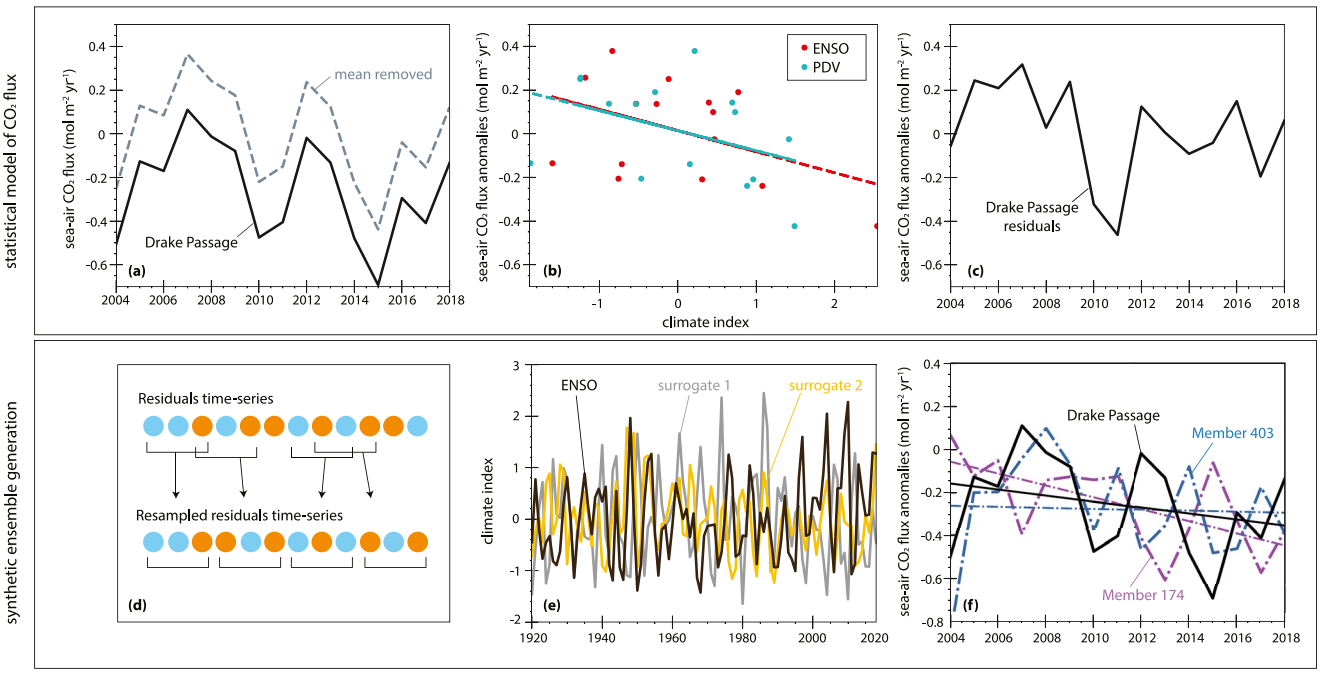


**Figure 1.** Synthetic ensemble construction. (a–c) Statistical model of annual mean Drake Passage sea-air $CO_2$ flux, as in Equation 1: (a) (solid) Time-series of $CO_2$ flux (black, $X^{i,t}$) and (dashed) $CO_2$ flux with temporal mean $\left(\beta_0^i\right)$ removed, (b) regression of $CO_2$ flux anomalies (temporal mean and response to external forcing (a simple linear trend in this example), $\beta_F{}^t$, removed) onto the El Niño-Southern Oscillation (ENSO) (red) and Pacific Decadal Variability (PDV) (blue) climate indices ($\beta^i_{\mathrm{ENSO}}$, $\beta^i_{\mathrm{PDV}_\perp}$), and (c) residual variability, $\varepsilon^{i,t}$. (d–f) Construction of the synthetic ensemble: (d) The block bootstrap process re-samples the residual variability, $\varepsilon^{i,t}$, (e) the IAFFT technique produces surrogate ENSO and $\mathrm{PDV}_\perp$ indices (ENSO shown here), and (f) two synthetic ensemble members show alternative phasing of internal variability and different long-term trends (dashed) than the original time-series (solid black line same as in a). Positive fluxes correspond to decreased oceanic carbon uptake. Panel (d) adapted from Elsworth et al. (2020).

derived from concentration-driven simulations (that is, their atmospheric $CO_2$ concentrations were prescribed and were thus unaltered by variability in sea-air $CO_2$ fluxes; Eyring et al., 2016). Our use of only three Earth system model ensemble means for this study has the potential to bias our estimate of external forcing. For this reason, we also test our approach using an alternative external forcing model (see next subsection).

### 2.5. Upper Ocean Box Model

We use a theoretical upper-ocean box model created by McKinley et al. (2020) to produce an estimate of externally forced variations in sea-air $CO_2$ flux driven by variations in atmospheric $pCO_2$ and volcanic eruptions alone. The box model solves for the time change of dissolved inorganic carbon (DIC) in the upper ocean (surface area for flux excludes ice-covered regions, as in the observation-based products) and contains a simple representation of overturning circulation. The rate of flux is estimated using the approach in Wanninkhof (2014). The biological pump is assumed constant over time. The box model uses the same $xCO_2^{atmosphere}$ data as that used to estimate sea-air flux in the observation-based products. Temperature is held at a constant global surface ocean value except in the years following the eruptions of El Chichón and Mt. Pinatubo, where the upper ocean cools according to that simulated by the CESM1-LE. The box model is spun up from 1959 to 1979 using observed $pCO_2^{atmosphere}$, and then time-stepped with monthly resolution over 1980–2018 with the external forcing described above.

## 3. Synthetic Ensemble Construction

Before analyzing the global sea-air $CO_2$ flux from observation-based products, we first illustrate our synthetic ensemble approach for the reader using a single time-series of annual mean $CO_2$ flux derived from observations collected in Drake Passage (Figure 1). Our method is built upon the approach developed in McKinnon et al., 2017; McKinnon and Deser (2018). We statistically model sea-air $CO_2$ flux as:

$$X^{i,t} = \beta_0^i + \beta_F^t + \beta^i_{\mathrm{ENSO}} M^t_{\mathrm{ENSO}} + \beta^i_{\mathrm{PDV}_\perp} M^t_{\mathrm{PDV}_\perp} + \varepsilon^{i,t} \quad (1)$$

where $X^{i,t}$ is the sea-air $CO_2$ flux at location $i$, representing a single grid cell, a regional integral or a global integral, and time $t$, resolved annually. In this model, sea-air $CO_2$ flux is described as a linear combination of the mean state $\beta_0{}^i$, the response to external forcing $\beta_F{}^t$ (which we assume to be spatially uniform globally), the response to climate modes $\beta^i_{\mathrm{ENSO}} M^t_{\mathrm{ENSO}}$ and $\beta^i_{\mathrm{PDV}_\perp} M^t_{\mathrm{PDV}_\perp}$, and the residual internal variability $\varepsilon^{i,t}$. The term $\beta_F{}^t$ in Equation 1 captures the response of $CO_2$ flux to external forcing, while $\beta^i_{\mathrm{ENSO}} M^t_{\mathrm{ENSO}}$ and $\beta^i_{\mathrm{PDV}_\perp} M^t_{\mathrm{PDV}_\perp}$ capture the role of these climate modes in sea-air $CO_2$ flux. Both ENSO and Pacific Decadal Variability (PDV) have been shown to influence sea-air $CO_2$ flux on global scales (McKinley et al., 2004, 2006, 2017). We address the covariance between ENSO and PDV by creating a time-series of PDV ($\mathrm{PDV}_\perp$) that is orthogonalized with respect to ENSO (method described in McKinnon & Deser, 2018).

Figure 1 (top row) illustrates our statistical model for Drake Passage $CO_2$ flux, as in Equation 1. Figure 1a shows the annual mean flux in this region over 2004 to 2018 ($X^{i,t}$) as a solid line, and anomalies in the flux once the the temporal mean flux $\left(\beta_0{}^i\right)$ has been subtracted as a dashed line. In this illustrative example, we model the external forcing $\left(\beta_F{}^t\right)$ as a simple linear trend (note that we model external forcing differently for the four global observation-based products, discussed later in this section). We model the influence of climate modes on sea-air $CO_2$ flux variability by calculating the linear regression between globally integrated $CO_2$ flux and the standardized indices for ENSO and orthogonalized PDV ($\beta^i_{\mathrm{ENSO}}$, $\beta^i_{\mathrm{PDV}_\perp}$; Figure 1b). The $CO_2$ flux residuals ($\varepsilon^{i,t}$) are modeled as the component of $X^{i,t}$ that is not captured by the external forcing or internal climate modes. As there is only a small linear correlation between sea-air $CO_2$ flux and the ENSO and PDV indices (Figure 1b), these residuals capture a large fraction of the internal variability (95%).

Figure 1 (bottom row) illustrates how we construct a synthetic ensemble from our statistical model of Drake Passage $CO_2$ flux. We use block bootstrapping to re-sample the residuals ($\varepsilon^{i,t}$) 1,000 times (Figure 1d). As described in Wilks (1997), we use a block length of 3 years to ensure that the time blocks are suitably large relative to the autocorrelation timescale, while also producing enough time blocks to generate sufficient variability between samples (Wilks, 1997; Equation 19). Block bootstrapping selects any contiguous 3-year block of sea-air $CO_2$ flux from the anomaly time-series and randomly samples these blocks with replacement to generate a new time-series with the same length as the original. Note that this approach maintains the temporal relationship within each block (Figure 1d). Next, we use the Iterative Amplitude Adjustment Fourier Transfer (IAAFT) technique (Schreiber & Schmitz, 1996, 2000) to produce 1,000 surrogate ENSO and PDV indices with similar spectral characteristics as the original climate indices (Figure 1e). For example, an IAFFT-generated surrogate ENSO index will exhibit a spectral peak in the 3- to 7-year time window, as the observed ENSO index does, but the amplitude and phasing of surrogate ENSO events will differ from the observed ENSO index. As described in McKinnon and Deser (2021), the IAAFT method produces surrogate time series that differ in their temporal evolution but share the same amplitude distribution and Fourier spectra as the original time series. We produce 1,000 unique synthetic ensemble members of Drake Passage $CO_2$ flux (2 members shown in Figure 1f) by combining the re-sampled residuals ($\varepsilon^{i,t}$), the $CO_2$ flux evolution due to the surrogate climate modes ($\beta^i_{\mathrm{ENSO}} M^t_{\mathrm{ENSO}}$ and $\beta^i_{\mathrm{PDV}_\perp} M^t_{\mathrm{PDV}_\perp}$), the external forcing $\left(\beta_F{}^t\right)$, and the temporal mean flux $\left(\beta_0{}^i\right)$. This technique produces 1,000 "alternative histories" of sea-air $CO_2$ flux in this region.

Figure 1f shows the temporal evolution of Drake Passage $CO_2$ flux from two synthetic ensemble members and the original observations over 2004–2018. Each synthetic ensemble member has statistical properties that are similar to the observational record and an identical externally forced signal, but a unique sequence of internal variability. Here, we see the clear influence of internal variability on the long-term trend: different phasing of internal variability in sea-air $CO_2$ flux between members is substantial enough to drive different estimates of the long-term trend (Figure 1f). The effect of internal variability on long-term trends is especially pronounced over the relatively short time period and at the regional scale of the Drake Passage observations (Hawkins & Sutton, 2009). While the observed $CO_2$ flux and synthetic ensemble members exhibit negative trends (more ocean carbon absorption with time), ensemble member 174 exhibits a much larger negative trend than the others over the same period. This outcome emphasizes the importance of internal variability in modifying the long-term trends in sea-air $CO_2$ fluxes in this region.

We use our statistical emulation technique to develop synthetic ensembles of globally and regionally integrated sea-air $CO_2$ flux for each of the observation-based products (CSIR-ML6, MPI-SOMFFN, JENA-MLS, and CMEMS-FFNN; Figure 2) and for the average of the four observation-based products. Our approach is identical to that described for the Drake Passage time-series, with the exception of our model for the externally forced

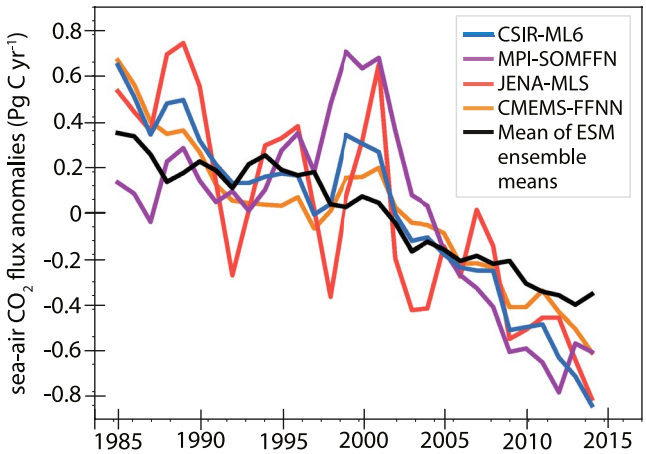


**Figure 2.** Global $CO_2$ flux variations. Temporal evolution of globally integrated sea-air $CO_2$ flux anomalies (temporal mean, $\beta_0^i$, removed) from the (blue) Council for Scientific and Industrial Research-Machine Learning ensemble (CSIR-ML6), (purple) Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN), (pink) Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS), and (orange) Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN) observation-based products. Black line shows the global $CO_2$ flux response to external forcing, $\beta_F{}^t$, estimated as the mean of three ensemble means from Earth system model (ESM) output contributed to the sixth Coupled Model Intercomparison Project (CMIP6) archive (CanESM5, IPSL-CM6, CESM2). Positive flux anomalies correspond to decreased oceanic carbon uptake.

signal $\left(\beta_F{}^t\right)$. Here, we model $\beta_F{}^t$ as the mean of three ensemble mean $CO_2$ flux estimates from historical simulations of CMIP6 Earth system models (Figure 2; see Section 2.4). As sea-air $CO_2$ flux is sensitive to variations in atmospheric $pCO_2$ and short-term volcanic forcing (McKinley et al., 2020), we use the ensemble mean time-series of $CO_2$ flux from Earth system models that are driven by these external forcing variations to isolate the temporal evolution of the forced signal ($\beta_F{}^t$; McKinley et al., 2016). This allows us to generate synthetic ensemble members that differ due to internal variability, rather than anthropogenic and natural external forcing. We further account for differences in model structure for estimation of the forced signal by averaging across ensemble means from three different Earth system models. In Section 5, we explore the sensitivity of our results to the statistical model of the externally forced signal.

## 4. Results

The synthetic ensemble of globally integrated sea-air $CO_2$ flux from the four observation-based products reveal multiple possible trajectories for the temporal evolution of ocean carbon uptake (Figure 3). While the ensemble mean trend is negative over 1985–2014 (increased ocean carbon absorption with time and driven by external forcing), different phasing of internal variability produces alternative $CO_2$ flux evolution across the synthetic ensemble (Figure 3). In Figure 3, we highlight the $CO_2$ flux from the original products in yellow and a single synthetic ensemble member in black (with the remaining 999 synthetic members as thinner, multi-hued lines) for each product to illustrate how the observed temporal evolution of $CO_2$ flux may not be replicated by the synthetic ensemble member, and the observed long-term trend may be amplified or muted in the synthetic ensemble member. This showcases the utility of the synthetic ensemble for quantifying the effects of internal variability on particular features of the time-series and the long-term trend.

The synthetic ensembles of globally integrated $CO_2$ flux from the four observation-based products display statistical properties that are different for each product (Figure 3). While all four ensembles show a long-term negative ensemble mean trend (increased ocean carbon absorption with time), the average ensemble spread ranges from 0.13 Pg C $yr^{-1}$ (1$\sigma$, CMEMS-FFNN) to 0.27 Pg C $yr^{-1}$ (1$\sigma$, JENA-MLS). This synthetic ensemble spread derives from the variance in the original observation-based product (Figure 2), and so it is not surprising that the product with the highest variance (JENA-MLS) exhibits the largest synthetic ensemble spread (Figure 3c), while the product with the lowest variance (CMEMS-FFNN) exhibits the lowest synthetic ensemble spread (Figure 3d).

If internal variability had been phased differently in the past, would we have observed the same decadal trends in sea-air $CO_2$ flux? We answer this question by analyzing the statistical properties of linear $CO_2$ flux trends over 1990–1999 and 2000–2009 from the four synthetic ensembles and displaying the results as probability density functions (PDFs; Figure 4). These decades were selected for analysis as they are associated with stagnation and growth of the ocean carbon sink, respectively, in several previous studies (see, for example, Ritter et al., 2017). The forced trend is nearly identical in each decade (−0.05 Pg C $yr^{-1}$ for both decades), and thus the differences in the observation-based product trends in each decade are due to internal variability. The PDFs in Figure 4 show the distributions of trends over the two decades in globally integrated $CO_2$ flux from 1,000 synthetic ensemble members of each observation-based product (purple line; kernel density estimation), with the 1$\sigma$ (67%) confidence intervals shaded in purple and the 2$\sigma$ (95%) confidence intervals shaded in pink. The width of these trend distributions vary across products, with MPI-SOMFFN exhibiting the widest distribution and CMEMS-FFNN exhibiting the narrowest (Table 2); for MPI-SOMFFN, internal variability alone can produce a wide range of trends for globally integrated flux (nearly 0.2 Pg C $yr^{-2}$ in a single decade, Table 2). For CMEMS-FFNN, the range of trends is nearly half of MPI-SOMFFN (~0.1 Pg C $yr^{-2}$ in a single decade, Table 2). The answer to the question posed at the beginning of this paragraph requires not only information about the width of the trend distributions, but also information about the center of the trend distributions. Our approach assumes that the center

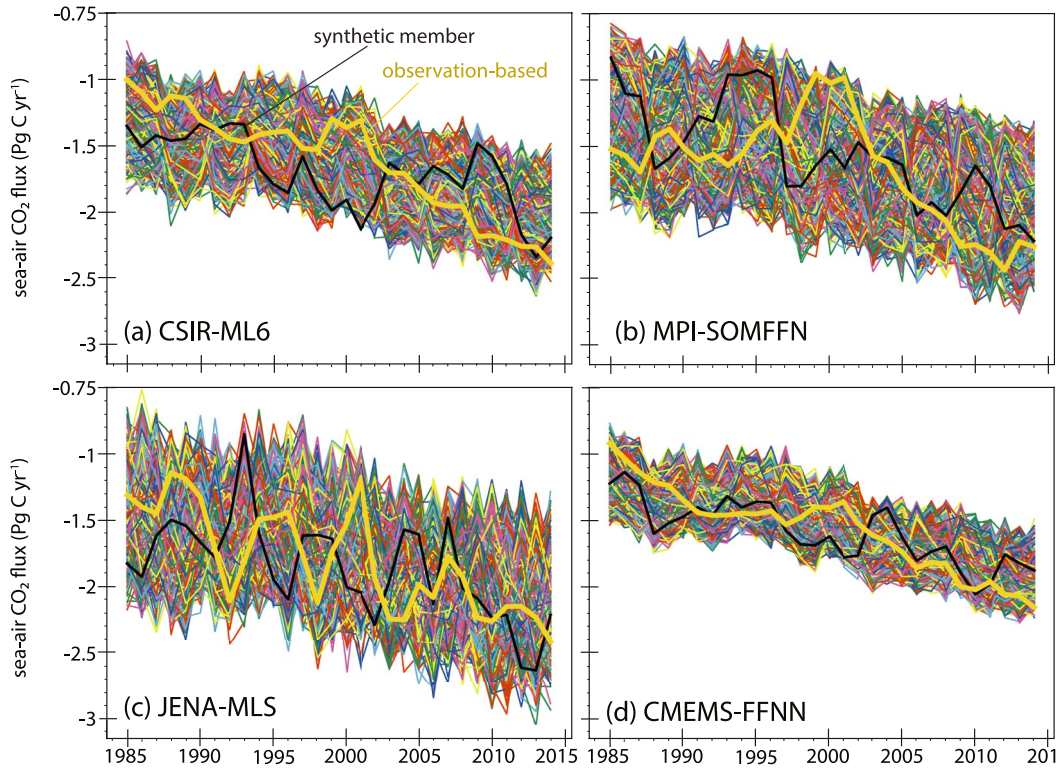


**Figure 3.** Synthetic ensembles of global sea-air $CO_2$ flux. Temporal evolution of globally integrated sea-air $CO_2$ flux from 1,000-member synthetic ensembles of the (a) Council for Scientific and Industrial Research-Machine Learning ensemble (CSIR-ML6), (b) Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN), (c) Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS), and (d) Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN) observation-based products. Yellow lines show the $CO_2$ flux evolution from the given observation-based product, and black line shows the temporal evolution of a single ensemble member with the remaining 999 members shown in thin multi-hued lines. Negative fluxes correspond to ocean carbon uptake.

of the trend distribution (vertical purple lines in Figure 4) is the mean of three Earth system model ensemble means and is thus identical for all of the observation-based synthetic ensembles in each time period (we examine this assumption in further detail in Section 5). Next, we quantify the probability of the observed decadal trends (vertical black lines in Figure 4) in the context of the synthetic ensemble trend distribution for each observation-based product (Table 2; trend probabilities estimated as the lower/upper cumulative distribution for a normal distribution). The observed trend in globally integrated $CO_2$ flux over 1990–1999 is a small negative number (more ocean carbon uptake with time) that falls within the $1\sigma$ range around the mean in the synthetic ensembles of three of the four products (CSIR-ML6, JENA-MLS, and CMEMS-FFNN) (Figures 4a–4d and Table 2). Whereas the observed trend in MPI-SOMFFN over 1990–1999 is a positive number (less ocean carbon uptake with time) that falls outside the $1\sigma$ range around the mean in the distribution of synthetic trends (Figures 4b and Table 2). Over 2000–2009, three of four observation-based products (CSIR-ML6, MPI-SOMFFN, and CMEMS-FFNN) exhibit large negative trends that are in the tails of the synthetic trend distributions (Figures 4e–4h and Table 2). Thus, the answer to the question we posed at the beginning of this paragraph is product and period dependent. Over 1990–1999, three of the four products have trends that are within one standard deviation from the mean in their respective synthetic ensembles, but over 2000–2009, three of the four products have trends that are nearly two standard deviations away from the mean with different phasing of internal variability, with only one observed trend (JENA-MLS) falling near the center of the distribution (Figure 4g and Table 2). For MPI-SOMFFN, however, the observed trends fall on the tails of the PDFs for both decades. This aligns with Gloege et al. (2021), who find that the MPI-SOMFFN product overestimates the amplitude of global $CO_2$ flux variability on decadal timescales. A synthetic ensemble generated from the average of all four observation-based products produces a relatively narrow synthetic trend distribution (0.11 Pg C $yr^{-2}$) and observed trends fall within

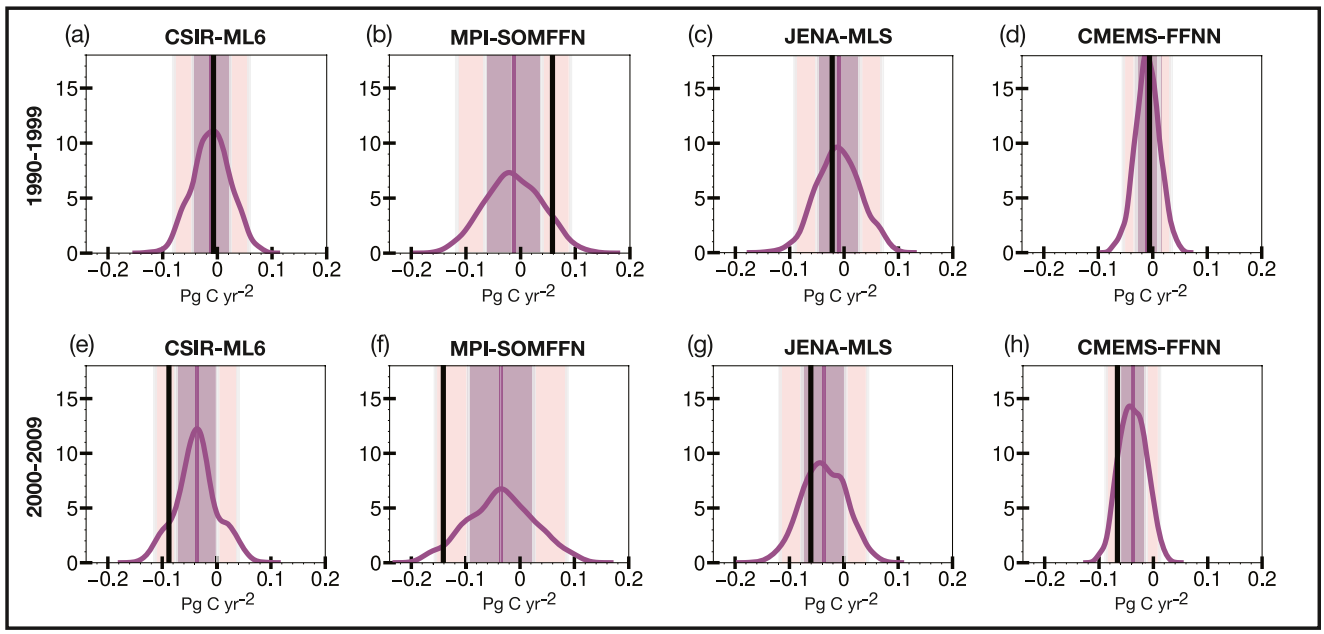


**Figure 4.** Distribution of decadal trends in global $CO_2$ flux derived from the synthetic ensembles. Probability density functions (kernel density estimation, purple curves) of decadal trends in globally integrated sea-air $CO_2$ flux for (first row) 1990–1999 and (second row) 2000–2009, as estimated from synthetic ensembles of the (first column) Council for Scientific and Industrial Research-Machine Learning ensemble (CSIR-ML6), (second column) Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN), (third column) Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS), and (fourth column) Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN) observation-based products. Purple vertical lines show the ensemble mean trend, and the $1\sigma$ (67%) and $2\sigma$ (95%) confidence intervals are shaded in purple and pink respectively. Black lines show the observed decadal trend from each product. Negative trends correspond to increased ocean carbon uptake with time.

$1\sigma$ in the synthetic ensemble in 1990–1999 and within $2\sigma$ in 2000–2009 (Figure S1 in Supporting Information S1; Table 2).

We create probability distributions for the range of decadal trends in regional, observation-based sea-air $CO_2$ flux over 1990–1999 and 2000–2009 by generating synthetic ensembles of $CO_2$ flux integrated over "super-biomes", that is, biomes that capture large-scale oceanographic regions (Canadell et al., 2021), and performing similar statistical analyses as for the globally integrated fluxes (Figure 5 and Figures S2–S5 in Supporting Information S1). We focus our discussion here on the Southern Ocean region, as previous work suggests large, opposite-signed decadal $CO_2$ flux trends in this region across the two decades of interest (Landschützer et al., 2015; Le Quéré et al., 2007; Lovenduski et al., 2008; Ritter et al., 2017). The observation-based synthetic ensembles of sea-air $CO_2$ flux integrated over the Southern Ocean Ice, Subpolar and Subtropical Seasonally Stratified biomes (SO ICE, SO-SPSS, SO-STSS; biomes defined in Fay & McKinley, 2014) produce mostly narrow PDFs of the decadal trends over 1990–1999 and 2000–2009 (MPI-SOMFFN is nearly double the other products) (Figure 5; all 4 observation-based products were masked identically to ensure consistency in a region where areas might be missed or miscalculated due to sea-ice). The externally forced trend (mean of Earth system model ensemble means) in the Southern Ocean is negative for both decades (more Southern Ocean carbon absorption with time, purple vertical lines in Figure 5). The 95% confidence interval of the synthetic trends ranges from −0.04 to 0.03 Pg C $yr^{-2}$ in each decade across the products (Figure 5), suggesting that both negative and positive trends are possible with different phasing of internal variability in both decades. However, the observed Southern Ocean flux trends range from negative to positive and do not always fall within $1\sigma$ of the PDFs (black vertical lines in Figure 5). For example, observed trends in the CSIR-ML6 and MPI-SOMFFN products over both decades fall outside the $1\sigma$ confidence interval of the PDFs, indicating a low chance of occurrence of the observed trend in the synthetic ensemble given different phasing of internal variability (Figures 5a, 5b, 5e and 5f). We note the remarkable consistency of all four observation-based products that agree the observed trends (black lines in Figure 5) were more positive than the forced trend (purple vertical lines in Figure 5) during the 1990s, and more negative than the forced trend during the 2000s.

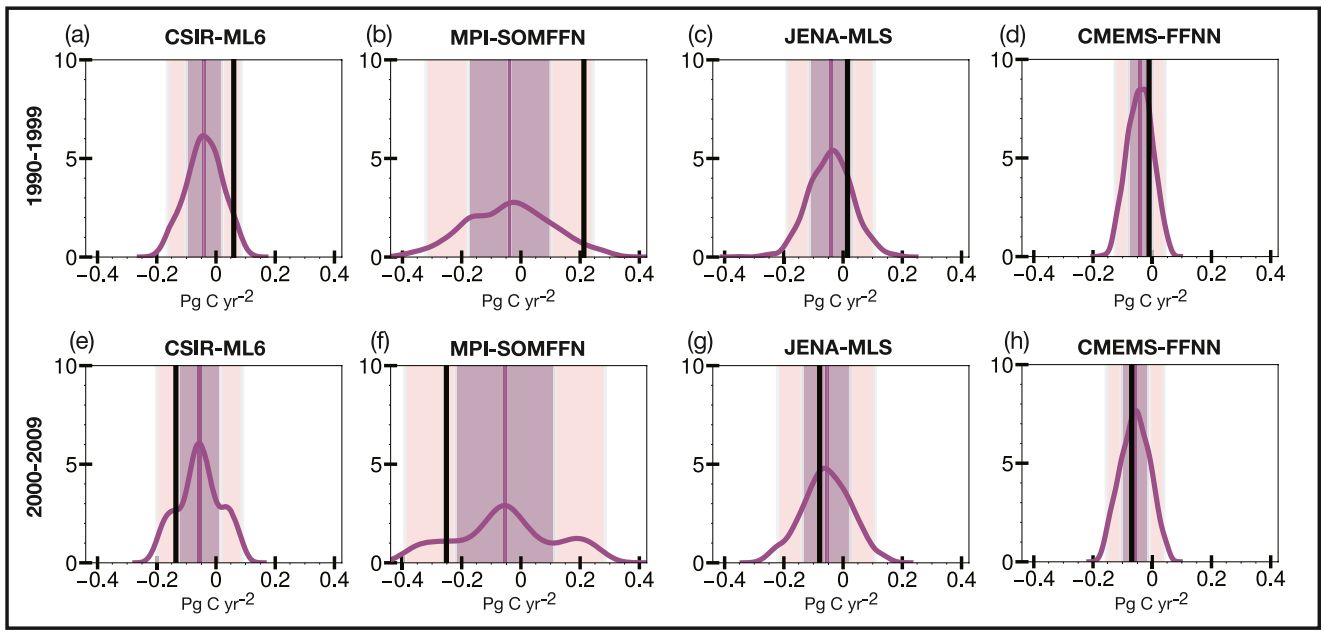


**Figure 5.** Distribution of decadal trends in Southern Ocean $CO_2$ flux derived from the synthetic ensembles. As in Figure 4, but for the $CO_2$ flux integrated over the Southern hemisphere super-biome, made up of Southern Ocean Ice, Subpolar and Subtropical Seasonally Stratified biomes (biomes defined in Fay & McKinley, 2014).

The observed trends and the distributions of synthetic trends in other super-biomes over 1990–1999 and 2000–2009 are shown in Figures S2–S5 in Supporting Information S1 and briefly described here. In the Northern hemisphere high latitude super-biome, the distribution of synthetic trends is very broad (−0.4 to 0.4 Pg C $yr^{-2}$; minimum and maximum values of PDFs) and, like the global fluxes, the observed trends of all observation-based

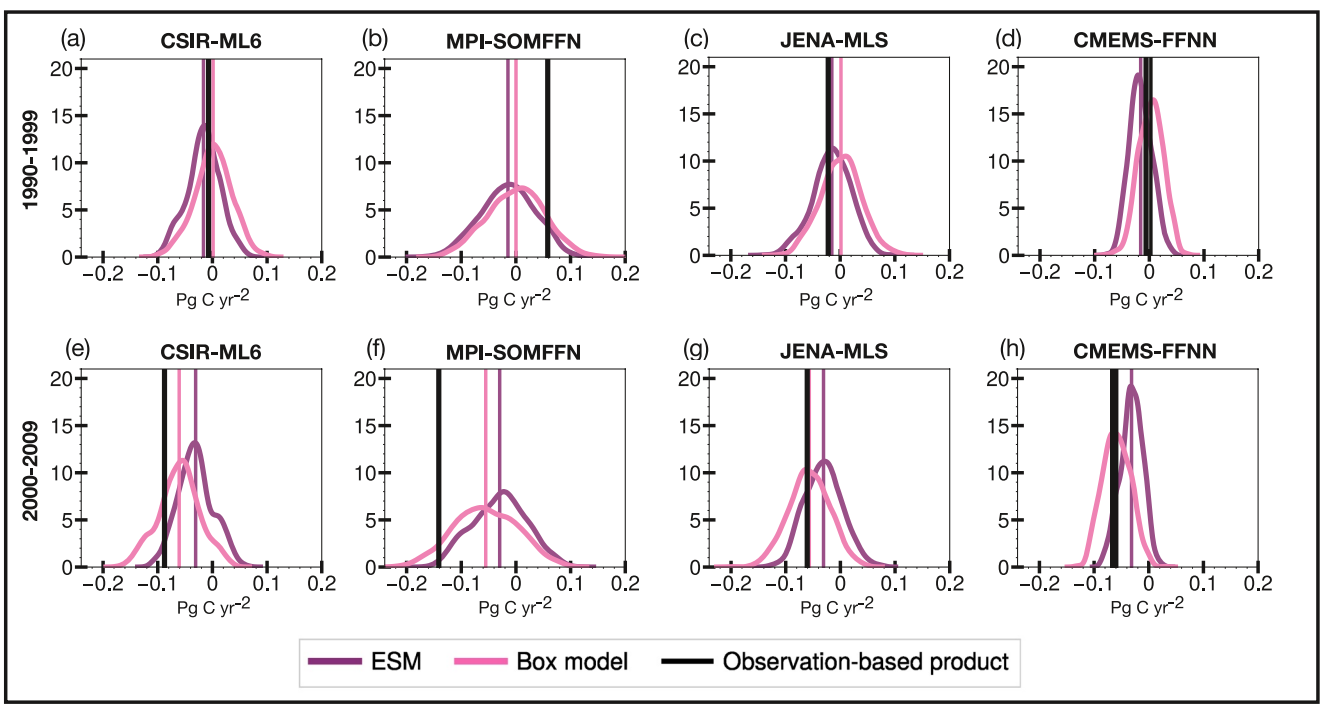


**Figure 6.** Synthetic ensemble of global $CO_2$ flux using Community Earth System Model Version 1 Large Ensemble (CESM1-LE). Temporal evolution of the globally integrated sea-air $CO_2$ flux from the (pastel colors) CESM1-LE and (gray) synthetic ensemble of CESM1-LE member 21 (shown in black). Negative fluxes correspond to ocean carbon uptake.

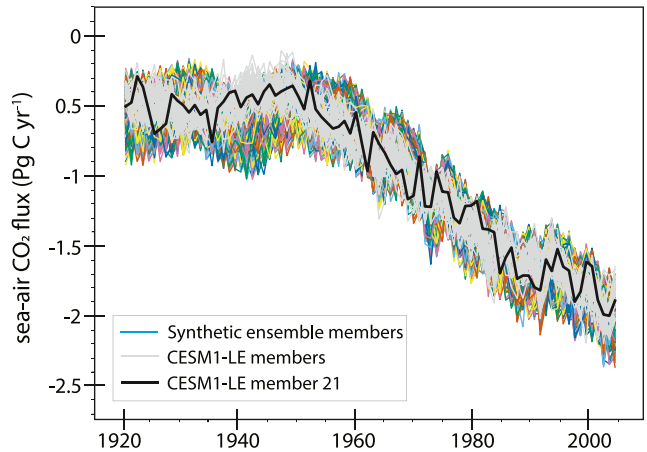


**Figure 7.** Distribution of decadal trends in global $CO_2$ flux with alternative model of external forcing. Probability density functions (kernel density estimation, purple and pink curves) of decadal trends in globally integrated sea-air $CO_2$ flux for (first row) 1990–1999 and (second row) 2000–2009, as estimated from synthetic ensembles of the (first column) Council for Scientific and Industrial Research-Machine Learning ensemble (CSIR-ML6), (second column) Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN), (third column) Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS), and (fourth column) Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN) observation-based products. Purple curves show probability density when the external signal is derived from the mean of three ensemble means from Earth system model (ESM) output submitted to the sixth Coupled Model Intercomparison Project (CMIP6) archive. Pink curves show probability density when the external signal is derived from the McKinley et al. (2020) upper ocean box model. Pink and purple vertical lines indicate the ensemble mean trend, and black vertical lines show the observed decadal trend from each product.

products over 2000–2009 fall within the $2\sigma$ range around the mean (Figure S2 in Supporting Information S1). The subtropical super-biomes in the Northern and Southern hemispheres exhibit narrow distributions of synthetic trends, due to lower interannual variability in $CO_2$ flux in these regions (−0.2 to 0.1 in the both hemispheres, Figures S3 and S4 in Supporting Information S1). The Equatorial super-biome synthetic ensemble produces a wide distribution of decadal trends over both periods (−0.5 to 0.4 Pg C $yr^{-2}$) and all observation-based product trends fall near the mean or within $1\sigma$ (Figure S5 in Supporting Information S1). Thus, the analysis of our super-biome synthetic ensembles suggests that our findings are also regionally dependent.

## 5. Testing Our Approach and Assumptions

We now consult the CESM1-LE as a testbed, where we apply our synthetic ensemble method to a single ensemble member to see if we reproduce the spread across the full ensemble. Using the globally integrated $CO_2$ flux from a single ensemble member of CESM1-LE, our statistical emulation technique generates a synthetic ensemble of $CO_2$ flux that is similar to the true CESM1-LE (Figure 6). We model the statistical properties of a particular CESM1-LE ensemble member (in this case, member 21, black line in Figure 6) using Equation 1 with the external signal $\left(\beta_F{}^t\right)$ modeled as the mean of three ensemble mean $CO_2$ flux estimates from historical simulations of CMIP6 Earth system models (see Section 2.4; note that a different external signal produces different results as discussed later in

**Table 1**
*Observation-Based Products of Sea-Air $CO_2$ Flux Used in This Study*

| Observation-based product | Abbreviation | Reference |
|---|---|---|
| Council for Scientific and Industrial Research-Machine Learning ensemble | CSIR-ML6 | Gregor et al. (2019) |
| Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network | MPI-SOMFFN | Landschützer et al. (2013, 2014, 2015, 2016) |
| Jena, Germany, Max Planck Institute for Biogeochemistry - Mixed Layer Scheme | JENA-MLS | Rödenbeck et al. (2014) |
| Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network | CMEMS-FFNN | Denvil-Sommer et al. (2019) |

*Note*. See Section 2.2 for details.

**Table 2**
*Width of Trend Distributions and Trend Percentiles for Synthetic Ensembles of Globally Integrated $CO_2$ Flux Produced From the Council for Scientific and Industrial Research-Machine Learning Ensemble (CSIR-ML6), Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network (MPI-SOMFFN), Max Planck Institute for Biogeochemistry-Mixed Layer Scheme (JENA-MLS), and Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network (CMEMS-FFNN) Observation-Based Products*

| Observation-based product | Width of trend distribution 1990–1999 | Width of trend distribution 2000–2009 | Percentile of observed trend within synthetic ensemble distribution 1990–1999 | Percentile of observed trend within synthetic ensemble distribution 2000–2009 |
|---|---|---|---|---|
| CSIR-ML6 | 0.13 | 0.13 | 64% | 19% |
| MPI-SOMFFN | 0.19 | 0.19 | 25% | 17% |
| JENA-MLS | 0.14 | 0.14 | 44% | 21% |
| CMEMS-FFNN | 0.08 | 0.09 | 72% | 20% |
| Average of all observation-based products | 0.11 | 0.11 | 22% | 15% |

*Note*. Widths estimated as $4\sigma$ (Pg C $yr^{-2}$).

this section), and the climate modes ($M^t_{ENSO}$ and $M^t_{PDV_\perp}$) produced from CESM1 for this ensemble member. We then generate a 1,000-member synthetic ensemble of ensemble member 21 using the steps described previously. Figure 6 illustrates the resulting synthetic ensemble in pastel colors overlain on the original CESM1-LE ensemble in gray. The envelope of variability in our synthetic ensemble is larger than that of the full CESM1-LE (Figure 6), as the average standard deviation (trend removed) of the synthetic ensemble is 11% different from that of full CESM1-LE over 1920–2005 (Table 3). We generate a synthetic ensemble of globally integrated sea-air $CO_2$ flux for each of the 34 CESM1-LE ensemble members and display the quotient of the two standard deviations (synthetic ensemble divided by the CMIP6 Earth system model mean of three ensemble means; Section 2.4) averaged over 1920–2005 in Table 3. This analysis reveals that the synthetic ensemble can overestimate the standard deviation by as much as 69% (ensemble member 8) depending upon the statistical properties of the original time-series used to generate the synthetic ensemble. On average across the 34 ensemble members, the overestimation bias in the standard deviation is 34% (Table 3). If, instead of modeling external forcing as the mean of CMIP6 model means, we instead use the CESM1-LE ensemble mean, the standard deviation bias averaged across all ensemble members is 4%. Thus, results from this analysis suggest that our statistical emulation technique for synthetic ensemble generation is reasonably unbiased and sensitive to the model of external forcing.

The length of the time-series used to generate the synthetic ensemble can have an influence on its statistical properties (Table 3). In this study, we generate a synthetic ensemble from observation-based products that are only 30 years long (1985–2014) and thus may not capture the full temporal spectrum of internal variability that occurs in the real world (McKinnon et al., 2017; McKinnon & Deser, 2018). We assess whether this shorter time-series can produce biased estimates of variance by generating synthetic ensembles of each CESM1-LE member over 1976–2005 (a 30-year period) and comparing their standard deviations to that of the full CESM1-LE over the same time period (Table 3). Interestingly, the synthetic ensembles generated from the shorter record produce smaller biases in the standard deviations of the synthetic ensembles than the synthetic ensembles generated from the longer record, with the average standard deviation bias as 15% (Table 3). If we instead use the CESM1-LE mean as the external forcing representation, the average standard deviation bias lessens to 2%.

The distributions of synthetic trends reported in the previous section are undoubtedly sensitive to the externally derived signal. Because the externally derived signal sets the mean value of the synthetic trend distribution, a different assumption about this signal can shift the distribution to the left/right and affect the placement of the observed trend within the synthetic ensemble distribution. Recall that our estimate of the externally forced signal is derived from the mean of three Earth system model ensemble means. McKinley et al. (2020) used an idealized upper-ocean box model to produce an estimate of externally forced variations in sea-air $CO_2$ flux driven by variations in atmospheric $pCO_2$ and volcanic eruptions alone. Figure 7 illustrates that the placement of observed trends within the synthetic ensemble distributions are similar regardless of whether we model the external forcing $\left(\beta_{F^t}\right)$ using the McKinley et al. (2020) box model or the mean of the three Earth system model ensemble means.

**Table 3**

*Standard Deviation Quotient (Synthetic Ensemble Standard Deviation Divided by External Forcing Representation). Column 1 Is the Sixth Coupled Model Intercomparison Project (CMIP6) Earth System Model Mean of Ensemble Means Standard Deviation and Column 2 Is the Community Earth System Model Version 1 Large Ensemble (CESM1-LE) Mean Standard Deviation for Synthetic Ensembles of Globally Integrated $CO_2$ Flux Produced From Each CESM1-LE Member for 1920–2005. Columns 4 (CMIP6 Earth System Model Mean of Ensemble Means) and 5 (CESM1-LE Mean) Are the Standard Deviation Quotients for Synthetic Ensemble Members for 1976–2005*

| Ensemble member | $\frac{\sigma_{SE}}{\sigma_{CMIP6}}$ 1920–2005 | $\frac{\sigma_{SE}}{\sigma_{CESM1-LE}}$ 1920–2005 | $\frac{\sigma_{SE}}{\sigma_{CMIP6}}$ 1976–2005 | $\frac{\sigma_{SE}}{\sigma_{CESM1-LE}}$ 1976–2005 |
|---|---|---|---|---|
| 1 | 1.54 | 1.14 | 1.22 | 1.08 |
| 2 | 1.17 | 0.91 | 1.26 | 1.10 |
| 3 | 1.45 | 1.22 | 0.94 | 0.83 |
| 4 | 1.20 | 0.99 | 1.02 | 0.90 |
| 5 | 1.53 | 1.21 | 1.05 | 0.92 |
| 6 | 1.33 | 0.95 | 1.01 | 0.88 |
| 7 | 1.32 | 0.92 | 0.78 | 0.69 |
| 8 | 1.69 | 1.40 | 1.21 | 1.07 |
| 9 | 1.28 | 1.08 | 1.36 | 1.18 |
| 10 | 1.31 | 1.02 | 1.38 | 1.21 |
| 11 | 1.29 | 1.18 | 1.27 | 1.10 |
| 12 | 1.26 | 1.06 | 1.38 | 1.22 |
| 13 | 1.21 | 0.80 | 0.88 | 0.78 |
| 14 | 1.33 | 0.97 | 1.20 | 1.05 |
| 15 | 1.32 | 1.03 | 1.29 | 1.14 |
| 16 | 1.15 | 1.04 | 0.95 | 0.84 |
| 17 | 1.20 | 0.93 | 1.06 | 0.93 |
| 18 | 1.31 | 1.03 | 1.27 | 1.10 |
| 19 | 1.45 | 0.90 | 0.88 | 0.78 |
| 20 | 1.26 | 0.96 | 1.07 | 0.94 |
| 21 | 1.11 | 0.99 | 1.07 | 0.93 |
| 22 | 1.47 | 1.10 | 1.43 | 1.26 |
| 23 | 1.29 | 1.08 | 1.09 | 0.94 |
| 24 | 1.62 | 1.19 | 1.34 | 1.17 |
| 25 | 1.32 | 1.04 | 1.18 | 1.02 |
| 26 | 1.39 | 0.99 | 1.06 | 0.94 |
| 27 | 1.19 | 0.88 | 0.83 | 0.72 |
| 28 | 1.45 | 1.10 | 1.77 | 1.54 |
| 29 | 1.40 | 1.09 | 1.41 | 1.24 |
| 30 | 1.48 | 1.06 | 1.15 | 1.00 |
| 31 | 1.46 | 1.06 | 1.17 | 1.03 |
| 32 | 1.44 | 1.11 | 1.16 | 1.03 |
| 33 | 1.44 | 1.30 | 1.14 | 0.99 |
| 34 | 1.31 | 0.89 | 1.06 | 0.93 |
| mean | 1.34 | 1.04 | 1.15 | 1.02 |

*Note*. The mean standard deviation across all ensemble members for two time periods is on the last line of the table.

## 6. Conclusions and Discussion

We develop synthetic large ensembles of global and regional sea-air $CO_2$ flux from four observation-based products using a statistical emulation technique. Much like a large initial condition ensemble of an Earth system model, the resulting synthetic ensemble members exhibit different phasing of internal variability and a common externally forced signal. We use these synthetic large ensembles to quantify the placement of decadal trends in $CO_2$ flux for each observation-based product within the synthetic ensemble trend distribution. We find that the phasing of internal variability creates unique features in the time-series of $CO_2$ flux and plays an important role in setting the multi-decadal trends in sea-air $CO_2$ flux for each synthetic ensemble member. The statistical properties of the synthetic large ensembles differ across the four observation-based products with JENA-MLS exhibiting the highest variance and CMEMS-FFNN exhibiting the lowest variance. Over the decade 1990–1999, three of the four products show negative observed trends in globally integrated sea-air $CO_2$ flux that fall within $1\sigma$ of the mean with different phasing of internal variability in the synthetic ensemble. However, over the decade 2000–2009, all four products show larger negative trends in sea-air $CO_2$ flux and three of the four products fall nearly two standard deviations away from the mean in the distribution of synthetic trends. The JENA-MLS product trends over these decades fall within $1\sigma$ of the mean within the synthetic ensemble, while the MPI-SOMFFN product trends over these decades fall outside $1\sigma$ range of the mean with different phasing of internal variability in the synthetic ensemble. In the Southern Ocean, CSIR-ML6 and MPI-SOMFFN exhibit large trends that fall outside the $1\sigma$ range of the mean with different phasing of internal variability in the synthetic ensemble. While the short length of the time-series used to construct the synthetic ensembles can bias the resulting statistical properties of the synthetic ensemble, the results of our study are similar whether we use an Earth system model or a box model to estimate the external signal, and are capable of producing robust estimates of the statistical properties when we construct the synthetic ensembles using longer time-series.

Our approach provides a new perspective on the important role of internal variability in short-term global and regional sea-air $CO_2$ flux trends estimated from the observational record. While we are not the first to demonstrate this point (see, for example, Fay & McKinley, 2013; McKinley et al., 2011, 2016), our synthetic ensembles provide visualizations of this variability. Further, the statistical properties of the synthetic ensembles provide a basis for examining the frequency of a particular global/regional trend in sea-air $CO_2$ flux given different phasing of internal variability over a particular time span and with no prior knowledge about the climate state. Finally, our work adds to the recent discussion about the role of internal versus external processes in interannual to decadal variations in sea-air $CO_2$ flux (DeVries et al., 2017; Landschützer et al., 2019; McKinley et al., 2020). While this study is successful in separating internal variability and external forcing, the external signal is an uncertain quantity. Regardless of how we model the externally forced signal, however, internal variability seems to play a key role in the distribution of decadal trends across our synthetic ensembles.

## Data Availability Statement

Our synthetic ensembles are available at https://doi.org/10.5281/zenodo.5227342. Underway DPT $pCO_2$ measurements are submitted to SOCAT and archived at NOAA's National Centers for Environmental Information (https://www.ncei.noaa.gov/access/ocean-carbon-data-system/oceans/VOS_Program/LM_gould.html). ENSO and PDV indices were provided by the NOAA Climate Prediction Center (https://origin.cpc.ncep.noaa.gov/). CESM ensemble output is available from the Earth System Grid (https://www.earthsystemgrid.org/dataset/ucar.cgd.ccsm4.CESM_CAM5_BGC_LE.html). CESM computing resources were provided by CISL at NCAR. CESM1-LE climate indices were calculated using the Climate Variability Diagnostics package (https://www.cesm.ucar.edu/working_groups/CVC/cvdp/). Coupled Model Intercomparison Project output was provided by the Earth System Grid Federation (https://esgf-node.ipsl.upmc.fr/projects/cmip6-ipsl/). Trend probabilities were calculated via the Keisan Online Calculator service provided by Casio Computer Co., Ltd. (https://keisan.casio.com/exec/system/1180573188).

**Acknowledgments**
The Surface Ocean $CO_2$ Atlas (SOCAT) is an international effort, supported by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biogeochemistry and Ecosystem Research program (IMBER), to deliver a uniformly quality-controlled surface ocean $CO_2$ database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT. The authors are grateful for the efforts of the marine and science support teams of the ARSV Laurence M. Gould, particularly Bruce Felix, Andy Nunn, and Kevin Pedigo. We are grateful for funding from the National Science Foundation (OCE-1558225, OCE-1752724, OCE-1948664, PLR-1543457, and the Graduate Research Fellowship Program). This work benefited from computational assistance from G. Elsworth, E. Maroon, and H. Zanowski. We are grateful to T. DeVries and an anonymous reviewer for providing feedback that helped to improve the manuscript.

## References

Anav, A., Friedlingstein, P., Kidston, M., Bopp, L., Ciais, P., Cox, P., et al. (2013). Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth System Models. *Journal of Climate*, *26*(18), 6801–6843. https://doi.org/10.1175/JCLI-D-12-00417.1

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., & Gombos, D. (2011). A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *Bulletin of the American Meteorological Society*, *92*(2), 157–174. https://doi.org/10.1175/2010BAMS2946.1

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., et al. (2016). A multi-decade record of high-quality $fCO_2$ data in version 3 of the Surface Ocean $CO_2$ Atlas (SOCAT). *Earth System Science Data*, *8*(2), 383–413. https://doi.org/10.5194/essd-8-383-2016

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002. https://doi.org/10.1029/2019MS002010

Canadell, J. G., Monteiro, P. M. S., Costa, M. H., da Cunha, L. C., Cox, P. M., Eliseev, A. V., et al. (2021). Global carbon and other biogeochemical cycles and feedbacks In V. P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, et al. (Eds.), *Climate change 2021: The physical science basis. Contribution of working group I to the sixth assessment Report of the intergovernmental Panel on climate change*.

Ciais, P., & Sabine, C. (2013). Chapter 6: Carbon and other biogeochemical cycles. In T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, et al. (Eds.) *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental Panel on climate change*. Cambridge University Press.

Danabasoglu, G., Bates, S. C., Briegleb, B. P., Jayne, S. R., Jochum, M., Large, W. G., et al. (2012). The CCSM4 ocean component. *Journal of Climate*, *25*(5), 1361–1389. https://doi.org/10.1175/jcli-d-11-00091.1

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The community Earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. https://doi.org/10.1029/2019MS001916

Denvil-Sommer, A., Gehlen, M., Vrac, M., & Mejia, C. (2019). LSCE-FFNN-v1: A two-step neural network model for the reconstruction of Surface Ocean $pCO_2^{oc}$ over the global ocean. *Geoscientific Model Development*, *12*(5), 2091–2105. https://doi.org/10.5194/gmd-12-2091-2019

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. https://doi.org/10.1038/s41558-020-0731-2

DeVries, T., Holzer, M., & Primeau, F. (2017). Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning. *Nature*, *542*(7640), 215–218. https://doi.org/10.1038/nature21068

DeVries, T., Le Quéré, C., Andrews, O., Berthet, S., Hauck, J., Ilyina, T., et al. (2019). Decadal trends in the ocean carbon sink. *Proceedings of the National Academy of Sciences*, *116*(24), 11646–11651. https://doi.org/10.1073/pnas.1900371116

Doney, S. C., Busch, D. S., Cooley, S. R., & Kroeker, K. J. (2020). The impacts of ocean acidification on marine ecosystems and reliant human communities. *Annual Review of Environment and Resources*, *45*(1), 83–112. https://doi.org/10.1146/annurev-environ-012320-083019

Dong, F., Li, Y., Wang, B., Huang, W., Shi, Y., & Dong, W. (2016). Global air–sea $CO_2$ flux in 22 CMIP5 models: Multiyear mean and interannual variability. *Journal of Climate*, *29*(7), 2407–2431. https://doi.org/10.1175/JCLI-D-14-00788.1

Elsworth, G. W., Lovenduski, N. S., McKinnon, K. A., Krumhardt, K. M., & Brady, R. X. (2020). Finding the fingerprint of anthropogenic climate change in marine phytoplankton abundance. *Current Climate Change Reports*, *6*(2), 37–46. https://doi.org/10.1007/s40641-020-00156-w

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Fay, A. R., Gregor, L., Landschützer, P., McKinley, G. A., Gruber, N., Gehlen, M., et al. (2021). Seaflux: Harmonization of air–sea $CO_2$ fluxes from surface $pCO_2$ data products using a standardized approach. *Earth System Science Data*, *13*(10), 4693–4710. https://doi.org/10.5194/essd-13-4693-2021

Fay, A. R., Lovenduski, N. S., McKinley, G. A., Munro, D. R., Sweeney, C., Gray, A. R., et al. (2018). Utilizing the Drake Passage time-series to understand variability and change in subpolar Southern Ocean $pCO_2$. *Biogeosciences*, *15*(12), 3841–3855. https://doi.org/10.5194/bg-15-3841-2018

Fay, A. R., & McKinley, G. A. (2013). Global trends in surface ocean $pCO_2$ from in situ data. *Global Biogeochemical Cycles*, *27*(2), 541–557. https://doi.org/10.1002/gbc.20051

Fay, A. R., & McKinley, G. A. (2014). Global open-ocean biomes: Mean and temporal variability. *Earth System Science Data*, *6*(2), 273–284. https://doi.org/10.5194/essd-6-273-2014

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C. E., Hauck, J., et al. (2021). Global carbon budget 2021. *Earth System Science Data Discussions*, 1–191. https://doi.org/10.5194/essd-2021-386

Gloege, L., McKinley, G. A., Landschützer, P., Fay, A. R., Frölicher, T. L., Fyfe, J. C., et al. (2021). Quantifying errors in observationally based estimates of ocean carbon sink variability. *Global Biogeochemical Cycles*, *35*(4), e2020GB006788. https://doi.org/10.1029/2020GB006788

Gregor, L., & Fay, A. (2021). SeaFlux: Harmonised Sea-air CO2 fluxes from surface pCO2 data products using a standardised approach, Zenodo. https://zenodo.org/record/5482547#.YpEitC-B30o

Gregor, L., Lebehot, A. D., Kok, S., & Scheel Monteiro, P. M. (2019). A comparative assessment of the uncertainties of global surface ocean $CO_2$ estimates using a machine-learning ensemble (CSIR-ML6 version 2019a)—Have we hit the wall? *Geoscientific Model Development*, *12*(12), 5113–5136. https://doi.org/10.5194/gmd-12-5113-2019

Hawkins, E., & Sutton, R. (2009). The potential to narrow uncertainty in regional climate predictions. *Bulletin America Meteorology Social*, *90*(8), 1095–1107. https://doi.org/10.1175/2009BAMS2607.1

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The community Earth system model: A framework for collaborative Research. *Bulletin America Meteorology Social*, *94*(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., et al. (2015). The community Earth system model (CESM) large ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bulletin America Meteorology Social*, *96*(8), 1333–1349. https://doi.org/10.1175/BAMS-D-13-00255.1

Landschützer, P., Gruber, N., & Bakker, D. C. E. (2016). Decadal variations and trends of the global ocean carbon sink. *Global Biogeochemical Cycles*, *30*(10), 1396–1417. https://doi.org/10.1002/2015GB005359

Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., et al. (2013). A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink. *Biogeosciences*, *10*(11), 7793–7815. https://doi.org/10.5194/bg-10-7793-2013

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C. E., van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224. https://doi.org/10.1126/science.aab2620

Landschützer, P., Ilyina, T., & Lovenduski, N. S. (2019). Detecting regional modes of variability in observation-based surface ocean $pCO_2$. *Geophysical Research Letters*, *46*(5), 2670–2679. https://doi.org/10.1029/2018GL081756

Landschützer, P., Laruelle, G. G., Roobaert, A., & Regnier, P. (2020). A uniform $pCO_2$ climatology combining open and coastal oceans. *Earth System Science Data*, *12*(4), 2537–2553. https://doi.org/10.5194/essd-12-2537-2020

Landschützer, P., Peter, N., Gruber, D. C. E. B., & Schuster, U. (2014). Recent variability of the global ocean carbon sink. *Global Biogeochemical Cycles*, *28*(9), 927–949. https://doi.org/10.1002/2014GB004853

Le Quéré, C., Andrew, R. M., Friedlingstein, P., Sitch, S., Pongratz, J., Manning, A. C., et al. (2018). Global carbon budget 2017. *Earth System Science Data*, *10*(1), 405–448. https://doi.org/10.5194/essd-10-405-2018

Le Quéré, C., Raupach, M. R., Canadell, J. G., Marland, G., Bopp, L., Ciais, P., et al. (2009). Trends in the sources and sinks of carbon dioxide. *Nature Geoscience*, *2*(12), 831–836. https://doi.org/10.1038/ngeo689

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., et al. (2007). Saturation of the Southern Ocean $CO_2$ sink due to recent climate change. *Science*, *316*(5832), 1735–1738. https://doi.org/10.1126/science.1136188

Lovenduski, N. S., Gruber, N., & Doney, S. C. (2008). Toward a mechanistic understanding of the decadal trends in the Southern Ocean carbon sink. *Global Biogeochemical Cycles*, *22*(3), GB3016. https://doi.org/10.1029/2007GB003139

Lovenduski, N. S., McKinley, G. A., Fay, A. R., Lindsay, K., & Long, M. C. (2016). Partitioning uncertainty in ocean carbon uptake projections: Internal variability, emission scenario, and model structure. *Global Biogeochemical Cycles*, *30*(9), 1276–1287. https://doi.org/10.1002/2016GB005426

McKinley, G. A., Fay, A. R., Eddebbar, Y. A., Gloege, L., & Lovenduski, N. S. (2020). External forcing explains recent decadal variability of the ocean carbon sink. *AGU Advances*, *1*(2), e2019AV000149. https://doi.org/10.1029/2019AV000149

McKinley, G. A., Fay, A. R., Lovenduski, N. S., & Pilcher, D. J. (2017). Natural variability and anthropogenic trends in the ocean carbon sink. *Annual Review of Marine Science*, *9*(1), 125–150. https://doi.org/10.1146/annurev-marine-010816-060529

McKinley, G. A., Fay, A. R., Takahashi, T., & Metzl, N. (2011). Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales. *Nature Geoscience*, *4*(9), 606–610. https://doi.org/10.1038/ngeo1193

McKinley, G. A., Follows, M. J., & Marshall, J. (2004). Mechanisms of air-sea $CO_2$ flux variability in the equatorial Pacific and the North Atlantic. *Global Biogeochemical Cycles*, *18*(2), C07S06. https://doi.org/10.1029/2003GB002179

McKinley, G. A., Pilcher, D. J., Fay, A. R., Lindsay, K., Long, M. C., & Lovenduski, N. S. (2016). Timescales for detection of trends in the ocean carbon sink. *Nature*, *530*(7591), 469–472. https://doi.org/10.1038/nature16958

McKinley, G. A., Takahashi, T., Buitenhuis, E., Chai, F., Christian, J. R., Doney, S. C., et al. (2006). North Pacific carbon cycle response to climate variability on seasonal to decadal timescales. *Journal of Geophysical Research*, *111*(C7), C07S06. https://doi.org/10.1029/2005JC003173

McKinnon, K. A., & Deser, C. (2018). Internal variability and regional climate trends in an observational large ensemble. *Journal of Climate*, *31*(17), 6783–6802. https://doi.org/10.1175/JCLI-D-17-0901.1

McKinnon, K. A., & Deser, C. (2021). The inherent uncertainty of precipitation variability, trends, and extremes due to internal variability, with implications for Western U.S. water resources. *Journal of Climate*, *34*(24), 9605–9622. https://doi.org/10.1175/JCLI-D-21-0251.1

McKinnon, K. A., Poppick, A., Dunn-Sigouin, E., & Deser, C. (2017). An "Observational Large Ensemble" to compare observed and modeled temperature trend uncertainty due to internal variability. *Journal of Climate*, *30*(19), 7585–7598. https://doi.org/10.1175/JCLI-D-16-0905.1

Metzl, N., Corbière, A., Reverdin, G., Lenton, A., Takahashi, T., Olsen, A., et al. (2010). Recent acceleration of the sea surface $fCO_2$ growth rate in the North Atlantic subpolar gyre (1993–2008) revealed by winter observations. *Global Biogeochemical Cycles*, *24*(4). https://doi.org/10.1029/2009GB003658

Moore, J. K., & Braucher, O. (2008). Sedimentary and mineral dust sources of dissolved iron to the world ocean. *Biogeosciences*, *5*(3), 631–656. https://doi.org/10.5194/bg-5-631-2008

Moore, J. K., & Doney, S. C. (2007). Iron availability limits the ocean nitrogen inventory stabilizing feedbacks between marine denitrification and nitrogen fixation. *Global Biogeochemical Cycles*, *21*(2). https://doi.org/10.1029/2006GB002762

Moore, J. K., Doney, S. C., & Lindsay, K. (2004). Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochemical Cycles*, *18*(4), GB4028. https://doi.org/10.1029/2004GB002220

Munro, D. R., Lovenduski, N. S., Stephens, B. B., Newberger, T., Arrigo, K. R., Takahashi, T., et al. (2015). Estimates of net community production in the Southern Ocean determined from time series observations (2002–2011) of nutrients, dissolved inorganic carbon, and surface ocean $pCO_2$ in Drake Passage. *Deep-Sea Res. II*, *114*(0), 49–63. https://doi.org/10.1016/j.dsr2.2014.12.014

Munro, D. R., Lovenduski, N. S., Takahashi, T., Stephens, B. B., Newberger, T., & Sweeney, C. (2015b). Recent evidence for a strengthening $CO_2$ sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002–2015). *Geophysical Research Letters*, *42*(18), 7623–7630. https://doi.org/10.1002/2015GL065194

NOAA Global Monitoring Laboratory (web) *NOAA greenhouse gas marine boundary layer reference*. Retrieved from https://gml.noaa.gov/ccgg/mbl/mbl.html

Ridge, S. M., & McKinley, G. A. (2021). Ocean carbon uptake under aggressive emission mitigation. *Biogeosciences*, *18*(8), 2711–2725. https://doi.org/10.5194/bg-18-2711-2021

Ritter, R., Landschützer, P., Gruber, N., Fay, A. R., Iida, Y., Jones, S., et al. (2017). Observation-based trends of the Southern Ocean carbon sink. *Geophysical Research Letters*, *44*(24), 12339–12348. https://doi.org/10.1002/2017GL074837

Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., et al. (2015). Data-based estimates of the ocean carbon sink variability—First results of the surface ocean $pCO_2$ mapping intercomparison (SOCOM). *Biogeosciences*, *12*(23), 7251–7278. https://doi.org/10.5194/bg-12-7251-2015

Rödenbeck, C., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., Cassar, N., et al. (2014). Interannual sea–air $CO_2$ flux variability from an observation-driven ocean mixed-layer scheme. *Biogeosciences*, *11*(17), 4599–4613. https://doi.org/10.5194/bg-11-4599-2014

Schreiber, T., & Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Physical Review Letters*, *77*(4), 635–638. https://doi.org/10.1103/PhysRevLett.77.635

Schreiber, T., & Schmitz, A. (2000). Surrogate time series. *Physica D: Nonlinear Phenomena*, *142*(3), 346–382. https://doi.org/10.1016/S0167-2789(00)00043-9

Schuster, U., Watson, A., Bates, N., Corbiere, A., Gonzalez-Davila, M., Metzl, N., et al. (2009). Trends in North Atlantic Sea surface $fCO_2$ from 1990 to 2006 North Atlantic Sea-surface fCO2 from 1990 to 2006. *Deep-Sea Res. II*, *56*(8–10), 620–629. https://doi.org/10.1016/j.dsr2.2008.12.011

Schuster, U., & Watson, A. J. (2007). A variable and decreasing sink for atmospheric $CO_2$ in the North Atlantic. *Journal of Geophysical Research*, *112*(C11), C11006. https://doi.org/10.1029/2006JC003941

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019). The Canadian Earth system model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Sweeney, C., Gloor, E., Jacobson, A., Key, R., McKinley, G., Sarmiento, J. L., & Wanninkhof, R. (2007). Constraining global air-sea gas exchange for $CO_2$ with recent bomb $^{1}4c$ measurements. *Global Biogeochemical Cycles*(GB2015), 21. https://doi.org/10.1029/2006gb002784

Takahashi, T., Sutherland, S., & Kozyr, A. (2018). *Global ocean surface water partial pressure of $CO_2$ database (LDEO database version 2019): Measurements performed during 1957-2019 (NCEI accession 0160492)*. NOAA National Centers for Environmental Information. Dataset. https://doi.org/10.3334/CDIAC/OTG.NDP088(V2015)

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin America Meteorology Social*, *93*(4), 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Thomas, H., Friederike Prowe, A. E., Lima, I. D., Doney, S. C., Wanninkhof, R., Greatbatch, R. J., et al. (2008). Changes in the North Atlantic oscillation influence $CO_2$ uptake in the North Atlantic over the past 2 decades. *Global Biogeochemical Cycles*, *22*(4), GB4027. https://doi.org/10.1029/2007GB003167

Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, *12*(6), 351–362. https://doi.org/10.4319/lom.2014.12.351

Watson, A. J., Schuster, U., Bakker, D. C. E., Bates, N. R., Corbière, A., González-Dávila, M., et al. (2009). Tracking the variable north Atlantic sink for Atmospheric $CO_2$. *Science*, *326*(5958), 1391–1393. https://doi.org/10.1126/science.1177394

Wilks, D. (1997). Resampling hypothesis tests for autocorrelated fields. *Journal of Climate*, *10*(1), 65–82. https://doi.org/10.1175/1520-0442(1997)010<0065:rhtfaf>2.0.co;2