Geophysical Research Letters

RESEARCH LETTER
10.1029/2022GL098632

# Variability in the Global Ocean Carbon Sink From 1959 to 2020 by Correcting Models With Observations

**Val Bennington[1,2], Lucas Gloege[3], and Galen A. McKinley[1]**

[1]Columbia University and Lamont-Doherty Earth Observatory New York, New York, NY, USA, [2]Makai Ocean Engineering, Waimanalo, HI, USA, [3]NASA-GISS, New York, NY, USA

**Key Points:**

- Climatological deviations of ocean hindcast models from observations are substantially larger than interannual deviations
- Climatological correction extends machine learning based hybrid model-observation product temporally for surface ocean $pCO_2$ (LDEO-Hybrid Physics Data product)
- The ocean carbon sink from 1959 to 2020 has responded to atmospheric $pCO_2$ growth and volcanic eruptions

**Supporting Information:**

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

**Correspondence to:**

V. Bennington,
Valerie.Bennington@makai.com

**Citation:**

Bennington, V., Gloege, L., & McKinley, G. A. (2022). Variability in the global ocean carbon sink from 1959 to 2020 by correcting models with observations. *Geophysical Research Letters*, *49*, e2022GL098632. https://doi.org/10.1029/2022GL098632

Received 9 MAR 2022
Accepted 23 JUN 2022

**Author Contributions:**

**Conceptualization:** Val Bennington, Lucas Gloege, Galen A. McKinley
**Funding acquisition:** Galen A. McKinley
**Investigation:** Val Bennington, Galen A. McKinley
**Methodology:** Val Bennington, Lucas Gloege, Galen A. McKinley
**Project Administration:** Galen A. McKinley
**Supervision:** Galen A. McKinley
**Writing – original draft:** Val Bennington, Lucas Gloege



**Abstract** The ocean reduces human impact on the climate by absorbing and sequestering $CO_2$. From 1950s to the 1980s, observations of $pCO_2$ and related ocean carbon variables were sparse and uncertain. Thus, global ocean biogeochemical models (GOBMs) have been the basis for quantifying the ocean carbon sink. The LDEO-Hybrid Physics Data product (LDEO-HPD) interpolates sparse surface ocean $pCO_2$ data to global coverage by using GOBMs as priors, and applying machine learning to estimate full-coverage corrections. The largest component of the GOBM corrections are climatological. This is consistent with recent findings of large seasonal discrepancies in GOBMs, but contrasts the long-held view that interannual variability is a major source of GOBM error. This supports extension of the LDEO-HPD $pCO_2$ product back to 1959, using a climatology of model-observation misfits prior to 1982. Consistent with previous studies for 1980 onward, air-sea $CO_2$ fluxes for 1959–2020 demonstrate response to atmospheric $pCO_2$ growth and volcanic eruptions.

**Plain Language Summary** The ocean removes carbon dioxide ($CO_2$) from the atmosphere and reduces climate change caused by humans. The magnitude of this removal can be estimated using computer models of ocean physics, chemistry, and biology, as well as statistical extrapolations of observations. The observational record is too sparse to directly reconstruct air-sea fluxes prior to 1982, but by combining models and a statistical approach, we make an estimate for 1959-present that is substantially informed by observations. The LDEO-Hybrid Physics Data product (LDEO-HPD) product for air-sea $CO_2$ exchange includes two periods, with the first previously published for 1982–2018 and extended here to end in 2020, and the second being this extension back in time. For 1959–1981, LDEO-HPD corrects models using the monthly average of data-based corrections derived from the observed period, a choice justified by our finding that these monthly means are the largest component of the needed corrections during the observed period. The LDEO-HPD product agrees much better with independent observations than the models alone, and can be used to understand what controls year to year changes in the ocean carbon sink.

## 1. Introduction

By absorbing and sequestering carbon dioxide from the atmosphere, the global oceans play a critical role in modulating climate change. The ocean has absorbed 37% of fossil carbon emissions since the start of the industrial age (Friedlingstein et al., 2022). Quantifying the redistribution of carbon emissions in the land biosphere, ocean and atmospheric reservoirs supports climate policy (Peters et al., 2017). In order to estimate air-sea fluxes of carbon dioxide, the driver of these fluxes, the partial pressure of carbon dioxide in the surface waters ($pCO_2$) must be estimated.

Global ocean biogeochemical models (GOBMs) explicitly simulate the physics, biology and chemistry of the ocean carbonate system based on equations that represent the physical and biogeochemical processes. Forced with winds and surface energy fluxes from observations for recent decades, the models estimate the state of the ocean physics and biogeochemistry for the same decades. Output from these models include a vast array of variables, including surface ocean $pCO_2$ and air-sea $CO_2$ flux. It has long been believed that GOBMs underrepresent the magnitude of interannual variability of the ocean carbon sink (DeVries et al., 2019; Gruber et al., 2019; Landschützer et al., 2015; Le Quéré et al., 2007), though the mechanisms of this proposed underrepresentation have not been identified. More recently, the community has paid greater attention to the significant seasonal biases in GOBMs (Hauck et al., 2020; Mongwe et al., 2018).

Observation-based products utilize sparse observations of $pCO_2$ from the Surface Ocean $CO_2$ ATlas (SOCAT) (Bakker et al., 2016), and train a machine learning algorithm to relate these data to full-coverage observations of

**Writing – review & editing:** Val Bennington, Lucas Gloege, Galen A. McKinley

associated variables, such that $pCO_2$ can be estimated at all points in space and time. Although these algorithms often do not explicitly include the known physics of the ocean carbonate system, the results do compare well to independent observations of $pCO_2$ (Bennington et al., 2022; Denvil-Sommer et al., 2019; Gregor et al., 2019; Landschützer et al., 2014). The mixed layer model of Rödenbeck et al. (2013, 2022) is another approach to creating an observation-based product. It combines statistical fits to data and explicit representations of physical processes.

Due to the sparsity of $pCO_2$ data, observation-based products have been limited to the period of in situ observations that started to become more numerous in the 1980s. For the period of 1959–1990, eight GOBMs were used to quantify the historical air-sea $CO_2$ flux in the Global Carbon Budget 2021 (Friedlingstein et al., 2022). For 1990–2020, the average of of eight GOBMs and seven data products was used as the basis for this estimate. In all prior Global Carbon Budget releases, for example, Friedlingstein et al. (2020), only the average of GOBMs were used to estimate the ocean carbon sink.

To directly incorporate the physical knowledge contained within GOBMs into an observation-based product, Gloege et al. (2022) utilized the machine-learning algorithm XGBoost (Chen & Guestrin, 2016) to learn model-observation misfits of GOBM simulated surface ocean $pCO_2$. The resulting data product (LDEO-HPD) showed an improved fit compared to the independent data over other data products. The resulting historical reconstruction of air-sea $CO_2$ fluxes from the extended LDEO-HPD is within the range of other data products, and in agreement with 2010–2020 mean flux estimates from the Global Carbon Budget 2021 (Friedlingstein et al., 2022).

LDEO-HPD estimated air-sea fluxes beginning in 1982. Here, we extend LDEO-HPD back in time by applying the climatology of 2000–2020 estimated GOBM-observation misfits to the GOBMs for 1959–1981. As discussed below, this approach is supported by the fact that much of the skill in LDEO-HPD against independent modern observations is due to the climatological correction. This paper is organized as follows. We present the methods and resulting estimated air-sea $CO_2$ fluxes for 1959–2020. We then briefly examine the resulting estimated flux variability in four basins and globally.

## 2. Methods

The LDEO-HPD data product (Gloege et al., 2022) utilizes the nearly global coverage of satellite sea surface temperature (SST) (Reynolds et al., 2002), sea surface salinity (SSS) (Good et al., 2013), chlorophyll-a (Maritorena et al., 2010), geographic location, time of year, the climatology of mixed layer depth (de Boyer Montégut et al., 2004), and the machine learning algorithm XGBoost (Chen & Guestrin, 2016) to create a nonlinear function between observations and the model-data misfit of surface ocean $pCO_2$. Misfits to observed ocean surface $pCO_2$ of the SOCATv2021 database (Bakker et al., 2016; Sabine et al., 2013) are calculated for each of eight (8) GOBMs (Friedlingstein et al., 2022) (Table 1) separately. The machine learning algorithm is trained to learn the relationship between driver data (SSS, SST, Chl-a, location, MLD, time of year) and observed $pCO_2$ misfit (SOCAT-GOBM) where SOCAT data are available. This algorithm is used to estimate the model-data misfit from the full-coverage driver data at all times and locations. The estimated model-specific misfits are full-coverage, time-varying estimates of how the model $pCO_2$ field should be modified to bring model $pCO_2$ into agreement with the real world. In other machine learning applications for this problem, statistics are used to directly estimate real-world $pCO_2$ at all points in space and time. Our approach is to estimate how each model's $pCO_2$ output needs to be adjusted to better represent reality. To clarify that our final step is to correct the models, we also use the term "corrections" for the model-data misfits.

As shown in the Results, interannual variability of model-data misfits is generally small compared to the climatological mean correction. Thus, we extend LDEO-HPD to the beginning of the model simulations at 1959 using the monthly climatology of the 2000–2020 model-data misfit as the correction. For 1982–2020, the monthly and interannually varying correction is used (Gloege et al., 2022). This misfit is separately calculated for, and applied as a correction to, each of eight GOBMs.

Each of the GOBMs are independently adjusted with its unique correction field that varies at 1° latitude by 1° longitude and monthly for 1982–2020, or climatologically for 1959–1981. The final $pCO_2$ reconstruction is the ensemble mean of the eight corrected GOBM $pCO_2$ estimates (modeled $pCO_2$ + reconstructed correction).

**Table 1**
*Global Ocean Biogeochemical Models, Data Products, and Their Corresponding References*

| Global ocean biogeochemical model | Data product | Reference |
|---|---|---|
| CESM-ETHZ | | Doney et al. (2009) |
| FESCOM2-REcoM | | Gurses et al. (2021) |
| MICOM-HAMOCC (NorESM1-OCv1.2) | | Schwinger et al. (2016) |
| MOM6-COBALT (Princeton) | | Adcroft et al. (2019) |
| MPIOM-HAMOCC6 (MPI) | | Paulsen et al. (2017) |
| NEMO-PlankTOM5 | | Buitenhuis et al. (2013) |
| NEMO-PISCES (IPSL) | | Aumont et al. (2015) |
| NEMO3.6-PISCESv2-gas (CNRM) | | Berthet et al. (2019) |
| | LDEO-HPD | Gloege et al. (2022), this paper |
| | JENA MLS | Rödenbeck et al. (2022) |
| | CSIR ML6 | Gregor et al. (2019) |
| | MPI SOMFFN | Landschützer et al. (2014) |
| | CMEMS FFNN | Denvil-Sommer et al. (2019) |
| | $pCO_2$ Residual | Bennington et al. (2022) |

The complete description of the LDEO-HPD method and the resulting data product can be found in Gloege et al. (2022).

### 2.1. $CO_2$ Flux Calculations

In our analysis of model-data misfits and of reconstruction skill against independent data, we consider $pCO_2$. To assess the global ocean carbon sink associated with these $pCO_2$ estimates, air-sea $CO_2$ exchange must be calculated. We use the same gas transfer velocity, solubility, winds, and ice for LDEO-HPD, other observation-based products, and the GOBMs so that differences in these calculations do not factor into the resulting comparison (Fay et al., 2021). EN4.2.2 salinity (Good et al., 2013); ERA5 winds, sea level pressure, and SST (Bell et al., 2019, 2020); the wind scaling factor for ERA5 (Gregor & Fay, 2021); and Hadley sea ice fractional coverage (Rayner et al., 2003) are used. Unreconstructed coastal areas in data products, which vary in area across the products, are filled with the scaled coastal $pCO_2$ climatology (Landschützer et al., 2020), also following Fay et al. (2021).

Air-sea $CO_2$ flux ($FCO_2$) is estimated using a bulk parameterization (Equation 1),

$$FCO_2 = K_w \cdot K_0 \cdot (1 - ice_{fraction}) \cdot \left(pCO_2^{sea} - pCO_2^{atm}\right) \quad (1)$$

Where $K_w$ is the gas-transfer velocity calculated from wind speeds, scaled to the 16.5 cm/hr 14C bomb flux estimate according to Wanninkhof (1992) and Sweeney et al. (2007) as in Gregor and Fay (2021); $K_0$ is the solubility calculated using salinity and SST; $pCO_2^{atm}$ is the water vapor corrected atmospheric partial pressure of $CO_2$ from CarboScope (Rödenbeck, 2005); and $pCO_2^{sea}$ is the surface ocean $pCO_2$.

Data products which incorporate observations of surface ocean $pCO_2$ include both natural and anthropogenic carbon in the resulting $pCO_2$ and $CO_2$ flux product. This is the net $CO_2$ flux ($F_{net} = F_{natural} + F_{ant}$). Global ocean biogeochemical models exclude the natural outgassing of riverine carbon ($F_{natural}$), which caused net $CO_2$ efflux from the preindustrial ocean (Aumont et al., 2001). To quantify the anthropogenic air-sea $CO_2$ flux, this $F_{natural}$ must be subtracted from our net flux, given that the models have been corrected toward $pCO_2$ observations consistent with $F_{net}$. Quantifying the global air-sea $CO_2$ flux due to decomposition and outgassing of riverine carbon remains uncertain and is the topic of current research. Here, as in Gloege et al. (2022) and Bennington et al. (2022), we use an average of three estimates: Jacobson et al. (2007) (0.45 ± 0.18 PgC/yr), Resplandy et al. (2018) (0.78 ± 0.41 PgC/yr), and Lacroix et al. (2020) (0.23 Pg C/yr +/− an assumed 50% uncertainty). The combined globally integrated efflux due to riverine carbon is 0.49 ± 0.26 Pg C/yr, and thus we remove the efflux

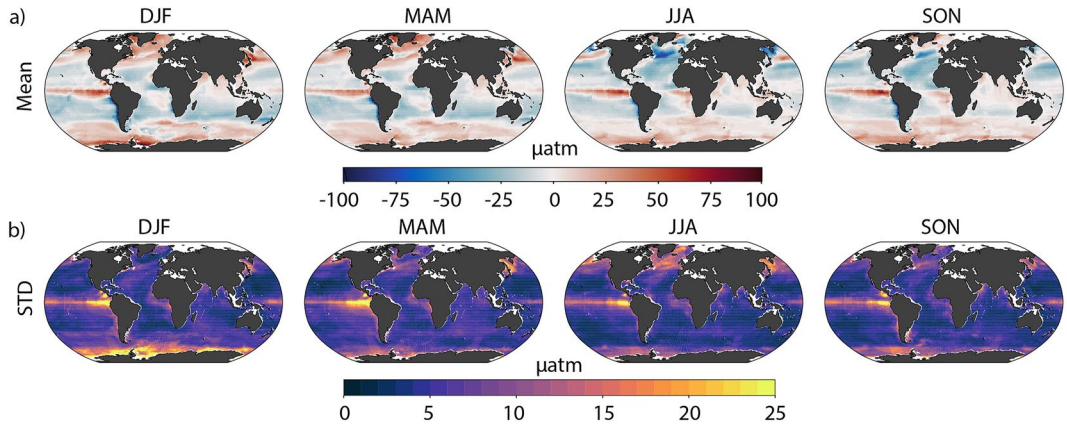


**Figure 1.** (a) Seasonal climatology (2000–2020) of model-data misfit in the Princeton model according to HPD. (b) Standard deviation of model-data misfit over 2000–2020 in the Princeton model, by season.

of 0.49 PgC/yr from the estimated annual air-sea $CO_2$ fluxes calculated using the LDEO-HPD and other data products' $pCO_2$ for comparison of the global flux timeseries.

### 2.2. Box Model

The box model of McKinley et al. (2020) estimates the global-mean air-sea $CO_2$ flux that occurs in response to the observed growth of atmospheric $pCO_2$. It also has the option to include upper ocean heat content anomalies driven by the three most climatically impactful volcanic eruptions of the last 60 years: Agung in 1963, El Chichon in 1982, and Mt Pinatubo in 1991 (Crisp et al., 2022). Comparing air-sea $CO_2$ fluxes estimated by the box model for 1960–2019 allows consideration of flux variability with and without large volcanic influences and puts LDEO-HPD into context with previous comparisons of the box model to observation-based products (McKinley et al., 2020).

## 3. Results

### 3.1. Analysis of Model-Data Misfit

Given the lack of surface ocean $pCO_2$ observations prior to the 1980s, we must determine what corrections (model-data misfits) to apply to the models prior to 1982. Extending the analysis of misfits begun by Gloege et al. (2022), we examine both the climatological misfits and the interannual variability of the misfits for 2000–2020. We choose only 2000–2020 to best capture interannual variability (Bennington et al., 2022), as chlorophyll-a observations do not start until 1998 and a climatology of chlorophyll-a must be used prior (Landschützer et al., 2014). Additionally, $pCO_2$ coverage is better for the decades after 2000 (Bakker et al., 2016), so we have greater confidence in the estimated misfits.

The seasonal climatology and standard deviation of the model-data misfit for the Princeton GOBM is a representative example of the climatological misfit (Figure 1). Mean misfits are large in all seasons in the subpolar, equatorial, and Southern Ocean regions (Figure 1a). Interannual variability in the model-data misfit is quantified as the misfit standard deviation (Figure 1b). Year-to-year changes in misfits are significantly smaller in magnitude than the mean, typically less than 5 μatm. Larger standard deviations can occur during the biologically productive seasons in the subpolar regions and Southern Ocean. The equatorial Pacific exhibits moderate interannual variability in all seasons. This comparison between the magnitude of the climatological and interannually variable misfit are similar across most of the ocean models (Gloege et al., 2022) (Figure S1 in Supporting Information S1).

### 3.2. Climatological Misfits Dominate Improvements

To understand how much skill we could gain in our reconstruction if we used only a climatological correction, comparison to independent data is required. Data for such an assessment do not exist in sufficient number or

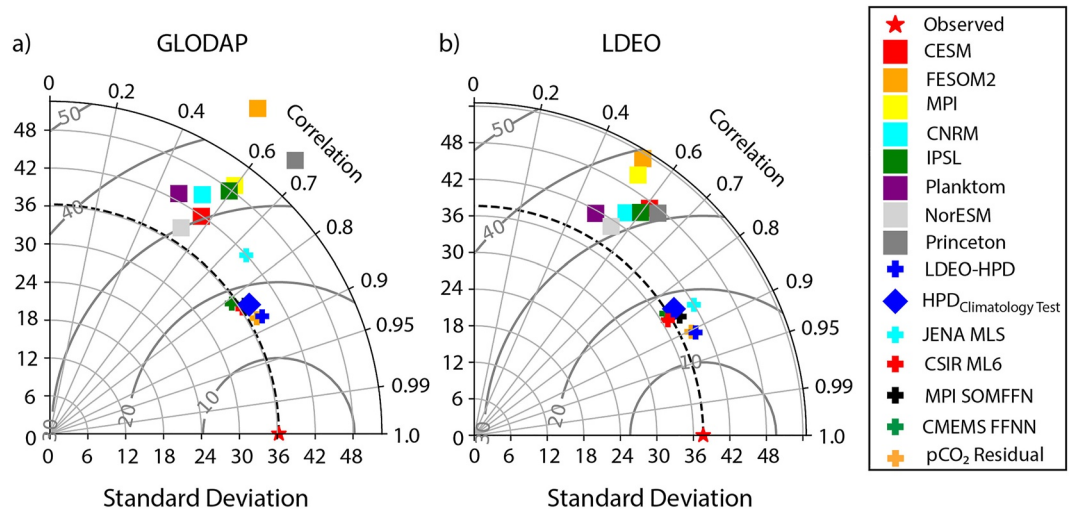


**Figure 2.** Taylor diagrams (Taylor, 2001) depict the skill of each ocean model (squares), previous data products (Table 1, crosses), LDEO-Hybrid Physics Data product (blue cross), and $HPD_{ClimatologyTest}$. The ability to capture observed $pCO_2$ variability for 1990–2020 is evaluated against two global datasets (a) GLODAP and (b) LDEO. The red star indicates the standard deviation of each data set. Distance along the radius represents the ability to capture observed variability (standard deviation). The distance along the circumference depicts correlation with the observations, and gray inlaid circles show unbiased RMSE.

quality prior to the 1980s (Key et al., 2004); and begin to be more available only in the 1990s. To make the needed comparison, we create an alternative reconstruction, $HPD_{ClimatologyTest}$, that applies the above-discussed climatology of the model-data misfit for 2000–2020 to the entire reconstruction period (1959–2020). With $HPD_{ClimatologyTest}$, we can assess the impact of a climatological correction on reconstruction skill by comparison to several datasets.

Figure 2 compares the original uncorrected GOBMs (squares), and five observation-based products (crosses) to GLODAP and LDEO observations for 1990–2020 (red stars). The SOCAT database does not contain these GLODAP or LDEO observations, and thus, this is an independent assessment of reconstruction skill. The observation-based products all have substantially greater skill than the uncorrected GOBMs, indicated by the fact that they lie much closer to the red stars in the Taylor diagrams, which represents a correlation of 1 and the product and data having the same variability.

$HPD_{ClimatologyTest}$ (solid blue diamond) lies almost as close to the observations (red star) as does LDEO-HPD. This leads to an important finding, which is that most of LDEO-HPD's skill is due to the correction of the GOBM's climatological mean state and seasonality (Fay & McKinley, 2021; Hauck et al., 2020; Mongwe et al., 2018) rather than their interannual variability. The additional skill achieved by adding interannual variability to the corrections (Figure 1b) is indicated by the difference between $HPD_{ClimatologyTest}$ and LDEO-HPD, which is modest for GLODAP (Figure 2a) and slightly larger for LDEO (Figure 2b). $HPD_{ClimatologyTest}$ has similar skill to other currently available observation-based products (Table 1), and the additional increment of skill from the interannual correction brings LDEO-HPD closest to the independent observations (Gloege et al., 2022).

These findings support the use of the 2000–2020 climatological correction as the basis for adjusting the GOBMs for 1959–1981. For 1982–2020, interannually varying corrections are used (Section 2).

### 3.3. $CO_2$ Fluxes

Air-sea $CO_2$ fluxes for 1959–2020 from LDEO-HPD, the eight GOBMs, previously published observation-based products, and $HPD_{ClimatologyTest}$ demonstrate a long-term increasing trend punctuated by interannual variability (Figure 3a). There most significant feature of this variability is the slowed growth in uptake during the 1990s (Fay & McKinley, 2013; Hauck et al., 2020; Landschützer et al., 2015; Le Quéré et al., 2007; Lovenduski et al., 2007, 2008).

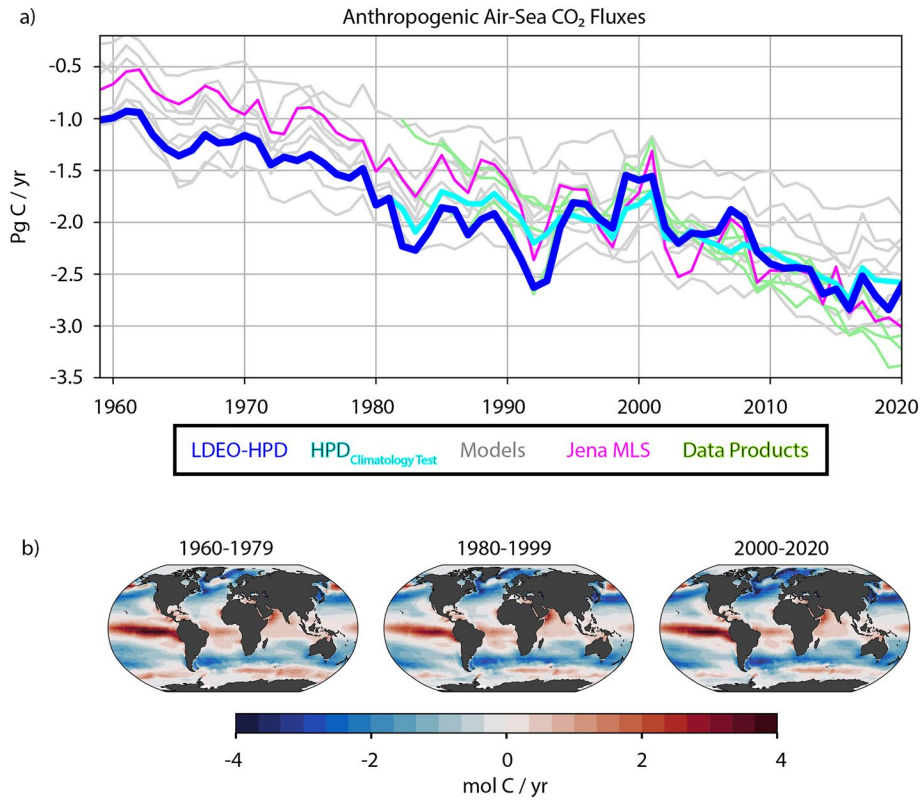


**Figure 3.** (a) Estimated air-sea $CO_2$ fluxes for 1959–2020 (Pg C/yr): LDEO-Hybrid Physics Data product (LDEO-HPD) (blue), $HPD_{ClimatologyTest}$ (cyan), unadjusted global ocean biogeochemical models (gray), Jena MLS (magenta), other observation-based products (green); comparisons shown in separate panels in Figure S1 in Supporting Information S1. $HPD_{ClimatologyTest}$ is identical to LDEO-HPD prior to 1982. (b) Map of mean air-sea $CO_2$ fluxes for 1960–1979, 1980–1999, and 2000–2020 according to LDEO-HPD (mol C/yr).

In LDEO-HPD, interannual variability prior to 1982 is driven by only the GOBMs; only the mean flux and seasonality have been adjusted with climatological model-data misfits (Figure S2 in Supporting Information S1). The adjustment leads to a larger mean flux than most of the GOBMs (Figure 3a). From 1982 onward, the flux in LDEO-HPD has similar anomaly timing to $HPD_{ClimatologyTest}$, but these anomalies are of larger amplitude. These differences are due to the interannually varying adjustments that are possible only during the observed period. This comparison indicates that LDEO-HPD likely underestimates the amplitude of interannual anomalies prior to 1982, which is to be expected when there are no data to directly drive the reconstruction toward extremes (Rödenbeck et al., 2022).

A riverine efflux of carbon is applied to the products so as to estimate the anthropogenic-only global-mean flux, based on the average of several recent estimates (0.49 PgC/yr, Section 2.1). Riverine efflux is highly uncertain and has only been estimated for the long-term mean. Other possible choices of this value would shift the long-term global mean of LDEO-HPD, as well as the other products shown in Figure 3a, by ±0.2 PgC/yr (Section 2.1), but would not impact variability or trends.

Examining the spatial patterns of the mean air-sea carbon dioxide fluxes for each 20 year period in Figure 3b, we see a reduced Pacific equatorial efflux during 1980–1999 compared to the other periods, consistent with the occurrence of multiple strong El Niño events in this period (e.g., 1982–1983, 1997–1998). In the Northern extra-tropics, the sink strengthening over time is evident.

Integrated flux anomalies at each latitude reveal the spatial distribution of interannual anomalies (Figure 4). Consistent with the global timeseries (Figure 3a), the dominant feature is the long-term growth (red to blue) of the ocean carbon sink at all latitudes.

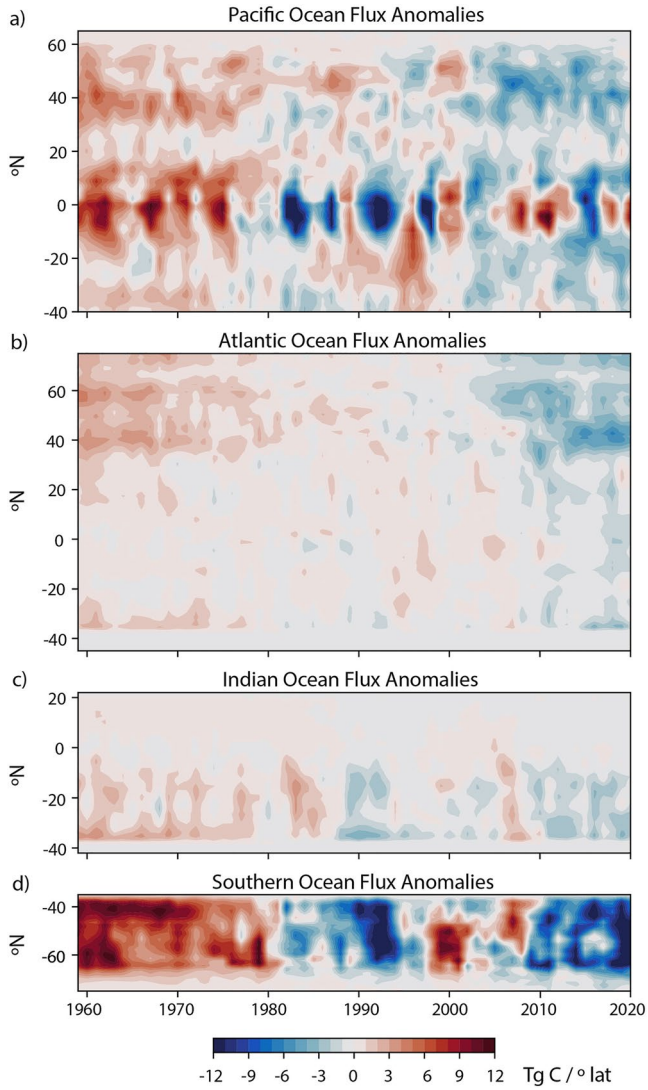


**Figure 4.** Air-sea $CO_2$ flux anomalies in four ocean basins (TgC/yr/°lat).

The Pacific Ocean has large flux variability, with significant anomalies occurring on interannual timescales within the equatorial region as a result of ENSO (McKinley et al., 2004, 2017; Rödenbeck et al., 2022). The Southern Ocean experiences significant carbon sink decadal variations (Gruber et al., 2019; Landschützer et al., 2015, 2016; Le Quéré et al., 2007; Lovenduski et al., 2007, 2008; McKinley et al., 2017; Ritter et al., 2017). Significant negative anomalies (greater uptake) occur in the 1980s to early 1990s, with anomalies of greatest intensity in 1992–1995. After 1997, a strong positive anomaly (reduced uptake) emerges. From 2009 on, the anomaly is again strongly negative in the Southern Ocean. These decadal variations remain after detrending (Figure S4 in Supporting Information S1). In the Atlantic, latitudes north of 40°N have the most intense fluxes. This basin is narrower than the others, and thus has a lower integrated flux and lower amplitude interannual variability. The Indian Ocean exhibits significant variability south of 10°S according to the reconstruction; however the region is particularly sparse in observations to guide the reconstruction, which should increase its uncertainty (Gloege et al., 2021).

Increased uptake occurs in the Pacific and Southern Oceans immediately following the eruptions of Agung (March 1963), El Chichon (March 1982) and Mt. Pinatubo (June 1991); also seen in the detrended fluxes (Figure S2 in Supporting Information S1). In the equatorial Pacific, the El Niño events that tend to follow these eruptions also drive significant flux anomalies (Eddebbar et al., 2019). After El Chichon and Pinatubo, slight negative anomalies also occur in the Southern Hemisphere Atlantic. The globally averaged box model of McKinley et al. (2020) parameterizes these eruptions as upper ocean heat content anomalies; and the estimated fluxes correlate highly with LDEO-HPD (Figure S3d in Supporting Information S1, $r = 0.82$). If the eruptions are neglected, the correlation decreases ($r = 0.64$). When both timeseries are detrended, the correlations remain significant only when the eruptions are included in the box model (with eruptions, $r = 0.51$, $p < 0.05$; without, $r = -0.23$, $p = 0.13$). Thus, both the box model and the spatial patterns of flux anomalies (Figure 4) indicates the potential for large volcanoes to impact interannual variability of the global ocean carbon sink since 1959. A more detailed study of this issue in the LDEO-HPD product will be presented elsewhere.

## 4. Discussion and Conclusions

This work temporally extends the LDEO-HPD data product back in time to begin in 1959. For 1982–2020, model-data misfits are calculated for each model and each month as in Gloege et al. (2022). For 1959–1981, the monthly climatology of this correction for 2000–2020 is applied independently to each of eight GOBMs. Across all years, the final LDEO-HPD $pCO_2$ estimate is the average across the eight corrected models.

While it would be ideal to compare to observations prior to 1982, the TA and DIC measured by the high-quality GEOSECS experiment is subject to sufficient inaccuracies that they were excluded from the GLODAP synthesis products (Key et al., 2004). Thus, we compare to independent data from the modern era. With these comparisons, we find that the substantial improvement over uncorrected GOBMs is due primarily to the correction of the model mean and seasonality; that is, the climatological correction. This finding contrasts to a long standing perception that the primary weakness of GOBMs is their representation of interannual varability (Gruber et al., 2019; Landschützer et al., 2015; Le Quéré et al., 2007), and suggests a need for reassessment of the mechanistic drivers of errors in the models' mean state.

There are significant regional biases in the mean and seasonality of many GOBMs (Fay & McKinley, 2021; Hauck et al., 2020; Mongwe et al., 2018) that this observation-based approach can reduce substantially, bringing the resulting estimates closer to observations (Figure 2). At the same time, this approach can preserve the GOBMs capability to represent interannual variability (Figure 3) that occurs in response to external forcing and internal ocean processes. Hauck et al. (2020) recently concluded that interannual variability is, in fact, well-represented by the GOBMs. By combining the strengths of models and observations with the LDEO-HPD approach, we have developed a robust approach to temporally extend this observation-based product back to 1959.

Another temporal extension on an observation-based product has recently been published, Jena MLS (Rödenbeck et al., 2022). Comparing JENA-MLS to LDEO-HPD, we find the two estimates to be significantly correlated ($r = 0.93$, $p = 0$; $r = 0.66$, $p = 0$ when detrended). The two reconstructions span the range of model flux estimates prior to 1990s (Figure 3b), after which observations better constrain the products. Jena-MLS has a significantly larger estimated trend in the ocean carbon sink over the reconstructed period. However, as discussed by Rödenbeck et al. (2022) (their section A2), Jena-MLS in its current version overestimates the trend; thus, it likely underestimates the sink for the pre-observation decades.

LDEO-HPD indicates that the ocean carbon sink increased over the last 60 years, consistent with the nearly exponential growth of atmospheric $pCO_2$ (McKinley et al., 2020; Raupach et al., 2014; Ridge & McKinley, 2021). Long-term growth of the sink has been punctuated by year-to-year variability. Consistent with many earlier studies, we find that the equatorial Pacific and Southern Ocean have the largest integrated impact on variations of the sink (Hauck et al., 2020; Landschützer et al., 2016; Le Quéré et al., 2000; McKinley et al., 2004, 2017; Resplandy et al., 2015). In the equatorial Pacific, variability is associated with ENSO. The Southern Ocean exhibits strong decadal timescale variations for which both internal and externally forced mechanisms have been proposed (Gruber et al., 2019; Landschützer et al., 2015; McKinley et al., 2020). Better understanding the variability of ocean carbon uptake in the Southern Ocean and across the globe is an important task that can be facilitated by observation-based products such as LDEO-HPD.

## Data Availability Statement

Open Data EN.4.2.2 data were obtained from https://www.metoffice.gov.uk/hadobs/en4/ and are © British Crown Copyright, Met Office (2022), provided under a Non-Commercial Government Licence: http://www.nationalarchives.gov.uk/doc/non-commercial-government-licence/version/2/ Project code (Python) freely available on Github to prepare observational data for the machine learning algorithm, utilize XGBoost, and analyze the resulting reconstructed misfits: https://github.com/valbennington/LDEO_HPD_extension Final reconstructed $CO_2$ fluxes available on Zenodo: https://zenodo.org/record/6647613.

**Acknowledgments**
The authors acknowledge support from NOAA (NA20OAR4310340), NSF through the LEAP STC (Award 2019625) and the Data Science Institute of Columbia University. L.G was sponsored by the National Aeronautics and Space Administration (NASA) through a contract with ORAU. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the National Aeronautics and Space Administration (NASA) or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein. We acknowledge that all data providers and quality controllers who work tirelessly to maintain the SOCAT database. We also acknowledge the ocean biogeochemical modelers who contribute to the Global Carbon Budget for sharing their output with us; and we thank J. Hauck specifically for her leadership on this effort.

## References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., et al. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167–3211. https://doi.org/10.1029/2019MS001726

Aumont, O., Ethé, C., Tagliabue, A., Bopp, L., & Gehlen, M. (2015). PISCES-v2: An ocean biogeochemical model for carbon and ecosystem studies. *Geoscientific Model Development*, *8*, 2465–2513. https://doi.org/10.5194/gmd-8-2465-2015

Aumont, O., Orr, J. C., Monfray, P., Ludwig, W., Amiotte-Suchet, P., & Probst, J.-L. (2001). Riverine-driven interhemispheric transport of carbon. *Global Biogeochemical Cycles*, *15*(2), 393–405. https://doi.org/10.1029/1999GB001238

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

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., et al. (2019). ERA5 monthly averaged data on single levels from 1979 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)*, *10*, 252–266. https://doi.org/10.24381/cds.f17050d7

Bell, B., Hersbach, H., Berrisford, P., Dahlgren, P., Horányi, A., Muñoz Sabater, J., et al. (2020). ERA5 monthly averaged data on single levels from 1950 to 1978 (preliminary version). Copernicus Climate Change Service (C3S) Climate Data Store (CDS).

Bennington, V., Galjanic, T., & McKinley, G. A. (2022). Estimating historical air-sea $CO_2$ fluxes: Incorporating physical knowledge within a data-only approach. *Journal of Advances in Modeling Earth Systems*, In Review. https://doi.org/10.1002/essoar.10510196.1

Berthet, S., Séférian, R., Bricaud, C., Chevallier, M., Voldoire, A., & Ethé, C. (2019). Evaluation of an online grid-coarsening algorithm in a global eddy-admitting ocean biogeochemical model. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1759–1783. https://doi.org/10.1029/2019MS001644

Buitenhuis, E. T., Hashioka, T., & Quéré, C. L. (2013). Combined constraints on global ocean primary production using observations and models. *Global Biogeochemical Cycles*, *27*(3), 847–858. https://doi.org/10.1002/gbc.20074

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *CoRR*, 02754. abs/1603.

Crisp, D., Dolman, H., Tanhua, T., McKinley, G. A., Hauck, J., Eggleston, S., & Aich, V. (2022). How well do we understand the land-ocean-atmosphere carbon cycle? *Reviews of Geophysics*, *60*(2). https://doi.org/10.1029/2021RG000736

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., & Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *Journal of Geophysical Research*, *109*(C12), C12003. https://doi.org/10.1029/2004JC002378

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 $pCO_2$ over the Global Ocean. *Geoscientific Model Development*, *12*(5), 2091–2105. https://doi.org/10.5194/gmd-12-2091-2019

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., Lima, I., Feely, R. A., Glover, D. M., Lindsay, K., Mahowald, N., et al. (2009). Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air–sea $CO_2$ fluxes: Physical climate and atmospheric dust. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8), 640–655. https://doi.org/10.1016/j.dsr2.2008.12.006

Eddebbar, Y. A., Rodgers, K. B., Long, M. C., Subramanian, A. C., Xie, S.-P., & Keeling, R. F. (2019). El niño–like physical and biogeochemical ocean response to tropical eruptions. *Journal of Climate*, *32*(9), 2627–2649. https://doi.org/10.1175/JCLI-D-18-0458.1

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., & McKinley, G. A. (2013). Global trends in surface ocean pCO2 from in situ data. *Global Biogeochemical Cycles*, *27*(2), 541–557. https://doi.org/10.1002/gbc.20051

Fay, A. R., & McKinley, G. A. (2021). Observed regional fluxes to constrain modeled estimates of the ocean carbon sink. *Geophysical Research Letters*, *48*(20), e2021GL095325. https://doi.org/10.1029/2021gl095325

Friedlingstein, P., Jones, M. W., O'Sullivan, M., Andrew, R. M., Bakker, D. C., Hauck, J., et al. (2022). Global carbon budget 2021. *Earth System Science Data*, *14*(4), 1917–2005. https://doi.org/10.5194/essd-14-1917-2022

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Hauck, J., Olsen, A., et al. (2020). Global carbon budget 2020. *Earth System Science Data*, *12*(4), 3269–3340. https://doi.org/10.5194/essd-12-3269-2020

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

Gloege, L., Yan, M., Zheng, T., & McKinley, G. A. (2022). Improved quantification of ocean carbon uptake by using machine learning to merge global models and $pCO_2$ data. *Journal of Advances in Modeling Earth Systems*, *14*(2), e2021MS002620. https://doi.org/10.1029/2021ms002620

Good, S. A., Martin, M. J., & Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *Journal of Geophysical Research: Oceans*, *118*(12), 6704–6716. https://doi.org/10.1002/2013JC009067

Gregor, L., & Fay, A. (2021). SeaFlux: Harmonised sea-air $CO_2$ fluxes from surface pCO2 data products using a standardised approach. *Zenodo*. https://doi.org/10.5281/zenodo.5482547

Gregor, L., Lebehot, A. D., Kok, S., & Monteiro, P. M. S. (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*, 5113–5136. https://doi.org/10.5194/gmd-12-5113-2019

Gruber, N., Landschützer, P., & Lovenduski, N. S. (2019). The variable Southern Ocean carbon sink. *Annual Review of Marine Science*, *11*(1), 159–186. https://doi.org/10.1146/annurev-marine-121916-063407

Gurses, O., Hauck, J., Zeising, M., & Oziel, L. (2021). Global Ocean biogeochemical modelling with FESOM2-REcoM. In *EGU general assembly conference abstracts* (pp. EGU21–14980).

Hauck, J., Zeising, M., Le Quéré, C., Gruber, N., Bakker, D. C. E., Bopp, L., et al. (2020). Consistency and challenges in the ocean carbon sink estimate for the global carbon budget. *Frontiers in Marine Science*, *7*. https://doi.org/10.3389/fmars.2020.571720

Jacobson, A. R., Mikaloff Fletcher, S. E., Gruber, N., Sarmiento, J. L., & Gloor, M. (2007). A joint atmosphere-ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Global Biogeochemical Cycles*, *21*(1). https://doi.org/10.1029/2005GB002556

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., et al. (2004). A global ocean carbon climatology: Results from global data analysis project (glodap). *Global Biogeochemical Cycles*, *18*(4). https://doi.org/10.1029/2004gb002247

Lacroix, F., Ilyina, T., & Hartmann, J. (2020). Oceanic $CO_2$ outgassing and biological production hotspots induced by pre-industrial river loads of nutrients and carbon in a global modeling approach. *Biogeosciences*, *17*(1), 55–88. https://doi.org/10.5194/bg-17-55-2020

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. (2014). Recent variability of the global ocean carbon sink. *Global Biogeochemical Cycles*, *28*(9), 927–949. https://doi.org/10.1002/2014GB004853

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., 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

Le Quéré, C., Orr, J., Monfray, P., Aumont, O., & Madec, G. (2000). Interannual variability of the oceanic sink of $CO_2$ from 1979 through 1997. *Global Biogeochemical Cycles*, *14*(4), 1247–1265. https://doi.org/10.1029/1999gb900049

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., 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., Gruber, N., Doney, S. C., & Lima, I. D. (2007). Enhanced $CO_2$ outgassing in the Southern Ocean from a positive phase of the southern annular mode. *Global Biogeochemical Cycles*, *21*(2), GB2026. https://doi.org/10.1029/2006gb002900

Maritorena, S., d'Andon, O. H. F., Mangin, A., & Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical model: Characteristics, benefits and issues. *Remote Sensing of Environment*, *114*(8), 1791–1804. https://doi.org/10.1016/j.rse.2010.04.002

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., 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). https://doi.org/10.1029/2003gb002179

Mongwe, N., Vichi, M., & Monteiro, P. (2018). The seasonal cycle of $pco_2$ and $CO_2$ fluxes in the Southern Ocean: Diagnosing anomalies in cmip5 Earth system models. *Biogeosciences*, *15*(9), 2851–2872. https://doi.org/10.5194/bg-15-2851-2018

Paulsen, H., Ilyina, T., Six, K. D., & Stemmler, I. (2017). Incorporating a prognostic representation of marine nitrogen fixers into the global ocean biogeochemical model HAMOCC. *Journal of Advances in Modeling Earth Systems*, *9*(1), 438–464. https://doi.org/10.1002/2016ms000737

Peters, G. P., Quéré, C. L., Andrew, R. M., Canadell, J. G., Friedlingstein, P., Ilyina, T., et al. (2017). Towards real-time verification of $CO_2$ emissions. *Nature Climate Change*, *7*(12), 848–850. https://doi.org/10.1038/s41558-017-0013-9

Raupach, M. R., Davis, S. J., Peters, G. P., Andrew, R. M., Canadell, J. G., Ciais, P., et al. (2014). Sharing a quota on cumulative carbon emissions. *Nature Climate Change*, *4*(10), 873–879. https://doi.org/10.1038/nclimate2384

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., & Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14), 4407. https://doi.org/10.1029/2002JD002670

Resplandy, L., Keeling, R. F., Rödenbeck, C., Stephens, B. B., Khatiwala, S., Rodgers, K. B., et al. (2018). Revision of global carbon fluxes based on a reassessment of oceanic and riverine carbon transport. *Nature Geoscience*, *11*(7), 504–509. https://doi.org/10.1038/s41561-018-0151-3

Resplandy, L., Séférian, R., & Bopp, L. (2015). Natural variability of $CO_2$ and $O_2$ fluxes: What can we learn from centuries-long climate models simulations? *Journal of Geophysical Research: Oceans*, *120*(1), 384–404. https://doi.org/10.1002/2014jc010463

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *Journal of Climate*, *15*(13), 1609–1625. https://doi.org/10.1175/1520-0442(2002)015<1609:aiisas>2.0.co;2

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., Iida, Y., Jones, S., et al. (2017). Observation-based trends of the Southern Ocean carbon sink. *Geophysical Research Letters*, *44*(24), 12–339. https://doi.org/10.1002/2017gl074837

Rödenbeck, C. (2005). *Estimating $CO_2$ sources and sinks from atmospheric mixing ratio measurements using a global inversion of atmospheric transport (Tech. Rep. No. 6)*. Max Planck Institute for Biogeochemistry. Retrieved from http://www.bgc-jena.mpg.de/CarboScope/

Rödenbeck, C., DeVries, T., Hauck, J., Le Quéré, C., & Keeling, R. (2022). Data-based estimates of interannual sea–air $CO_2$ flux variations 1957–2020 and their relation to environmental drivers. *Biogeosciences Discussions*, *19*(10), 2627–2652. https://bg.copernicus.org/articles/19/2627/2022/

Rödenbeck, C., Keeling, R. F., Bakker, D. C. E., Metzl, N., Olsen, A., Sabine, C., & Heimann, M. (2013). Global surface-ocean $pCO_2$ and sea-air $CO_2$ flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Science*, *9*(2), 193–216. https://doi.org/10.5194/os-9-193-2013

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., et al. (2013). Surface ocean $CO_2$ atlas (socat) gridded data products. *Earth System Science Data*, *5*(1), 145–153. https://doi.org/10.5194/essd-5-145-2013

Schwinger, J., Goris, N., Tjiputra, J. F., Kriest, I., Bentsen, M., Bethke, I., et al. (2016). Evaluation of NorESM-OC (versions 1 and 1.2), the ocean carbon-cycle stand-alone configuration of the Norwegian Earth System Model (NorESM1). *Geoscientific Model Development*, *9*(8), 2589–2622. https://doi.org/10.5194/gmd-9-2589-2016

Sweeney, C., Gloor, E., Jacobson, A. R., Key, R. M., McKinley, G., Sarmiento, J. L., & Wanninkhof, R. (2007). Constraining global air-sea gas exchange for $CO_2$ with recent bomb 14C measurements. *Global Biogeochemical Cycles*, *21*(2). https://doi.org/10.1029/2006gb002784

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Wanninkhof, R. (1992). Relationship between wind speed and gas exchange over the ocean. *Journal of Geophysical Research*, *97*(C5), 7373–7382. https://doi.org/10.1029/92JC00188