Replaying the tape of history: Synthetic large ensembles of sea-air carbon dioxide (CO2) flux

Holly Olivarez[1], Nicole Lovenduski[1], Riley Brady[1], Amanda Fay[2], Marion Gehlen[3], Luke Gregor[4], Peter Landschützer[5], Galen McKinley[6], Karen McKinnon[7], and DAVID MUNRO[1]

[1]University of Colorado Boulder
[2]Lamont-Doherty Earth Observatory, Columbia University
[3]Laboratoire des sciences du climat et de l'environnement
[4]ETH Zürich Institute for Operations Research
[5]Universität Hamburg,Max Planck Institute for Meteorology,University of East Anglia
[6]Columbia University and Lamont-Doherty Earth Observatory
[7]University of California, Los Angeles

November 23, 2022

Abstract

We use a statistical emulation technique to construct synthetic ensembles of global and regional sea-air carbon dioxide (CO2) 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 CO2 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 likelihood of the observed trend given the distribution of synthetic trends during these two periods. Over the decade 1990-1999, three of the four observation-based products exhibit small negative trends in globally integrated sea-air CO2 flux (i.e., enhanced ocean CO2 absorption with time) that are highly probable (44-72% chance of occurrence) in their respective synthetic trend distributions. Over the decade 2000-2009, however, three of the four products show large negative trends in globally integrated sea-air CO2 flux that are somewhat improbable (17-19% chance of occurrence). Our synthetic ensembles suggest that the largest observation-based positive trends in global and Southern Ocean CO2 flux over 1990-1999 and the largest negative trends over 2000-2009 are somewhat improbable (<30% chance of occurrence). Our approach provides a new understanding of the role of internal and external processes in driving sea-air CO2 flux variability.

Hosted file

`agusupporting-information.docx` available at https://authorea.com/users/529539/articles/601262-replaying-the-tape-of-history-synthetic-large-ensembles-of-sea-air-carbon-dioxide-co2-flux

Replaying the tape of history: 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], David R. Munro[9]

[1]Environmental Studies Program, 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

Corresponding author: Holly C. Olivarez, `holly.olivarez@colorado.edu`

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 likelihood of the observed trend given the distribution of synthetic trends during these two periods. Over the decade 1990-1999, three of the 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 highly probable (44-72% chance of occurrence) in their respective synthetic trend distributions. Over the decade 2000-2009, however, three of the four products show large negative trends in globally integrated sea-air CO_2 flux that are somewhat improbable (17-19% chance of occurrence). Our synthetic ensembles suggest that the largest observation-based positive trends in global and Southern Ocean CO_2 flux over 1990-1999 and the largest negative trends over 2000-2009 are somewhat improbable (<30% chance of occurrence). 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 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 [*Friedlingstein et al.*, 2020]. 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 [*Ciais and Sabine*, 2013], though the magnitude of the globally integrated sea-air CO_2 flux will largely depend on our emissions trajectory [*Lovenduski et al.*, 2016; *Ridge and 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 ($p$$CO_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 [*Le Quéré et al.*, 2009; *Rödenbeck et al.*, 2015; *Landschützer et al.*, 2016; *DeVries et al.*, 2019], followed by intensification of ocean carbon absorption over the decade 2000-2009 [*Fay and McKinley*, 2013; *Rödenbeck et al.*, 2015; *Landschützer et al.*, 2016; *DeVries et al.*, 2019]. 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 [e.g., *Fay and McKinley*, 2013]. 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, sea-air

CO_2 flux variability in CESM1-LE and other Earth system models may not match that of the real world [*Hauck et al.*, 2020], 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 observation-based estimates of sea-air CO_2 flux into a large ensemble framework by constructing synthetic ensembles of observed sea-air CO_2 flux. 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.

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 $p\mathrm{CO_2}$ ($p\mathrm{CO}_2^{oc}$) 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 6^{th} Coupled Model Intercomparison Project (CMIP6). Finally, we use output from the CESM1-LE 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 $p\mathrm{CO}_2^{oc}$ collected as part of the Drake Passage Time-series program over 2004-2018 [Figure 1a; *Munro et al.*, 2015a,b; *Fay et al.*, 2018]. Each annual mean estimate of sea-air CO_2 flux is calculated from monthly means of all underway $p\mathrm{CO}_2^{oc}$ 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 sea-air CO_2 flux products

Our synthetic ensembles of sea-air CO_2 flux are derived from observation-based estimates that rely upon sparse $p\mathrm{CO}_2^{oc}$ 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 $p\mathrm{CO}_2^{oc}$ where and when measurements are not available (Table 1).

The Council for Scientific and Industrial Research-Machine Learning ensemble [CSIR-ML6; *Gregor et al.*, 2019] uses an ensemble of two-step neural network methods where two types of clusters and three types of regressions are used to interpolate $p\mathrm{CO}_2^{oc}$ from SOCAT v2020 using chlorophyll-a, sea surface temperature, absolute dynamic topography, mixed layer depth, sea ice, and sea surface salinity. The final product uses an ensemble average of six machine-learning models.

The Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network [MPI-SOMFFN; *Landschützer et al.*, 2013, 2014, 2015, 2016] uses a two-step neural network method to gap-fill $p\mathrm{CO}_2^{oc}$. In the first step, a self-organizing map is used to subdivide the ocean into 16 provinces with similar climatological biogeochemical properties. In the second step, a feed-forward neural network is used to predict the non-linear relationships between driver variables and SOCAT v2020 observations in each province. Driver variables for MPI-SOMFFN include sea surface temperature, mixed layer depth, satellite derived chlorophyll-a concentration, sea surface salinity, and atmospheric $p\mathrm{CO}_2$.

The Max Planck Institute for Biogeochemistry-Mixed Layer Scheme [JENA-MLS; *Rödenbeck et al.*, 2014] combines ocean mixed layer biogeochemistry with $p\mathrm{CO}_2^{oc}$ data from SOCAT v2020 and seasonal, interannual and short-term (daily) variations of sea surface temperature, mixed layer depth, ice-free fraction, salinity, wind speed, and alkalinity.

The Copernicus Marine Environment Monitoring Service Feed-Forward Neural Network [CMEMS-FFNN; *Denvil-Sommer et al.*, 2019] uses a two-step process that first reconstructs monthly climatologies of global $p\mathrm{CO}_2^{oc}$ from the LDEO database, and then reconstructs monthly anomalies using the SOCAT v5 grid. The driver variables used are chlorophyll-a, sea surface temperature, mixed layer depth, sea surface salinity, the atmospheric CO_2 mole fraction ($\chi\mathrm{CO}_2$), and sea surface height.

In this study, we use annual mean sea-air CO_2 flux estimates spanning the common observation-based product period of 1985 to 2014. As in *McKinley et al.* [2020] and *Fay et al.* [2021], we correct for the spatial coverage differences in each observation-based product 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, and wind speed inputs from the SeaFlux product [*Gregor and Fay*, 2021; *Fay et al.*, 2021] and a quadratic flux parameterization [*Ho et al.*, 2006; *Wanninkhof*, 2014]. This preprocessing ensures that differences in decadal trends or interannual variance in our synthetic ensemble are due solely to differences in the pCO$_2$ products rather than from statistical artifacts or flux calculation parameters.

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 and Doney*, 2007; *Moore and Braucher*, 2008]. 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 5^{th} Coupled Model Intercomparison Project [CMIP5; *Taylor et al.*, 2012; *Kay et al.*, 2015]. 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 derived from concentration-driven simulations [*i.e.*, their atmospheric CO_2 concentrations were prescribed and were thus unaltered by variability in sea-air CO_2 fluxes; *Eyring et al.*, 2016].

3 Synthetic ensemble construction

Before analyzing the global observation-based sea-air CO_2 flux 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] and *McKinnon and Deser* [2018]. We statistically model sea-air CO_2 flux as:

$$X^{i,t} = \beta_0^i + \beta_F^t + \beta_{\mathrm{ENSO}}^i M_{\mathrm{ENSO}}^t + \beta_{\mathrm{PDV}_\perp}^i M_{\mathrm{PDV}_\perp}^t + \varepsilon^{i,t} \tag{1}$$

where $X^{i,t}$ is the sea-air CO_2 flux at location i and time t. 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 and 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 (${\beta_0}^i$) has been subtracted as a dashed line. In this illustrative example, we model the external forcing (${\beta_F}^t$) 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

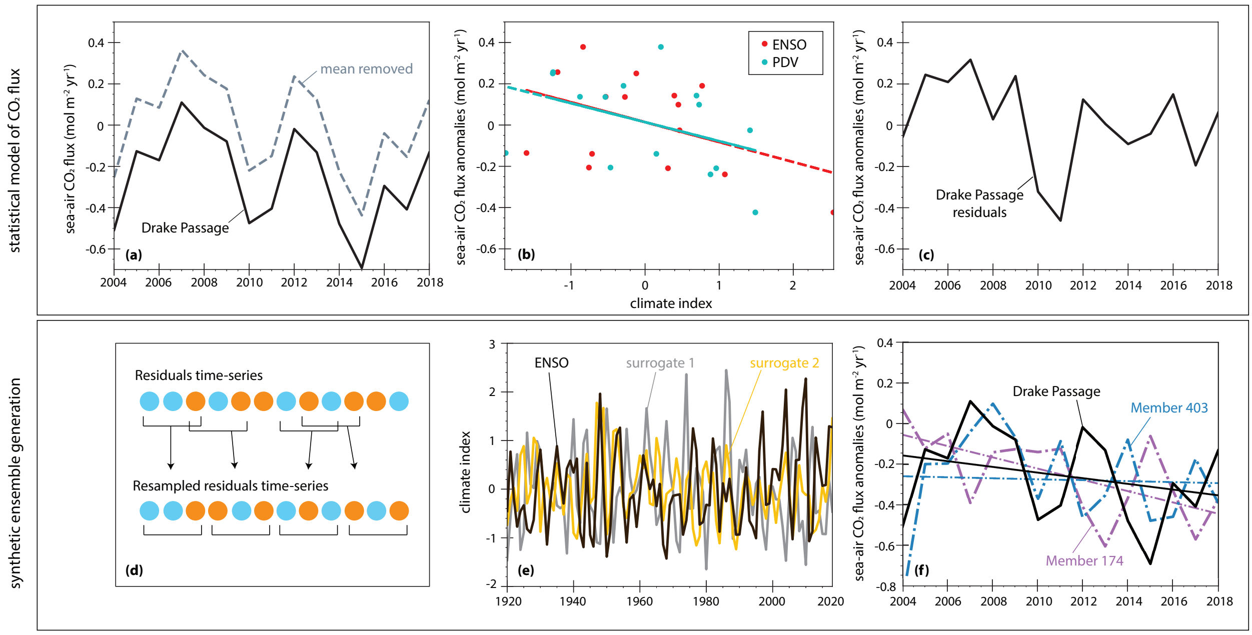

Figure 1. Synthetic ensemble construction. (a-c) Statistical model of annual mean Drake Passage sea-air CO_2 flux, as in Equation 1: (a) (solid) Time-series of CO_2 flux (black, $X^{i,t}$) and (dashed) CO_2 flux with temporal mean (β_0^i) removed, (b) regression of CO_2 flux anomalies (temporal mean and response to external forcing (a simple linear trend in this example), ${\beta_F}^t$, removed) onto the ENSO (red) and 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].

(${\beta^i}_{\mathrm{ENSO}}$, ${\beta^i}_{\mathrm{PDV}_\perp}$; 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, and these residuals are quite large in the Drake Passage region (Figure 1c), suggesting only a small role for ENSO and PDV in CO_2 flux here.

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 with a block length of 3 years to re-sample the residuals ($\varepsilon^{i,t}$) 1,000 times [Figure 1d; block length according to *Wilks*, 1997]. 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. This technique and block length preserves some of the temporal characteristics (*e.g.*, year-to-year variations) of the residuals (Figure 1d). Next, we use the Iterative Amplitude Adjustment Fourier Transfer (IAAFT) technique [*Schreiber and 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. 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 (${\beta_F}^t$), and the temporal mean flux (${\beta_0}^i$). 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 and 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 for interpretation of 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, JENA-MLS, CMEMS-FFNN, and MPI-SOMFFN; Figure 2) and for the average of the 4 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 signal (${\beta_F}^t$). 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*

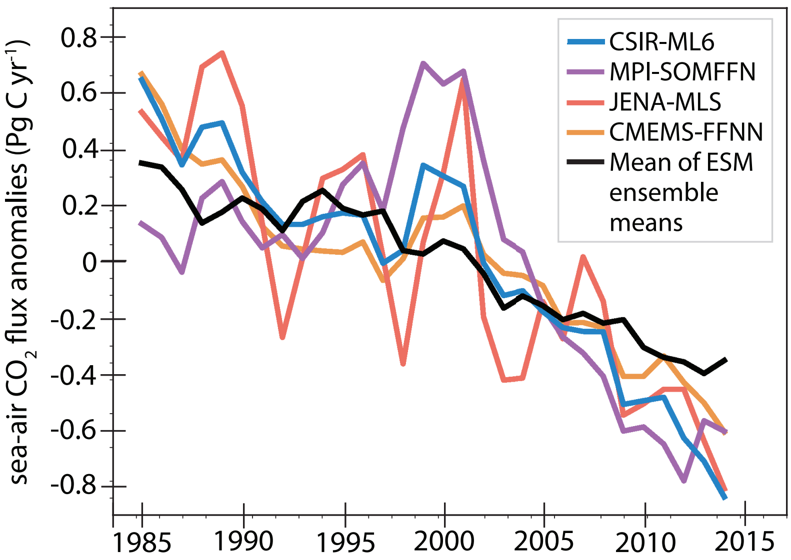

Figure 2. Global CO_2 flux variations. Temporal evolution of globally integrated sea-air CO_2 flux anomalies (temporal mean, β_0^i, removed) from the (blue) CSIR-ML6, (purple) MPI-SOMFFN, (pink) JENA-MLS, and (orange) CMEMS-FFNN observation-based products. Black line shows the global CO_2 flux response to external forcing, ${\beta_F}^t$, estimated as the mean of three ensemble means from Earth system model (ESM) output contributed to the CMIP6 archive (CanESM5, IPSL-CM6, CESM2). Positive flux anomalies correspond to decreased oceanic carbon uptake.

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, likely driven by external forcing), different phasing of internal variability produces a different 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.

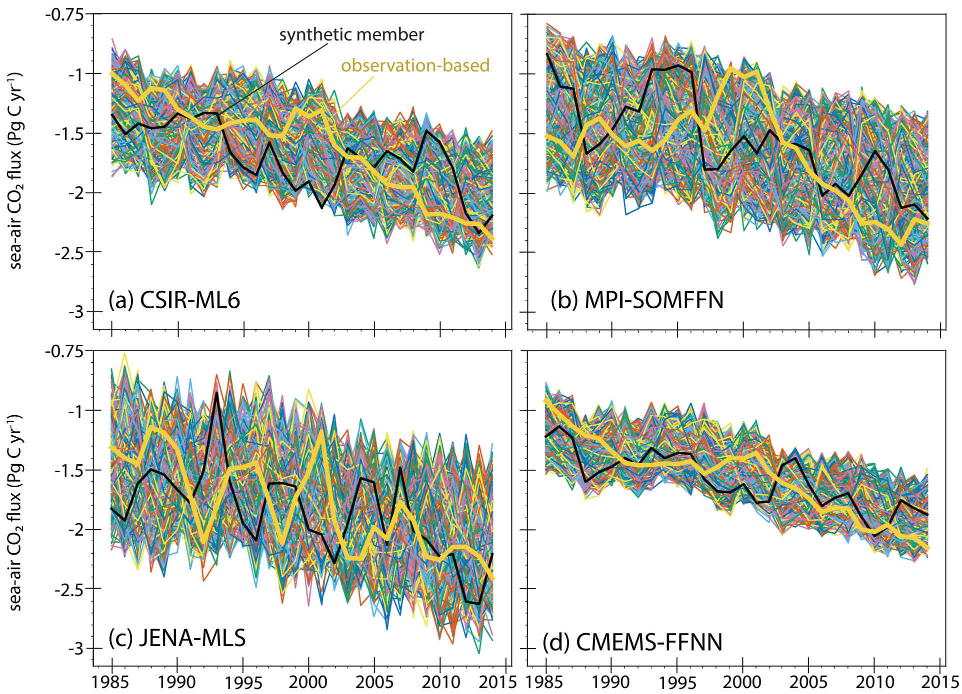

Figure 3. Synthetic ensembles of global sea-air CO_2 flux. Temporal evolution of globally integrated sea-air CO_2 flux from 1,000-member synthetic ensembles of the (a) CSIR-ML6, (b) MPI-SOMFFN, (c) JENA-MLS, and (d) CMEMS-FFNN observation-based products. Yellow lines show the CO_2 flux evolution from 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.

The synthetic ensembles of globally integrated CO_2 flux from the four observation-based CO_2 flux 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}

(1σ, CMEMS-FFNN) to 0.27 Pg C yr^{-1} (1σ, 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).

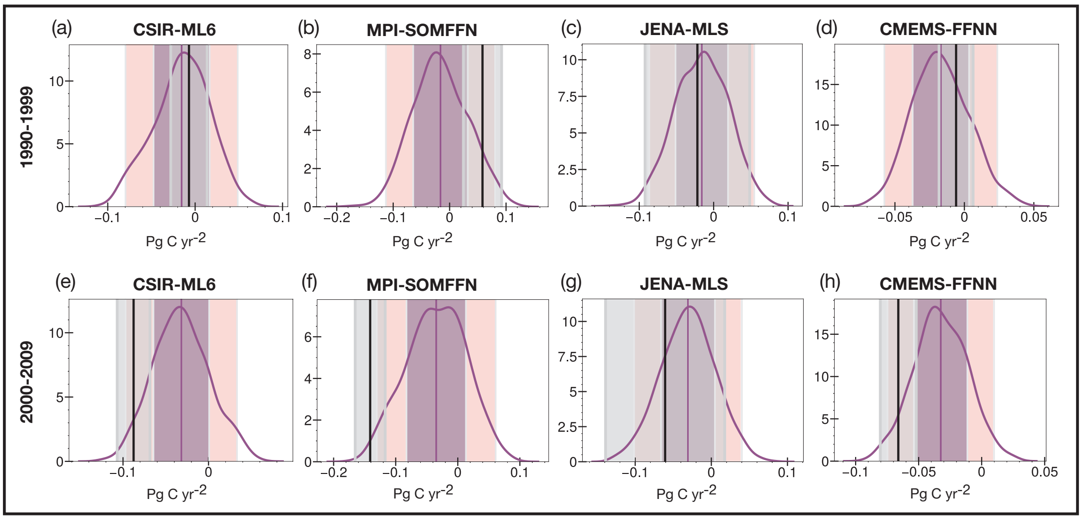

Figure 4. Probability of decadal trends in global CO_2 flux. Probability density functions (kernel density estimation, purple curves) of decadal trends in globally integrated sea-air CO_2 flux for (first row) 1990-1999 and (second row) 2000-2009, as estimated from synthetic ensembles of the (first column) CSIR-ML6, (second column) MPI-SOMFFN, (third column) JENA-MLS, and (fourth column) CMEMS-FFNN observation-based products. Purple vertical lines show the ensemble mean trend, and the 1σ (67%) and 2σ (95%) confidence intervals are shaded in purple and pink respectively. Black lines show the observed decadal trend from each product with its 95% confidence interval shaded in gray. Negative trends correspond to increased ocean carbon uptake with time. Note that the x- and y-axes differ between panels.

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, e.g., *Ritter* *et al.*, 2017]. 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 1σ (67%) confidence intervals shaded in purple and the 2σ (95%) confidence intervals shaded in pink (note the different x-axes for each product). 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 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). Armed with this information, we can now quantify the probability of the observed decadal trends (vertical black lines and associated gray shading 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). Using the observation-based trend mean value and 1σ or 2σ values, the observed trend in globally integrated CO_2 flux over 1990-1999 is a small negative number (more ocean carbon uptake with time) that is likely to occur (>40% chance of occurrence) in three of the four products (CSIR-ML6, JENA-MLS, and CMEMS-FFNN) within the distribution of synthetic trends (Figure 4, Table 2). Whereas, the observed trend in MPI-SOMFFN over 1990-1999 is a positive number (less ocean carbon uptake with time) that has a low probability of occurrence (25%) within the distribution of synthetic trends (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 (<25% chance of occurrence (Figure 4, Table 2), calling into question the Earth system model representation of external forcing in this period. 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 product ensembles indicate high probability of the observed trends, but over 2000-2009, three of the four products indicate that the observed trends are somewhat improbable with different

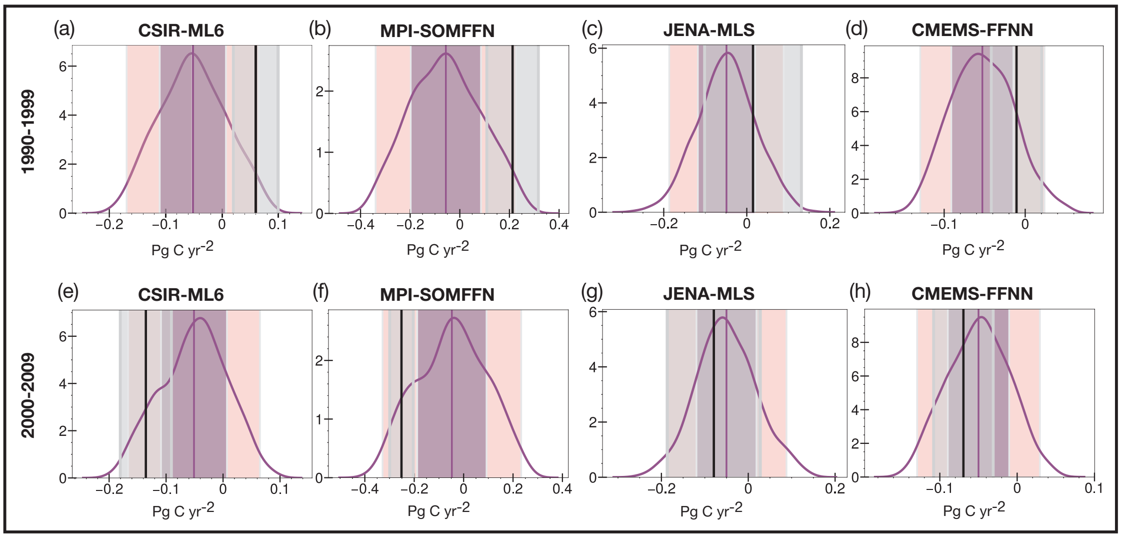

Figure 5. Probability of decadal trends in Southern Ocean CO_2 flux. As in Figure 4, but for the CO_2 flux integrated over the Southern hemisphere super-biome, made up of Southern Ocean Ice, Subpolar and Subtropical Seasonally Stratified biomes [biomes defined in *Fay and McKinley*, 2014]

.

phasing of internal variability, with only one (JENA-MLS) slightly probable (Figure 4, Table 2). For MPI-SOMFFN, however, the observed trends fall on the tails of the PDFs for both decades (25% chance of occurrence in 1990-1999 and 17% chance of occurrence in 2000-2009; Table 2), suggesting that different phasing of internal variability would likely have produced different observed trends in this product. A synthetic ensemble generated from the average of all four observation-based products produces a narrow synthetic trend distribution (0.11 Pg C yr^{-2}) and low probabilities for observed trends in both decades (Figure S1; Table 2), much like the MPI-SOMFFN product.

We estimate the probability of observed trends in regional sea-air CO_2 flux over 1990-1999 and 2000-2009 by creating synthetic ensembles of CO_2 flux integrated over "super-biomes", i.e., biomes that capture large-scale oceanographic regions [*Canadell et al.*, 2021], and performing similar statistical analyses as for the globally integrated fluxes (Figures 5 and S2-S5). 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 [*Le Quéré et al.*, 2007; *Lovenduski et al.*, 2008; *Landschützer et al.*, 2015; *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 and McKinley*, 2014] produce narrow PDFs of the decadal trends over 1990-1999 and 2000-2009 (Figure 5). 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 -0.05 to 0.05 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 1σ 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 1σ confidence interval of the PDFs, indicating low chance of occurrence given different phasing of internal variability (Figures 5a,b and 5e,f). Thus, results from this analysis suggest that the magnitudes of observed decadal trends in the Southern Ocean carbon sink discussed in the literature [e.g., *Landschützer et al.*, 2015] are not consistent across the observation-based products (as also noted by *Ritter et al.* [2017] and *DeVries et al.* [2019]), and for CSIR-ML6 and MPI-SOMFFN, are somewhat improbable given the distribution of synthetic trends and an assumed externally forced signal. This cross-product inconsistency is expected from the sparse pCO_2^{oc} measurements in the Southern Ocean [*Bakker et al.*, 2016; *Gloege et al.*, 2021], and the improbable nature of the observed trends advocates for a more refined approach, like the one presented in this study, to report on the likelihood of trends experienced in this region.

The distribution of synthetic trends and the probabilities of observed trends in other super-biomes over 1990-1999 and 2000-2009 are shown in Supporting Information Figures S2-S5 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, similar to the global fluxes, the observed trends over 2000-2009 are somewhat improbable within this distribution (Figure S2). 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 Northern hemisphere, -0.15 and 0.1 in the Southern hemisphere, Figures S3 and S4). The Equatorial super-biome synthetic ensemble produces a wide distribution of decadal trends over both

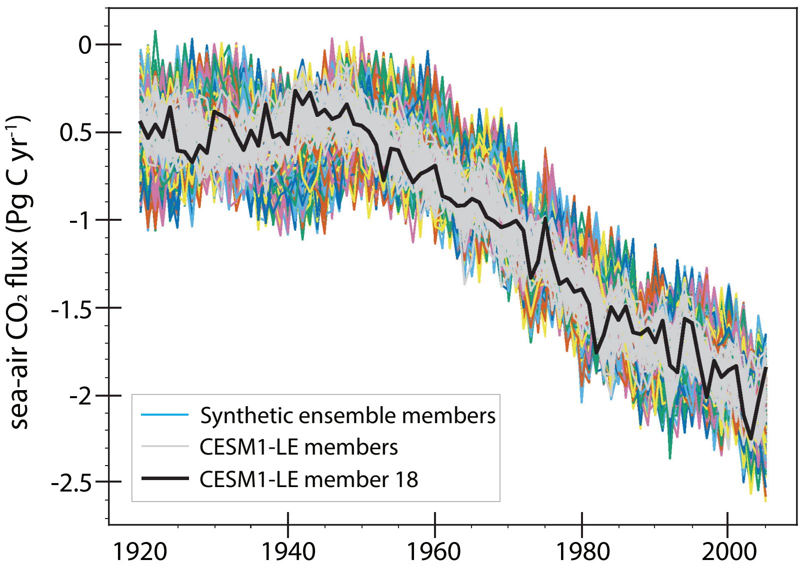

Figure 6. Synthetic ensemble of global CO_2 flux using CESM1-LE. Temporal evolution of the globally integrated sea-air CO_2 flux from the (pastel colors) CESM1-LE and (gray) synthetic ensemble of CESM1-LE member 18 (shown in black). Negative fluxes correspond to ocean carbon uptake.

periods (-0.5 to 0.4 Pg C yr^{-2}) and highly probable observed trends (Figure S5). 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 18, black line in Figure 6) using Equation 1 with the external signal (${\beta_F}^t$) modeled as the mean of three ensemble mean CO_2 flux estimates from historical simulations of CMIP6 Earth system models (see Section 2.4), and the climate modes (${M^t}_{\mathrm{ENSO}}$ and ${M^t}_{\mathrm{PDV}_\perp}$) produced from CESM1 for this ensemble member. We then generate a 1,000-member synthetic ensemble of ensemble member 18 as before. Figure 6 illustrates the resulting synthetic ensemble in pastel colors overlain on the original CESM1-LE ensemble in gray. While not exactly identical, the envelope of variability in our synthetic ensemble is a close match to that of the full CESM1-LE (Figure 6), as the average standard deviation (trend removed) of the synthetic ensemble is only 2% 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 CESM1-LE) averaged over 1920-2005 in Table 3. This analysis reveals that the synthetic ensemble can overestimate the standard deviation by as much as 29% (ensemble member 8) and can underestimate the standard deviation by as much as 15% (ensemble member 21), depending upon the statistical properties of the original time-series used to generate the synthetic ensemble. On average across the 34 ensemble members, however, the underestimation bias in the standard deviation is relatively small (2%; Table 3). Thus, results from this analysis suggest that our statistical emulation technique for synthetic ensemble generation is relatively unbiased.

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 observational 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 and 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). The synthetic ensembles generated from the shorter record produce larger biases in the standard deviations of the synthetic ensembles than the synthetic ensembles generated from the longer record, with overestimates as large as 54% and underestimates as small as 31% (Table 3). This finding lends support to continued and new observations of $p\mathrm{CO}_2^{oc}$ from which long records of sea-air CO_2 flux variability can be derived.

The probabilities of observed trends reported in the previous section are undoubtedly sensitive to the externally derived signal. Because the externally derived signal sets the center value of the synthetic trend distribution, a different assumption about this signal can shift the distribution to the left/right and affect the probability of the observed trend. 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 $p\mathrm{CO}_2$ and volcanic eruptions alone. Figure 7 illustrates that the probability of

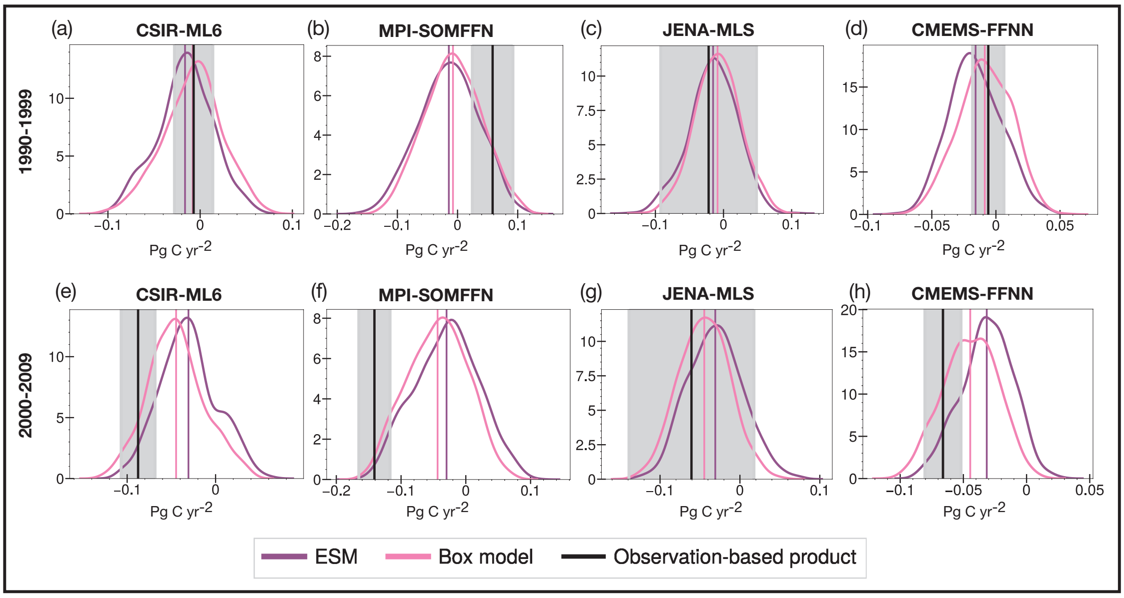

Figure 7. Probability of decadal trends in global CO_2 flux with alternative model of external forcing. Probability density functions (kernel density estimation, purple and pink curves) of decadal trends in globally integrated sea-air CO_2 flux for (first row) 1990-1999 and (second row) 2000-2009, as estimated from synthetic ensembles of the (first column) CSIR-ML6, (second column) MPI-SOMFFN, (third column) JENA-MLS, and (fourth column) 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 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 with 95% confidence intervals in gray shading. Note that the x- and y-axes differ between panels.

observed trends are similar regardless of whether we model the external forcing (${\beta_F}^t$) using the *McKinley et al.* [2020] box model or the mean of the three Earth system model ensemble means.

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 probability of decadal trends in CO_2 flux for each observation-based product. We further comment on the likelihood of the observed decadal trends given the synthetic trend probability 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 are highly probable given different phasing of internal variability. However, over the decade 2000-2009, three of the four products show somewhat improbable larger negative trends in sea-air CO_2 flux, calling into question the Earth system model estimate of external forcing in this period. The JENA-MLS product trends over these decades are highly probable, while the MPI-SOMFFN product trends over these decades are unlikely given different phasing of internal variability. The signs of the observed decadal trends in Southern Ocean sea-air CO_2 flux are inconsistent across the four observation-based products and, in the case of MPI-SOMFFN and CSIR-ML6, their magnitude is somewhat improbable given different phasing of internal variability. 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, e.g., *McKinley et al.*, 2011; *Fay and McKinley*, 2013; *McKinley et al.*, 2016], our synthetic ensembles provide stirring visualizations of this variability. Further, the statistical properties of the synthetic ensembles provide a basis for examining the likelihood of observed global and regional trends in sea-air CO_2 flux given different phasing of internal variability. 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 [*Landschützer et al.*, 2019; *DeVries et al.*, 2017; *McKinley et al.*, 2020]. Regardless of how we model the externally forced signal, internal variability seems to play a key role in driving the observed decadal trends across our synthetic ensembles.

Sustained and new observations of $p\mathrm{CO}_2$ across the global ocean and the continued development and refinement of observation-based gap-filled products will further expand our understanding of sea-air CO_2 flux variations. This understanding is critical for near-term predictions of the global carbon cycle [e.g., *Ilyina et al.*, 2021] and for our community's ability to inform international emission reduction efforts [*Peters et al.*, 2017].

Acknowledgments

Our synthetic ensembles are available at https://doi.org/10.5281/zenodo.5227342.

The Surface Ocean CO2 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. Underway DPT $p\mathrm{CO}_2$ 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/). CMIP6 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). 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.

References

Atlas, R., R. N. Hoffman, J. Ardizzone, S. M. Leidner, J. C. Jusem, D. K. Smith, and D. Gombos (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, doi:10.1175/2010BAMS2946.1.

Bakker, D. C. E., B. Pfeil, C. S. Landa, N. Metzl, K. M. O'Brien, A. Olsen, K. Smith, C. Cosca, S. Harasawa, S. D. Jones, S.-I. Nakaoka, Y. Nojiri, U. Schuster, T. Steinhoff, C. Sweeney, T. Takahashi, B. Tilbrook, C. Wada, R. Wanninkhof, S. R. Alin, C. F. Balestrini, L. Barbero, N. R. Bates, A. A. Bianchi, F. Bonou, J. Boutin, Y. Bozec, E. F. Burger, W.-J. Cai, R. D. Castle, L. Chen, M. Chierici, K. Currie, W. Evans, C. Featherstone, R. A. Feely, A. Fransson, C. Goyet, N. Greenwood, L. Gregor, S. Hankin, N. J. Hardman-Mountford, J. Harlay, J. Hauck, M. Hoppema, M. P. Humphreys, C. W. Hunt, B. Huss, J. S. P. Ibánhez, T. Johannessen, R. Keeling, V. Kitidis, A. Körtzinger, A. Kozyr, E. Krasakopoulou, A. Kuwata, P. Landschützer, S. K. Lauvset, N. Lefèvre, C. Lo Monaco, A. Manke, J. T. Mathis, L. Merlivat, F. J. Millero, P. M. S. Monteiro, D. R. Munro, A. Murata, T. Newberger, A. M. Omar, T. Ono, K. Paterson, D. Pearce, D. Pierrot, L. L. Robbins, S. Saito, J. Salisbury, R. Schlitzer, B. Schneider, R. Schweitzer, R. Sieger, I. Skjelvan, K. F. Sullivan, S. C. Sutherland, A. J. Sutton, K. Tadokoro, M. Telszewski, M. Tuma, S. M. A. C. van Heuven, D. Vandemark, B. Ward, A. J. Watson, and S. Xu (2016), A multi-decade record of high-quality $f\mathrm{CO}_2$ data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT), *Earth Syst. Sci. Data*, *8*(2), 383–413, doi:10.5194/essd-8-383-2016.

Boucher, O., J. Servonnat, A. L. Albright, O. Aumont, Y. Balkanski, V. Bastrikov, S. Bekki, R. Bonnet, S. Bony, L. Bopp, P. Braconnot, P. Brockmann, P. Cadule, A. Caubel,

F. Cheruy, F. Codron, A. Cozic, D. Cugnet, F. D'Andrea, P. Davini, C. de Lavergne, S. Denvil, J. Deshayes, M. Devilliers, A. Ducharne, J.-L. Dufresne, E. Dupont, C. Éthé, L. Fairhead, L. Falletti, S. Flavoni, M.-A. Foujols, S. Gardoll, G. Gastineau, J. Ghattas, J.-Y. Grandpeix, B. Guenet, E. Guez, Lionel, E. Guilyardi, M. Guimberteau, D. Hauglustaine, F. Hourdin, A. Idelkadi, S. Joussaume, M. Kageyama, M. Khodri, G. Krinner, N. Lebas, G. Levavasseur, C. Lévy, L. Li, F. Lott, T. Lurton, S. Luyssaert, G. Madec, J.-B. Madeleine, F. Maignan, M. Marchand, O. Marti, L. Mellul, Y. Meurdesoif, J. Mignot, I. Musat, C. Ottlé, P. Peylin, Y. Planton, J. Polcher, C. Rio, N. Rochetin, C. Rousset, P. Sepulchre, A. Sima, D. Swingedouw, R. Thiéblemont, A. K. Traore, M. Vancoppenolle, J. Vial, J. Vialard, N. Viovy, and N. Vuichard (2020), Presentation and evaluation of the IPSL-CM6A-LR climate model, *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS002,010, doi:doi.org/10.1029/2019MS002010, e2019MS002010 10.1029/2019MS002010.

Canadell, J. G., P. M. S. Monteiro, M. H. Costa, L. C. da Cunha, P. M. Cox, A. V. Eliseev, S. Henson, M. Ishii, S. Jaccard, C. Koven, A. Lohila, P. K. Patra, S. Piao, J. Rogelj, S. Syampungani, S. Zaehle, and K. Zickfeld (2021), Global Carbon and other Biogeochemical Cycles and Feedbacks, *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]*.

Ciais, P., and C. Sabine (2013), Chapter 6: Carbon and Other Biogeochemical Cycles, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, p. 1535 pp, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Danabasoglu, G., S. C. Bates, B. P. Briegleb, S. R. Jayne, M. Jochum, W. G. Large, S. Peacock, and S. G. Yeager (2012), The CCSM4 Ocean Component, *J. Climate*, *25*(5), 1361–1389.

Danabasoglu, G., J. F. Lamarque, J. Bacmeister, D. A. Bailey, A. K. DuVivier, J. Edwards, L. K. Emmons, J. Fasullo, R. Garcia, A. Gettelman, C. Hannay, M. M. Holland, W. G.

Large, P. H. Lauritzen, D. M. Lawrence, J. T. M. Lenaerts, K. Lindsay, W. H. Lipscomb, M. J. Mills, R. Neale, K. W. Oleson, B. Otto-Bliesner, A. S. Phillips, W. Sacks, S. Tilmes, L. van Kampenhout, M. Vertenstein, A. Bertini, J. Dennis, C. Deser, C. Fischer, B. Fox-Kemper, J. E. Kay, D. Kinnison, P. J. Kushner, V. E. Larson, M. C. Long, S. Mickelson, J. K. Moore, E. Nienhouse, L. Polvani, P. J. Rasch, and W. G. Strand (2020), The Community Earth System Model version 2 (CESM2), *J. Adv. Model. Earth Syst.*, *12*(2), e2019MS001,916, doi:10.1029/2019MS001916.

Denvil-Sommer, A., M. Gehlen, M. Vrac, and C. Mejia (2019), LSCE-FFNN-v1: a two-step neural network model for the reconstruction of surface ocean $p\mathrm{co}_2^{oc}$ over the global ocean, *Geosci. Model Dev.*, *12*(5), 2091–2105, doi:10.5194/gmd-12-2091-2019.

Deser, C., F. Lehner, K. B. Rodgers, T. Ault, T. L. Delworth, P. N. DiNezio, A. Fiore, C. Frankignoul, J. C. Fyfe, D. E. Horton, J. E. Kay, R. Knutti, N. S. Lovenduski, J. Marotzke, K. A. McKinnon, S. Minobe, J. Randerson, J. A. Screen, I. R. Simpson, and M. Ting (2020), Insights from Earth system model initial-condition large ensembles and future prospects, *Nature Clim. Change*, *10*(4), 277–286, doi:10.1038/s41558-020-0731-2.

DeVries, T., M. Holzer, and F. Primeau (2017), Recent increase in oceanic carbon uptake driven by weaker upper-ocean overturning, *Nature*, *542*(7640), 215–218.

DeVries, T., C. Le Quéré, O. Andrews, S. Berthet, J. Hauck, T. Ilyina, P. Landschützer, A. Lenton, I. D. Lima, M. Nowicki, J. Schwinger, and R. Séférian (2019), Decadal trends in the ocean carbon sink, *Proceedings of the National Academy of Sciences*, *116*(24), 11,646–11,651, doi:10.1073/pnas.1900371116.

Doney, S. C., D. S. Busch, S. R. Cooley, and K. J. Kroeker (2020), The impacts of ocean acidification on marine ecosystems and reliant human communities, *Annu. Rev. Env. Resour.*, doi:10.1146/annurev-environ-012320-083019.

Elsworth, G. W., N. S. Lovenduski, K. A. McKinnon, K. M. Krumhardt, and R. X. Brady (2020), Finding the fingerprint of anthropogenic climate change in marine phytoplankton abundance, *Curr. Clim. Chang. Rep.*, *6*(2), 37–46, doi:10.1007/s40641-020-00156-w.

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

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

Fay, A. R., and G. A. McKinley (2014), Global open-ocean biomes: mean and temporal variability, *Earth Syst. Sci. Data*, *6*(2), 273–284, doi:10.5194/essd-6-273-2014.

Fay, A. R., N. S. Lovenduski, G. A. McKinley, D. R. Munro, C. Sweeney, A. R. Gray, P. Landschützer, B. B. Stephens, T. Takahashi, and N. Williams (2018), Utilizing the Drake Passage Time-series to understand variability and change in subpolar Southern Ocean $p\mathrm{CO}_2$, *Biogeosciences*, *15*(12), 3841–3855, doi:10.5194/bg-15-3841-2018.

Fay, A. R., L. Gregor, P. Landschützer, G. A. McKinley, N. Gruber, M. Gehlen, Y. Iida, G. G. Laruelle, C. Rödenbeck, and J. Zeng (2021), Harmonization of global surface ocean pco_2 mapped products and their flux calculations; an improved estimate of the ocean carbon sink, *Earth System Science Data Discussions*, *2021*, 1–32, doi:10.5194/essd-2021-16.

Friedlingstein, P., M. O'Sullivan, M. W. Jones, R. M. Andrew, J. Hauck, A. Olsen, G. P. Peters, W. Peters, J. Pongratz, S. Sitch, C. Le Quéré, J. G. Canadell, P. Ciais, R. B. Jackson, S. Alin, L. E. O. C. Aragão, A. Arneth, V. Arora, N. R. Bates, M. Becker, A. Benoit-Cattin, H. C. Bittig, L. Bopp, S. Bultan, N. Chandra, F. Chevallier, L. P. Chini, W. Evans, L. Florentie, P. M. Forster, T. Gasser, M. Gehlen, D. Gilfillan, T. Gkritzalis, L. Gregor, N. Gruber, I. Harris, K. Hartung, V. Haverd, R. A. Houghton, T. Ilyina, A. K. Jain, E. Joetzjer, K. Kadono, E. Kato, V. Kitidis, J. I. Korsbakken, P. Landschützer, N. Lefèvre, A. Lenton, S. Lienert, Z. Liu, D. Lombardozzi, G. Marland, N. Metzl, D. R. Munro, J. E. M. S. Nabel, S.-I. Nakaoka, Y. Niwa, K. O'Brien, T. Ono, P. I. Palmer, D. Pierrot, B. Poulter, L. Resplandy, E. Robertson, C. Rödenbeck, J. Schwinger, R. Séférian, I. Skjelvan, A. J. P. Smith, A. J. Sutton, T. Tanhua, P. P. Tans, H. Tian, B. Tilbrook, G. van der Werf, N. Vuichard, A. P. Walker, R. Wanninkhof, A. J. Watson, D. Willis, A. J. Wiltshire, W. Yuan, X. Yue, and S. Zaehle (2020), Global Carbon Budget 2020, *Earth Syst. Sci. Data*, *12*(4), 3269–3340, doi:10.5194/essd-12-3269-2020.

Gloege, L., G. A. McKinley, P. Landschützer, A. R. Fay, T. L. Frölicher, J. C. Fyfe, T. Ilyina, S. Jones, N. S. Lovenduski, K. B. Rodgers, S. Schlunegger, and Y. Takano (2021), Quantifying errors in observationally based estimates of ocean carbon sink variability, *Global Biogeochem. Cycles*, *35*(4), e2020GB006,788, doi: https://doi.org/10.1029/2020GB006788.

Gregor, L., and A. Fay (2021), SeaFlux: harmonised sea-air CO_2 fluxes from surface $p\mathrm{CO}_2$ data products using a standardised approach, *Zenodo*.

Gregor, L., A. D. Lebehot, S. Kok, and P. M. Scheel Monteiro (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?, *Geosci. Model Dev.*, *12*(12), 5113–5136, doi:10.5194/gmd-12-5113-2019.

Hauck, J., M. Zeising, C. Le Quéré, N. Gruber, D. C. E. Bakker, L. Bopp, T. T. T. Chau, Ö. Gürses, T. Ilyina, P. Landschützer, A. Lenton, L. Resplandy, C. Rödenbeck, J. Schwinger, and R. Séférian (2020), Consistency and challenges in the ocean carbon sink estimate for the Global Carbon Budget, *Front. Mar. Sci.*, *7*, 852, doi: 10.3389/fmars.2020.571720.

Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *B. Am. Meteorol. Soc.*, *90*(8), 1095–1107, doi:10.1175/2009BAMS2607.1.

Ho, D. T., C. S. Law, M. J. Smith, P. Schlosser, M. Harvey, and P. Hill (2006), Measurements of air-sea gas exchange at high wind speeds in the Southern Ocean: Implications for global parameterizations, *Geophysical Research Letters*, *33*(16), doi: https://doi.org/10.1029/2006GL026817.

Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. J. Kushner, J. F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, W. H. Lipscomb, M. C. Long, N. Mahowald, D. R. Marsh, R. B. Neale, P. Rasch, S. Vavrus, M. Vertenstein, D. Bader, W. D. Collins, J. J. Hack, J. Kiehl, and S. Marshall (2013), The Community Earth System Model: A Framework for Collaborative Research, *B. Am. Meteorol. Soc.*, *94*(9), 1339–1360, doi: 10.1175/BAMS-D-12-00121.1.

Ilyina, T., H. Li, A. Spring, W. A. Müller, L. Bopp, M. O. Chikamoto, G. Danabasoglu, M. Dobrynin, J. Dunne, F. Fransner, P. Friedlingstein, W. Lee, N. S. Lovenduski, W. J. Merryfield, J. Mignot, J. Y. Park, R. Séférian, R. Sospedra-Alfonso, M. Watanabe, and S. Yeager (2021), Predictable variations of the carbon sinks and atmospheric CO_2 growth in a multi-model framework, *Geophys. Res. Lett.*, *48*(6), e2020GL090,695, doi: doi.org/10.1029/2020GL090695.

Kay, J. E., C. Deser, A. Phillips, A. Mai, C. Hannay, G. Strand, J. M. Arblaster, S. C. Bates, G. Danabasoglu, J. Edwards, M. Holland, P. Kushner, J. F. Lamarque, D. Lawrence, K. Lindsay, A. Middleton, E. Munoz, R. Neale, K. Oleson, L. Polvani, and M. Vertenstein (2015), The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability, *B. Am. Meteorol. Soc.*, *96*(8), 1333–1349, doi:10.1175/BAMS-D-13-00255.1.

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

Landschützer, P., N. Gruber, F. A. Haumann, C. Rödenbeck, D. C. E. Bakker, S. van Heuven, M. Hoppema, N. Metzl, C. Sweeney, T. Takahashi, B. Tilbrook, and R. Wanninkhof (2015), The reinvigoration of the Southern Ocean carbon sink, *Science*, *349*(6253), 1221–1224.

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

Landschützer, P., T. Ilyina, and N. S. Lovenduski (2019), Detecting regional modes of variability in observation-based surface ocean pCO_2, *Geophys. Res. Lett.*, *46*(5), 2670–2679, doi:10.1029/2018GL081756.

Landschützer, P., G. G. Laruelle, A. Roobaert, and P. Regnier (2020), A uniform $p\mathrm{co}_2$ climatology combining open and coastal oceans, *Earth Syst. Sci. Data*, *12*(4), 2537–2553, doi:10.5194/essd-12-2537-2020.

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

Le Quéré, C., C. Rödenbeck, E. T. Buitenhuis, T. J. Conway, R. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metzl, N. Gillett, and M. Heimann (2007), Saturation of the Southern Ocean CO_2 sink due to recent climate change, *Science*, *316*(5832), 1735–1738.

Le Quéré, C., R. M. Andrew, P. Friedlingstein, S. Sitch, J. Pongratz, A. C. Manning, J. I. Korsbakken, G. P. Peters, J. G. Canadell, R. B. Jackson, T. A. Boden, P. P. Tans, O. D. Andrews, V. K. Arora, D. C. E. Bakker, L. Barbero, M. Becker, R. A. Betts, L. Bopp, F. Chevallier, L. P. Chini, P. Ciais, C. E. Cosca, J. Cross, K. Currie, T. Gasser, I. Harris, J. Hauck, V. Haverd, R. A. Houghton, C. W. Hunt, G. Hurtt, T. Ilyina, A. K. Jain, E. Kato, M. Kautz, R. F. Keeling, K. Klein Goldewijk, A. Körtzinger, P. Landschützer, N. Lefèvre, A. Lenton, S. Lienert, I. Lima, D. Lombardozzi, N. Metzl, F. Millero, P. M. S. Monteiro, D. R. Munro, J. E. M. S. Nabel, S.-I. Nakaoka, Y. Nojiri, X. A. Padin, A. Peregon, B. Pfeil, D. Pierrot, B. Poulter, G. Rehder, J. Reimer, C. Rödenbeck, J. Schwinger, R. Séférian, I. Skjelvan, B. D. Stocker, H. Tian, B. Tilbrook, F. N. Tubiello, I. T. van der Laan-Luijkx, G. R. van der Werf, S. van Heuven, N. Viovy, N. Vuichard, A. P. Walker,

A. J. Watson, A. J. Wiltshire, S. Zaehle, and D. Zhu (2018), Global Carbon Budget 2017, *Earth Syst. Sci. Data*, *10*(1), 405–448, doi:10.5194/essd-10-405-2018.

Le Quéré, C., C., M. R. Raupach, J. G. Canadell, and G. Marland et al. (2009), Trends in the sources and sinks of carbon dioxide, *Nature Geosci.*, *2*(12), 831–836.

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

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

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

McKinley, G. A., T. Takahashi, E. Buitenhuis, F. Chai, J. R. Christian, S. C. Doney, M.-S. Jiang, K. Lindsay, J. K. Moore, C. Le Quéré, I. Lima, R. Murtugudde, L. Shi, and P. Wetzel (2006), North Pacific carbon cycle response to climate variability on seasonal to decadal timescales, *J. Geophys. Res. Oceans*, *111*(C7), doi:10.1029/2005JC003173.

McKinley, G. A., A. R. Fay, T. Takahashi, and N. Metzl (2011), Convergence of atmospheric and North Atlantic carbon dioxide trends on multidecadal timescales, *Nature Geosci.*, *4*(9), 606–610.

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

McKinley, G. A., A. R. Fay, N. S. Lovenduski, and D. J. Pilcher (2017), Natural variability and anthropogenic trends in the ocean carbon sink, *Annu. Rev. Mar. Sci.*, *9*(1), 125–150, doi:10.1146/annurev-marine-010816-060529.

McKinley, G. A., A. R. Fay, Y. A. Eddebbar, L. Gloege, and N. S. Lovenduski (2020), External forcing explains recent decadal variability of the ocean carbon sink, *AGU Advances*, *1*(2), e2019AV000,149, doi:10.1029/2019AV000149.

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

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

Moore, J. K., and O. Braucher (2008), Sedimentary and mineral dust sources of dissolved iron to the world ocean, *Biogeosciences*, *5*(3), 631–656.

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

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

Munro, D. R., N. S. Lovenduski, B. B. Stephens, T. Newberger, K. R. Arrigo, T. Takahashi, P. D. Quay, J. Sprintall, N. M. Freeman, and C. Sweeney (2015a), 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, doi:10.1016/j.dsr2.2014.12.014.

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

Peters, G. P., C. Le Quéré, R. M. Andrew, J. G. Canadell, P. Friedlingstein, T. Ilyina, R. B. Jackson, F. Joos, J. I. Korsbakken, G. A. McKinley, S. Sitch, and P. Tans (2017), Towards real-time verification of CO_2 emissions, *Nature Clim. Change*, *7*(12), 848–850, doi:10.1038/s41558-017-0013-9.

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

Ritter, R., P. Landschützer, N. Gruber, A. R. Fay, Y. Iida, S. Jones, S. Nakaoka, G.-H. Park, P. Peylin, C. Rödenbeck, K. B. Rodgers, J. D. Shutler, and J. Zeng (2017), Observation-based trends of the Southern Ocean carbon sink, *Geophys. Res. Lett.*, *44*(24), 12,339–12,348, doi:10.1002/2017GL074837.

Rödenbeck, C., D. C. E. Bakker, N. Metzl, A. Olsen, C. Sabine, N. Cassar, F. Reum, R. F. Keeling, and M. Heimann (2014), Interannual sea–air CO_2 flux variability from an observation-driven ocean mixed-layer scheme, *Biogeosciences*, *11*(17), 4599–4613, doi:

10.5194/bg-11-4599-2014.

Rödenbeck, C., D. C. E. Bakker, N. Gruber, Y. Iida, A. R. Jacobson, S. Jones, P. Landschützer, N. Metzl, S. Nakaoka, A. Olsen, G.-H. Park, P. Peylin, K. B. Rodgers, T. P. Sasse, U. Schuster, J. D. Shutler, V. Valsala, R. Wanninkhof, and J. Zeng (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, doi:10.5194/bg-12-7251-2015.

Schreiber, T., and A. Schmitz (1996), Improved surrogate data for nonlinearity tests, *Phys. Rev. Lett.*, *77*, 635–638, doi:10.1103/PhysRevLett.77.635.

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

Swart, N. C., J. N. S. Cole, V. V. Kharin, M. Lazare, J. F. Scinocca, N. P. Gillett, J. Anstey, V. Arora, J. R. Christian, S. Hanna, Y. Jiao, W. G. Lee, F. Majaess, O. A. Saenko, C. Seiler, C. Seinen, A. Shao, M. Sigmond, L. Solheim, K. von Salzen, D. Yang, and B. Winter (2019), The Canadian Earth System Model version 5 (CanESM5.0.3), *Geosci. Model Dev.*, *12*(11), 4823–4873, doi:10.5194/gmd-12-4823-2019.

Takahashi, T., S. Sutherland, and A. Kozyr (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*.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *B. Am. Meteorol. Soc.*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

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

Wilks, D. (1997), Resampling hypothesis tests for autocorrelated fields, *J. Climate*, *10*(1), 65–82.

Observation-based product	Abbreviation	Methodology	Reference
Council for Scientific and Industrial Research-Machine Learning ensemble	CSIR-ML6	Ensemble of two-step neural network methods: two types of clusters and three types of regressions are used to interpolate pCO_2^{oc} from SOCAT v2020 using chlorophyll-a, sea surface temperature, absolute dynamic topography, mixed layer depth, sea ice, and sea surface salinity. The final product uses an ensemble average of six machine-learning models.	*Gregor et al.* [2019]
Max Planck Institute Self-Organizing Map-Feed-Forward Neural Network	MPI-SOMFFN	uses a two-step neural network method to gap-fill pCO_2^{oc}. In the first step, a self-organizing map is used to subdivide the ocean into 16 provinces with similar climatological biogeochemical properties. In the second step, a feed-forward neural network is used to predict the non-linear relationships between driver variables and SOCAT v2020 observations in each province. Driver variables for MPI-SOMFFN include sea surface temperature, mixed layer depth, satellite derived chlorophyll-a concentration, sea surface salinity, and atmospheric pCO_2.	*Landschützer et al.* [2013, 2014, 2015, 2016]
Jena, Germany, Max Planck Institute for Biogeochemistry - Mixed Layer Scheme	JENA-MLS	Combines ocean mixed layer biogeochemistry with pCO_2^{oc} data from SOCAT v2020 and seasonal, interannual and short-term (daily) variations of sea surface temperature, mixed layer depth, ice-free fraction, salinity, wind speed, and	*Rödenbeck et al.* [2014]

Observation-based product	Width of trend distribution 1990-1999	Width of trend distribution 2000-2009	Probability of observed trend 1990-1999	Probability of observed trend 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%

Table 2. Width of trend distributions and trend probabilities for synthetic ensembles of globally integrated CO_2 flux produced from the CSIR-ML6, MPI-SOMFFN, JENA-MLS, and CMEMS-FFNN observation-based products. Widths estimated as 4σ (Pg C yr^{-2}). Probabilities estimated as the lower/upper cumulative distribution for a normal distribution.

Ensemble member	$\frac{\sigma_{\mathrm{SE}}}{\sigma_{\mathrm{CESM1-LE}}}$ 1920-2005	$\frac{\sigma_{\mathrm{SE}}}{\sigma_{\mathrm{CESM1-LE}}}$ 1976-2005
1	1.16	1.08
2	0.90	1.10
3	1.09	0.83
4	0.91	0.90
5	1.15	0.92
6	0.86	0.88
7	1.00	0.69
8	1.29	1.07
9	0.96	1.18
10	0.99	1.21
11	0.98	1.10
12	0.93	1.22
13	0.93	0.78
14	1.01	1.05
15	1.01	1.14
16	0.88	0.84
17	0.91	0.93
18	1.00	1.10
19	1.10	0.78
20	0.95	0.94
21	0.85	0.93
22	1.11	1.26
23	0.97	0.94
24	0.90	1.17
25	1.01	1.02
26	1.05	0.94
27	0.87	0.72
28	1.10	1.54
29	1.06	1.24
30	1.12	1.00
31	1.10	1.03
32	1.10	1.03
33	1.11	0.99
34	0.99	0.93
mean	1.02	1.01

Table 3. Standard deviation quotient (synthetic ensemble standard deviation divided by Earth system model