AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2017GL074837

Key Points:
- We compare decadal trends of an ensemble of nine observation-based Delta pCO_2 products in the Southern Ocean
- Eight out of nine products reveal a weakening sink trend calculated for the 1990s, and a strengthening sink in the 2000s
- The spatial pattern of the multiproduct mean trends is rather uniform in both periods

Supporting Information:
- Supporting Information S1
- Data Set S1
- Data Set S2
- Data Set S3

Correspondence to:
P. Landschützer,
peter.landschuetzer@mpimet.mpg.de

Citation:
Ritter, R., Landschützer, P., Gruber, N., Fay, A. R., Iida, Y., Jones, S., … Zeng, J. (2017). Observation-based trends of the Southern Ocean carbon sink. *Geophysical Research Letters*, *44*, 12,339–12,348. https://doi.org/10.1002/2017GL074837

Received 6 JUL 2017
Accepted 28 NOV 2017
Accepted article online 4 DEC 2017
Published online 26 DEC 2017

Observation-Based Trends of the Southern Ocean Carbon Sink

R. Ritter[1], P. Landschützer[1,2], N. Gruber[1], A. R. Fay[3], Y. Iida[4], S. Jones[5,6], S. Nakaoka[7], G.-H. Park[8], P. Peylin[9], C. Rödenbeck[10], K. B. Rodgers[11], J. D. Shutler[12], and J. Zeng[7]

[1]Institute of Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland, [2]Max Planck Institute for Meteorology, Hamburg, Germany, [3]Lamont-Doherty Earth Observatory, Columbia University, New York, NY, USA, [4]Global Environment and Marine Department, Japan Meteorological Agency, Tokyo, Japan, [5]Geophysical Institute, University of Bergen, Bergen, Norway, [6]Bjerknes Centre for Climate Research, Bergen, Norway, [7]National Institute for Environmental Studies, Tsukuba, Japan, [8]East Sea Research Institute, Korea Institute of Ocean Science and Technology, Uljin, Korea, [9]Laboratoire des Sciences du Climat et de l'Environnement, Gif sur Yvette, France, [10]Max Planck Institute for Biogeochemistry, Jena, Germany, [11]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA, [12]College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom

Abstract The Southern Ocean (SO) carbon sink has strengthened substantially since the year 2000, following a decade of a weakening trend. However, the surface ocean pCO_2 data underlying this trend reversal are sparse, requiring a substantial amount of extrapolation to map the data. Here we use nine different pCO_2 mapping products to investigate the SO trends and their sensitivity to the mapping procedure. We find a robust temporal coherence for the entire SO, with eight of the nine products agreeing on the sign of the decadal trends, that is, a weakening CO_2 sink trend in the 1990s (on average 0.22 $\pm$ 0.24 pg C yr^{-1} $decade^{-1}$), and a strengthening sink trend during the 2000s (-0.35 ± 0.23 pg C yr^{-1} $decade^{-1}$). Spatially, the multiproduct mean reveals rather uniform trends, but the confidence is limited, given the small number of statistically significant trends from the individual products, particularly during the data-sparse 1990–1999 period.

Plain Language Summary The Southern Ocean plays an important role in regulating Earth's climate as it takes up a substantial amount of carbon dioxide from the atmosphere, thereby limiting the effect of global warming. However, this part of the global ocean is also the least well observed and observational data are sparse. Therefore, to study Southern Ocean carbon uptake, data interpolation methods are used to estimate the variability of the carbon uptake from the few existing observations. This poses the question on how reliable these estimates are. The Surface Ocean CO_2 Mapping intercomparison project aims to do exactly that, that is, test how reliable current estimates are by comparing results from different methods. Here we compare the results from nine data interpolation methods in the Southern Ocean from 1990 to 2010 and find a broad and encouraging agreement regarding decadal carbon uptake signals, whereas a spatially more refined analysis reveals much less agreement locally, illustrating the need to continue the measurement effort in the Southern Ocean.

1. Introduction

The Southern Ocean (SO), defined here as, the region south of 30°S, is the ocean's strongest sink for anthropogenic carbon dioxide (CO_2), that is, for the CO_2 emitted to the atmosphere as a result of human activities (Frölicher et al., 2015; Gruber et al., 2009; Khatiwala et al., 2013; Sabine et al., 2004). Averaged over the last three decades, the SO sink for anthropogenic CO_2 of about 1 pg C yr^{-1} (Frölicher et al., 2015; Mikaloff Fletcher et al., 2006) accounts for about 40% of the ocean carbon sink as a whole (Le Quéré et al., 2016). Despite its crucial role, the SO is among the least sampled oceans (Bakker et al., 2014; Landschützer et al., 2015; Majkut et al., 2014) and also perhaps the least understood (Lenton et al., 2013).

A considerable debate exists in the literature regarding the recent variability of this SO carbon sink. The first studies conducted in the early 2000 proposed that the sink weakened substantially during the latter part of the twentieth century, deviating strongly from a strengthening trend that is expected as a result of the continuing increase in atmospheric CO_2 (Lenton et al., 2013; Le Quéré et al., 2007; Lovenduski et al., 2007, 2008).

These studies, which relied on forward ocean models and atmospheric inversions, attributed the weakening trend to a southward shift and strengthening of the westerly winds that caused an increase of the upwelling of inorganic carbon from deeper layers of the ocean. This drives up the surface ocean partial pressure of CO_2, that is, pCO_2, the most important quantity controlling the direction and strength of the air-sea CO_2 flux (Sarmiento & Gruber, 2006). Some of these studies also pointed out that this weakening trend, that is, a trend toward a "saturation of the SO sink" might be a harbinger of the future, as the wind shift that occurred in the 1990s is congruent with the changes expected in a future warm world (Le Quéré et al., 2007).

More recent studies, largely based on observations of the surface ocean pCO_2 (Fay et al., 2014; Landschützer et al., 2015; Munro et al., 2015; Xue et al., 2015), have found that the sink trend appears to have reversed around 2000, leading to a "reinvigoration of the SO carbon sink" in the first decade of this millennium. In fact, Landschützer et al. (2015) showed that by the early 2010s, the sink strength was back to where it was expected based on the atmospheric CO_2 trend even though the westerly winds continued to remain stronger than usual, implying that the SO carbon sink should have remained weak. The driver for this reversal of the sink strength appears to be a combination of changes in sea surface temperature and a weakening of the upper-ocean overturning circulation (DeVries et al., 2017), while the wind-driven upwelling plays a lesser role (Landschützer et al., 2015). This reversal suggests the presence of rather strong intrinsic (natural) variability in the SO carbon sink that is superimposed on any anthropogenic trend. This has notable consequences for our understanding of the ocean carbon cycle, since the decadal variations in the SO carbon sink dominate the variations of the global ocean carbon sink (Landschützer et al., 2016). These variations are also critically important when assessing the time of emergence of any anthropogenic trends in the ocean carbon cycle (e.g., McKinley et al., 2016).

Even though these recent studies benefited greatly from the growing number of surface ocean observations and the substantial community-led effort to combine them in quality-controlled and publicly available data collections (Bakker et al., 2014, 2016; Pfeil et al., 2013; Takahashi et al., 2014), the SO remains sparsely observed. This requires careful attention to how data are aggregated or extrapolated in order to determine reliable trends over the entire region. Fay et al. (2014) averaged the data across a number of specified biomes, taking into account the spatial mean offset between the location of the observation and the biome average. Landschützer et al. (2015) employed a sophisticated data mapping method to extrapolate the observations to a continuous grid in time and space. Concretely, they employed a two-step neural network-based method, consisting of a self-organizing map (SOM) method to first identify a series of biogeochemical provinces and then a feed-forward network (FFN) method to identify nonlinear relationships between a series of independent variables (such as sea surface temperature and mixed layer depth) within each province (Landschützer et al., 2013). Landschützer et al. (2015) contrasted the results of this ETH-SOMFFN method with another independent method that uses a local data assimilation approach (Rödenbeck et al., 2013). They found good agreement between the two methods when they analyzed the decadal trends in the SO, even though these two methods deviate substantially from each other on a year-to-year basis.

Here we aim to expand on this comparison and to test the coherence of observation-based decadal trend estimates for the SO across a much larger set of estimates. For this purpose, we will be making use of the ensemble of estimates collected by the Surface Ocean pCO_2 Mapping intercomparison (SOCOM) initiative (see Rödenbeck et al., 2015). This ensemble consists of the results of independent and largely complementary mapping methods for the sea surface pCO_2. This permits us to conduct an intercomparison study of the products and to assess the congruence of the decadal carbon sink trends in one of the most important, but also least sampled, oceanic regions.

2. Materials and Methods

2.1. Observation-Based pCO_2 Mapping Methods

We use a subset of 9 out of the 14 pCO_2 mapping products collected by the SOCOM project (Rödenbeck et al., 2015). This selection is based on our aim to analyze decadal trends over a period of at least 2 decades. This eliminates four products due to the shortness of their records and one product owing to it not having been designed to reconstruct trends (based on Table 4 of Rödenbeck et al., 2015). The following nine members of the SOCOM ensemble are retained for the analysis: ETH-SOMFFN, Jena-MLS, PU-MCMC, UEA-SI, AOML-EMP, CARBONES-NN, NIES-NN, CU-SCSE and JMA-MLR (Iida et al., 2015; Jones et al., 2015; Landschützer et al., 2016; Majkut et al., 2014; Park et al., 2010; Rödenbeck et al., 2013; Zeng et al., 2014). Here we briefly summarize the

most relevant information regarding this ensemble. For a more detailed description please refer to Rödenbeck et al. (2015).

All ensemble members extrapolate the available sea surface pCO_2 observations both in time and space, that is, they close data gaps where no observations exist, thereby providing continuous pCO_2 maps. These maps then serve to estimate the partial pressure difference between the ocean and the atmosphere ($\Delta pCO_2 = pCO_2^{sea} - pCO_2^{atm}$), from which the air-sea flux can be computed. Six out of the nine selected products are based on the Surface Ocean CO_2 Atlas (SOCAT) data set (Bakker et al., 2014, 2016; Pfeil et al., 2013; Sabine et al., 2013) or its gridded subproduct. The remaining three use various releases of the Lamont-Doherty Earth Observatory (LDEO) data set (Takahashi et al., 2014). While there is a large overlap between these databases, the number of measurements used for the buildup of the individual methods differ by almost 1 order of magnitude (see Table 3 in Rödenbeck et al., 2015 for more details). Thus, in addition to the methodological differences (see supporting information), the two databases provide a second source of difference between the individual methods. The complementary nature of this ensemble allows us to identify ΔpCO_2 trends independent of the chosen data extrapolation or mapping method.

We analyze the decadal trends in these data products using the output from January 1990 to December 2009, the longest common period in the retained ensemble members. This period differs slightly from the analysis period in Landschützer et al. (2015), but we do not expect that this affects the results in a substantial manner. Prior to analysis, all product suppliers interpolated their pCO_2 products to the same monthly $1° \times 1°$ grid using their own interpolation techniques. The analyses are computed both on the resulting common grid as well as on a regionally aggregated level. To this end, we use the biome definitions of Fay and McKinley (2014), who divided the global open ocean into biomes based on four climatological criteria: maximum mixed layer depth, spring/summer chlorophyll concentration, sea surface temperature, and sea ice fractional coverage. Our analysis is conducted within their three SO biomes (subtropical seasonally stratified (SO STSS), subpolar seasonally stratified (SO SPSS) and ice (SO ICE) - see Figure 3).

2.2. Evaluation

The selected nine methods have been extensively validated and tested in their individual publications (see Rödenbeck et al., 2015). We thus investigate here only the temporal residuals to assess whether any method exhibits a systematic bias that could impact our trend estimates. To this end, we calculate the difference between the individual maps and the gridded observations from the SOCATv2 (Bakker et al., 2014) data set ($pCO_2^{est} - pCO_2^{SOCATv2}$, averaged over the SO) (Sabine et al., 2013). Further, we analyze the mean biases of 5 year intervals as the data sparseness of the SO precludes a year-to-year assessment in the early 1990s.

The resulting biases and spreads (within the 95% confidence intervals) stay around zero with no clear indication of any trend in the bias for the selected methods with the exception of the NIES-NN product (Figure S2). But its trends are not significantly different from those calculated from the other methods, suggesting that it can be retained in our multiproduct ensemble. Figure S2 also reveals that some methods tend to slightly underestimate the pCO_2 values (e.g., JMA-MLR and CARBONES-NN) or overestimate them (e.g., AOML-MLR and CU-SCSE).

As the majority of the selected products were created using the SOCATv2 data set (see Table 3 in Rödenbeck et al., 2015), this test (or another test using the LDEO database) does not provide an independent assessment of each method, but rather a first-order test of whether the selected methods are suitable for a trend analysis. Despite all discrepancies, we conclude that the selected methods are capable of reproducing the observations in time in absence of any systematic trend in the bias or spread, based on the available observations.

2.3. Trend Calculations

The trends for the individual products are computed on a pixel by pixel basis, using the deseasonalized air-sea difference in pCO_2, that is, $\Delta pCO_2 = pCO_2^{sea} - pCO_2^{atm}$. The atmospheric pCO_2 is estimated using the marine boundary layer product of GLOBALVIEW-CO2 (GLOBALVIEW-CO2, 2013) for the atmospheric mixing ratio and the National Centers for Environmental Prediction (NCEP) monthly mean sea level pressure (Kalnay et al., 1996) for the atmospheric pressure, assuming 100% relative humidity. The data are deseasonalized using a 12 month running average filter, which has the advantage that it also removes any higher-frequency variability. We then use the resulting smoothed ΔpCO_2 time series to determine the trends for each $1° \times 1°$ pixel and for the two time periods of interest, that is, 1990 through 1999 and 2000 through 2009, respectively, using the slope of a normal least squares regression fit for the individual methods. The slopes of the linear regressions are regarded as significantly different from 0 when the applied F statistics show a p value < 0.05. We substantially reduce

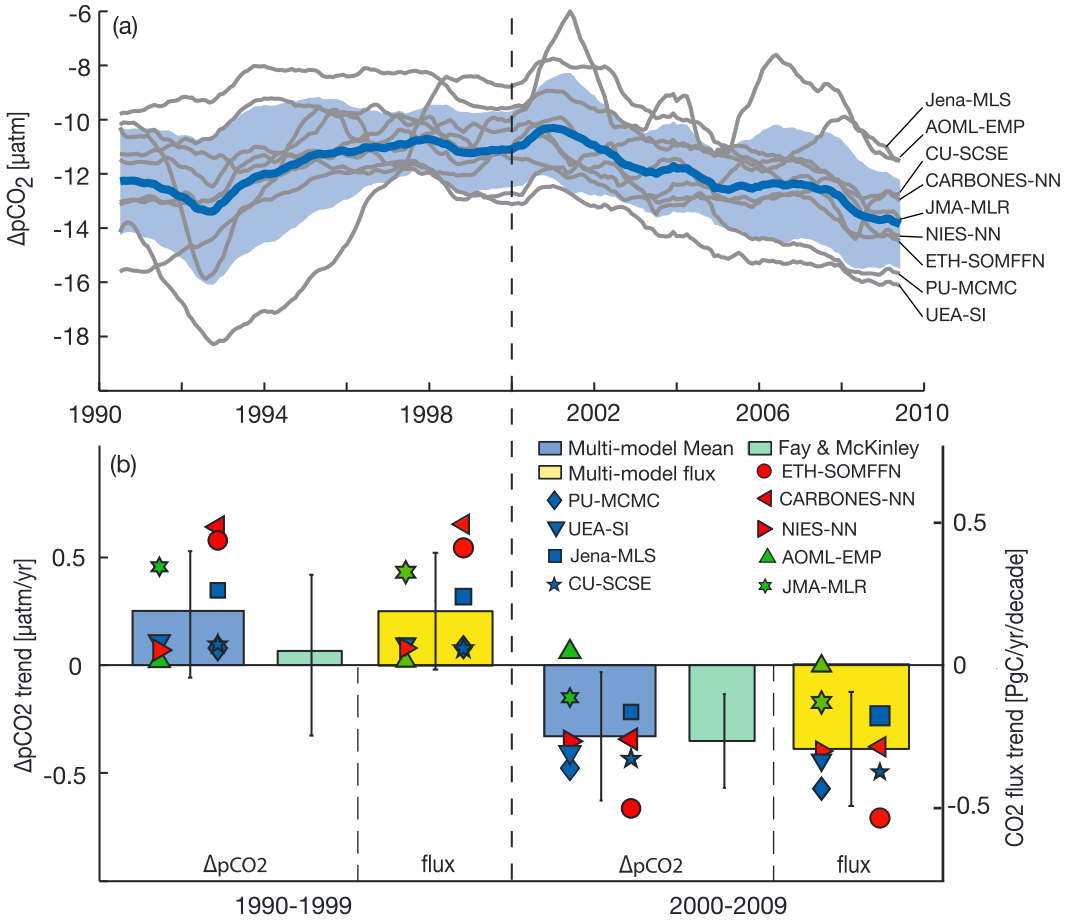

Figure 1. (a) SO deseasonalized time series of monthly ΔpCO_2 (pCO_2^{sea} - pCO_2^{atm}) of the selected nine mapping methods (grey), the multiproduct mean (blue) and its standard deviation (blue shading) in microatmospheres for the period from July 1990 to June 2009. (b) Estimated mean trends (ice-weighted) in µatm/yr of the nine mapping methods and the resulting multiproduct mean compared to the results reported by Fay & McKinley (2013) for the SO for the first and second decade. Also shown are the corresponding air-sea flux of CO_2 trends. Negative values indicate an increasing uptake of CO_2. The error bars represent the standard deviation (the standard deviation of Fay & McKinley are based on confidence intervals of 68.3%).

the degrees of freedom in this test, that is, use an effective number of degrees of freedom of 9, instead of the 108 monthly entries that remained after the running average filter removed the first and last 6 months of the original time series. This choice, which essentially means that we are using only the annual mean data, is a consequence of the smoothed time series having strong serial autocorrelation at any timescale shorter than 12 months. Furthermore, we only take into account ice-free ocean areas, determined from the ice-free fraction given by the sea ice product of Rayner et al. (2003). The trends for the aggregated regions, that is, biomes or for the entire SO, are computed by using an ice-free-weighted average. We further calculate the air-sea CO_2 flux trends corresponding to the observed ΔpCO_2 trends using high-resolution winds from the ERA-Interim (Dee et al., 2011) wind speed product as well as a quadratic gas formulation scaled to fit a long-term mean gas transfer velocity of 16 cm/h (Wanninkhof et al., 2013) (see Figure 1b).

The multiproduct mean and its spread are given by the arithmetic mean of all nine methods at each grid point and its standard deviation. The multiproduct mean trends and their spread are computed by first averaging the oceanic pCO_2 from the selected nine products for each grid point in space and time and then applying a York fit (York et al., 2004), that is, a fit accounting for the error stemming from the multimodel average in our regression model. We choose here to use the arithmetic mean rather than the median in accordance to the standard procedures used by the Intergovernmental Panel on Climate Change for multimodel analyses (Stocker et al., 2010). In practice the difference between the median and mean turns out to be small.

3. Results and Discussions

3.1. Decadal Trends

Our results reveal relatively congruent ΔpCO_2 trends in the SO across the SOCOM ensemble with the 1990s characterized by a reduction of the air-sea pCO_2 gradient, implying a weakening carbon sink, and the first decade of this millennium characterized by an increase of the air-sea pCO_2 gradient, implying a strengthening sink (Figure 1a). But there is a substantial amount of variance around the temporal evolution of the multiproduct mean with some products having very strong decadal trends and little year-to-year variability and other products having very low decadal trends but stronger year-to-year variations. The individual estimates also differ substantially with regard to their mean, with differences in the (deseasonalized) monthly mean ΔpCO_2 of more than 10 µatm in 1993 between individual methods. Some of these differences might be due to differences in the assumptions made by the individual method developers with regard to the upper-ocean thermal and chemical structures to take into account that the majority of the pCO_2 observations are taken at 5 m depth and below (Woolf et al., 2016). But the more likely source of these differences are differences in the methodologies of the data interpolation methods. Attributing these differences to the various causes is, however, well beyond the scope of this paper. Even accounting for the large spread, the multiproduct mean ΔpCO_2 shows a clear maximum around the year 2000, with more negative values before and after.

An analysis in the form of linear trends for each decade and for each product demonstrates this more quantitatively (Figure 1b). All nine selected mapping methods agree that the 1990s are characterized by a positive trend of the ΔpCO_2, although the multiproduct mean of 0.26 ± 0.30 µatm yr^{-1} is not distinguishable from zero. Eight of the nine selected methods further agree that this trend reversed in the 2000s, with a multiproduct mean of -0.33 ± 0.30 µatm yr^{-1}. These ΔpCO_2 trends for the nine methods imply a carbon sink that decreased from 1990 through 1999 at a rate between 0.01 to 0.41 pg C yr^{-1} $decade^{-1}$ (1990–1999) with a multiproduct mean flux trend of 0.22 ± 0.24 pg C yr^{-1} $decade^{-1}$ (Figure 1b). The post-2000 rebound in the carbon sink is uniformly predicted to occur with the reinvigoration trend varying between −0.01 to −0.53 pg C yr^{-1} $decade^{-1}$ with a multiproduct mean flux trend of -0.35 ± 0.23 pg C yr^{-1} $decade^{-1}$. The strongest trends and also strongest trend changes are reconstructed by the ETH-SOMFFN and CARBONES-NN methods, while on the other side of the spectrum it is the AOML-EMP method that diagnoses essentially no trend and also no change in the trend between the two decades.

Results based on the analysis of Fay and McKinley (2013) support our multiproduct mean trends of ΔpCO_2 as well as the changes in these trends (Figure 1b). Unlike the SOCOM methods, the Fay and McKinley (2013) estimate does not rely on any data interpolation algorithm to calculate trends but also use data corrections to avoid spatial aliasing (Figure 1b). While the reported slope from their analysis during the 1990s is small and not significant, the reversal of the trend after 2000 is well captured and also agrees well in magnitude with the SOCOM-based multiproduct mean. This implies that the trends captured by the SOCOM products are contained in the noninterpolated original data, giving further confidence in the trend reversal inferred from the interpolated products.

This result, despite slightly different time intervals, strongly supports the conclusions of Landschützer et al. (2015), which were based on the ETH-SOMFFN and Jena-MLS methods only. While the trends and trend changes diagnosed by the ETH-SOMFFN method tend to be substantially larger than those of the multiproduct mean, those of the Jena-MLS method are very close to that mean.

3.2. Robustness

The multiproduct mean was computed without consideration of how well the models are able to fit the available observations, and whether these differences have an impact on the estimated trends and their changes. To assess whether this is the case, we plot in Figure 2 the decadal trends of each method as a function of its (a) bias, (b) quality of fit represented by the root-mean-square error (RMSE) and (c) the difference between the mean biases of the two decades, all calculated from the gridded SOCATv2 (Bakker et al., 2014) data set. The individual methods are combined into three color groups representing common methodologies, that is, linear regression, nonlinear regression and statistical interpolation, and assimilation methods, recognizing that within each of these groups, the degree of complexity differs substantially.

While Figure 2a illustrates a wide bias range, with the 1990s mainly biased high and the 2000s biased low, it reveals that the magnitude of the decadal mean bias shows no clear effect on the direction and strength of the trend. The same conclusion applies to RMSE (Figure 2b), that is, the sign of the trend is independent of the RMSE. This is valid across all methods or groups of methods. The lower RMSE during the 1990s compared

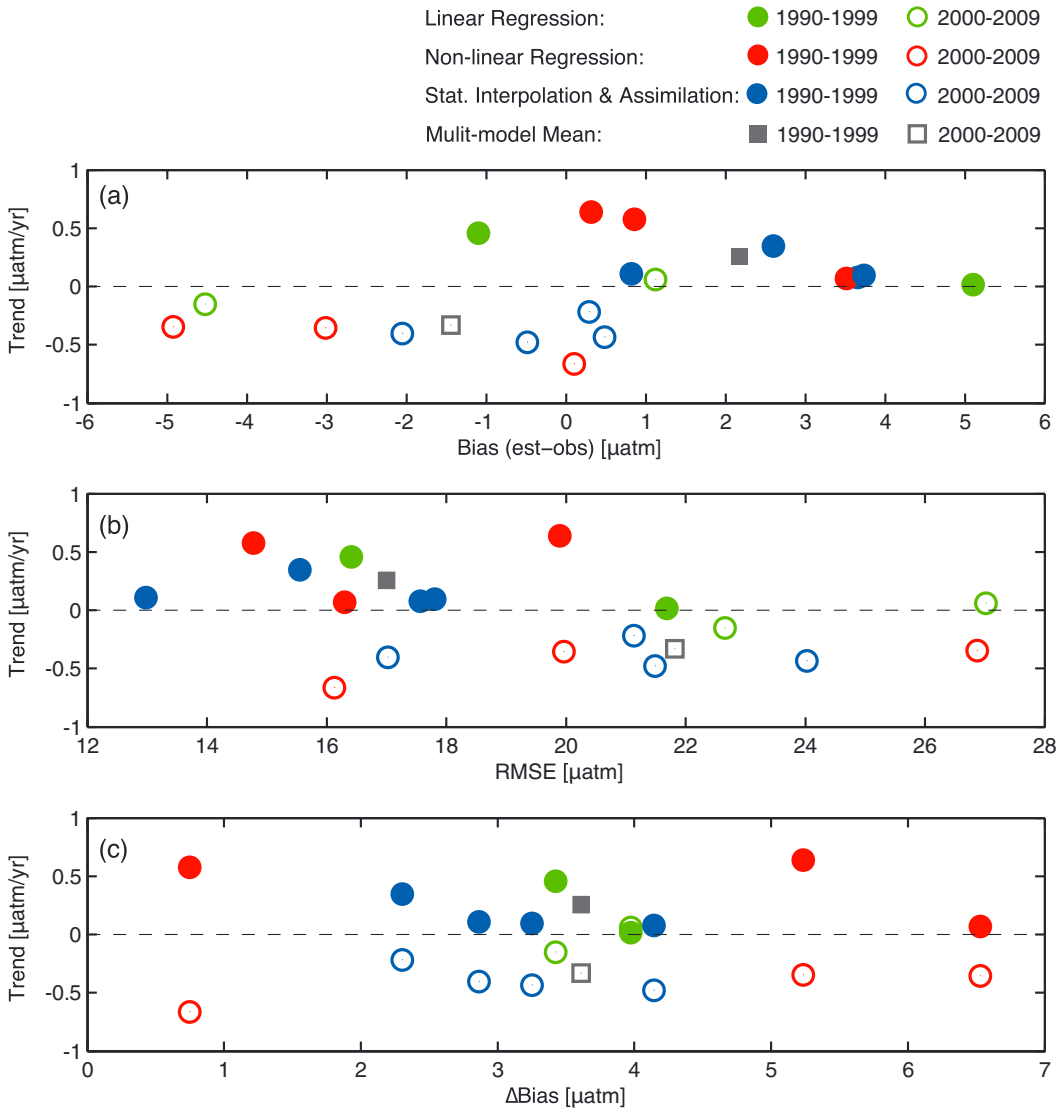

Figure 2. Estimated decadal mean trends (ΔpCO_2) of all nine methods in µatm/yr (1990–1999: filled circles; 2000–2009: open circles) and the multimodel mean trends (1990–1999: filled squares; 2000–2009: open squares) versus (a) the corresponding mean bias ($bias^{est}$ - $bias^{SOCATv2}$ of pCO_2) and (b) versus the corresponding RMSE (of pCO_2) as well as (c) versus the Δmean bias (mean $bias^{1990-1999}$ − mean $bias^{2000-2009}$ of pCO_2) in microatmospheres of all nine mapping methods and the Δmultimodel mean bias of pCO_2 averaged over the SO. The selected nine methods are consolidated into three color groups: Green represents all linear regression methods (AOML-EMP, JMA-MLR), red stands for the nonlinear regression mapping methods (ETH-SOMFFN, NIES-NN, and CARBONES-NN), and blue indicates the statistical interpolation and assimilation methods (Jena-MLS, UEA-SI, CU-SCSE, and PU-MCMC).

to the 2000s is mainly due to the latter period sampling a wider range of data, leading to a higher degree of variance, which causes also more challenges for the individual methods.

A further test for the robustness of the finding is whether the trend and particularly the change in the trend from one decade to the other depends on the change in the quality of the fit across the two decades. But as was the case for the other metrics, Figure 2c shows that there is no clear relationship between either the decadal trends or their change and the decadal difference in bias.

This result highlights that the different SOCOM products—despite their methodological differences, the different sets of observations used for the buildup of the results, and their rather large spread of mismatches vis-à-vis the observations—agree remarkably well with regard to the decadal fluctuations of the SO carbon sink. Yet there are substantial unresolved differences in the exact magnitude of the diagnosed trends. The lack

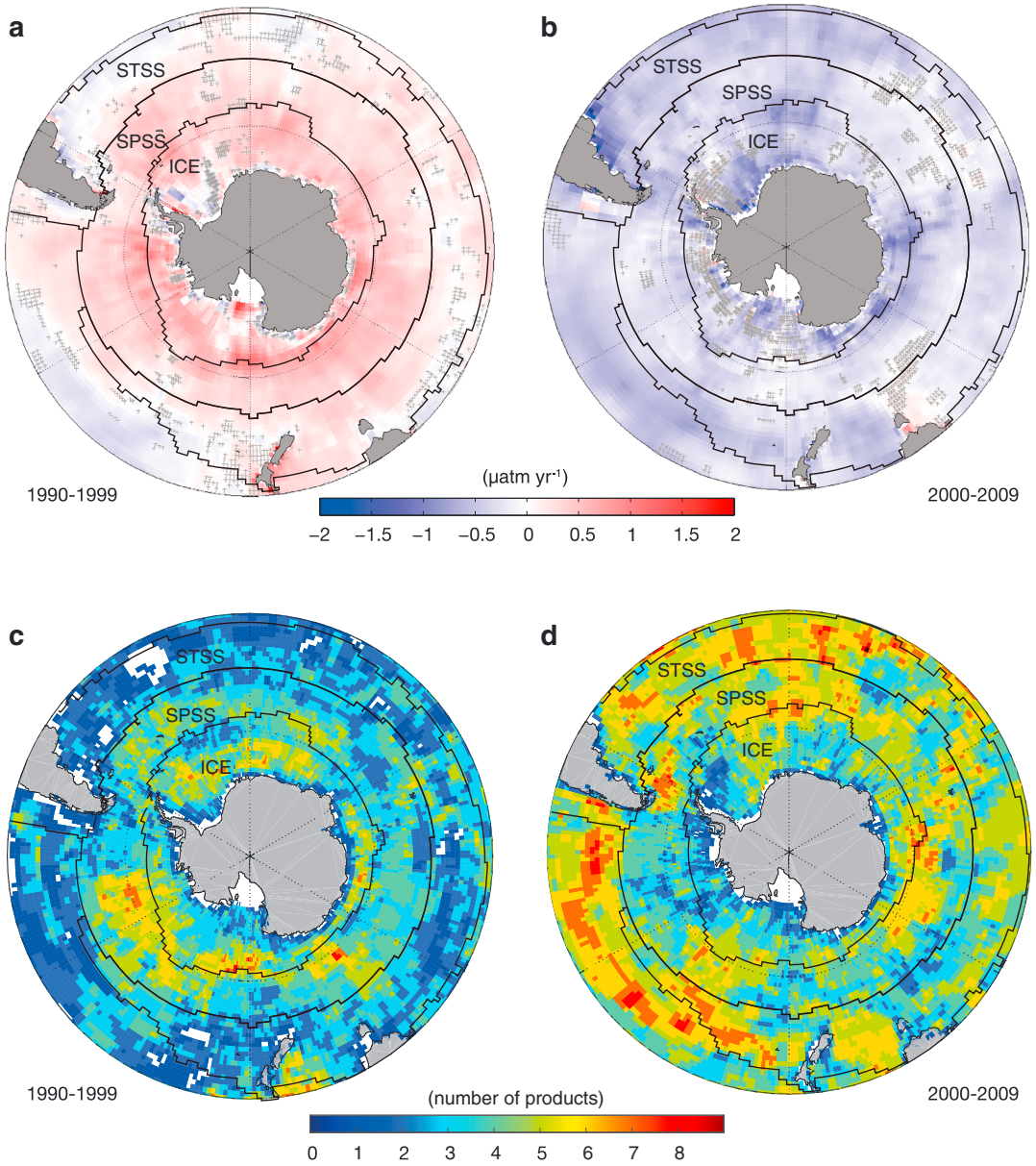

Figure 3. Spatial analysis of the trends for the 1990–1999 and 2000–2009 periods. (a) Map depicting the trends of the multiproduct mean ΔpCO_2 in μatm yr^{-1} for the period 1990 through 1999. (b) Same as Figure 3a but for the period 2000 through 2009. (c) Number of models with significant trends for the period 1990 through 1999. (d) Same as Figure 3c but for the period 2000 through 2009. Hatched areas in Figures 3a and 3b describe regions where less than 66% of the methods agree on the ΔpCO_2 trend sign. The black lines depict the borders of the SO biomes from Fay and McKinley (2014) (ICE, SPSS, and STSS biomes).

of clear relationships between the trend magnitudes and different metrics of quality as well as the lack of clear differences in methodologies implies that other sources must drive the spread. Truly independent data, that is, data that have not been used by any of the methods, would be of great value for developing better metrics of quality.

3.3. Spatial Structure of Trends

The multiproduct mean trends reveal for both periods spatially rather uniform ΔpCO_2 trends (Figures 3a and 3b). During the first decade (1990–1999), the ΔpCO_2 increased throughout the SO, with perhaps a tendency of larger changes in the most southern part of the domain, that is, in the ICE biome (Figure 3a). The situation is nearly reversed during the second decade (2000–2009), with nearly the entire SO having negative

ΔpCO_2 trends, with a tendency of a somewhat stronger trend in the regions north of the Antarctic Polar Front (STSS biome) (Figure 3). We have some confidence in this multiproduct mean trend map as it emerges from spatially rather congruent trends across the different products. For nearly all of the SO, at least six out of the nine methods (66%) agree on the sign of the trend in both decades, that is, are not hatched in Figures 3a and 3b. But most of the trends of the individual methods are not statistically significant (Figures 3c and 3d). Across the entire SO, we find that there exist only a limited number of regions for which the majority of the mapping methods have identified temporal trends that are significantly different from 0 (Figures 3c and 3d). This is particularly the case for the first decade (1990–1999), where, on average, perhaps only two out of nine methods have statistically significant positive trends. Only in the Pacific sector of the SPSS and some part of the ICE biomes we do find regions with more than five methods having statistically significant positive trends. The situation drastically improves in the second decade, that is, 2000–2009, when the majority of methods identify SO trends that are statistically different from 0 in agreement with the multimodel mean.

A striking observation regarding the spatial congruence of the trends is that there is no clear link to the density of observations (cf. Figure S1). For example, only a small number of statistically significant trends are found in the relatively data-rich area south of Tasmania in both periods, where Xue et al. (2015) reported, in fact, clear trends based on an analysis of the data without any interpolation. At the same time, there is a relatively good congruence among the methods in the Pacific sector for the period 1990–1999, even though this is a relatively data-sparse region. An exception is the Drake Passage, where the density of observations is high thanks to a dedicated observing network (Munro et al., 2015), and the methods agree very well on the sign of trend. This striking observation extends also to the temporal domain, although less pronounced. While the lower level of congruence among methods for the first decade is quite likely a consequence of the relative data paucity during this decade, the mean trend still emerges quite clearly, and consistently among all those methods that detect a trend. This lack of clear connection to data density suggests that the many other differences between the methods, such as those stemming from the input data and how the interpolation method deals with the large gaps in the data, matter quite substantially as well. Assessing all causes of disagreement goes well beyond the scope of this study. Nevertheless, our analysis provides a first-order indication that new and independent data will play an important role in improving the situation.

In our analysis we treated every method equally. However, there are distinct differences between the various estimates. While SOCATv2 contains more than 10 million observations, the LDEO database, which was used to create the CARBONES estimate only contains roughly 1 million data points, that is, an order of magnitude less. This, in addition to the methodological difference, partly explains the lack of trend significance observed in Figure 3. Considering this, it is even more astonishing that the decadal trends and their signs agree so well within the SOCOM ensemble. But there are additional sources of uncertainty that potentially influence our trend estimates. Sources of uncertainty that we do not account for are the temperature skin effect (Woolf et al., 2016), and our incomplete knowledge of the atmospheric CO_2 mixing ratio and the water vapor pressure. We do not expect them to bias the trend estimates much, however, they add to the uncertainty in the absolute value. Perhaps more relevant, at least for the estimation of the air-sea flux trends, are the uncertainties associated with the choices of the wind product and the gas transfer velocity (Wanninkhof & Triñanes, 2017), as these could induce biases in the trends. However, Landschützer et al. (2015) tested for this contribution with their product and found actually a rather small sensitivity of the trends to these choices.

Despite the substantial uncertainties associated with the spatial pattern of the trends, the relative uniformity of the ΔpCO_2 trends is remarkable. They also support the analyses of Landschützer et al. (2015) who demonstrated that the relative uniform trend pattern during the first decade is largely driven by a relatively spatially uniform trend in the nonthermal component of pCO_2, likely reflecting the enhanced upwelling of inorganic carbon-rich waters from below. This pattern thus provides further evidence that the southward shift of the westerly winds associated with the trend toward a positive phase of the Southern Annular Mode was the dominant driver for the weakening SO carbon sink, that is, supporting the classical interpretation (Lenton et al., 2013; Le Quéré et al., 2007; Lovenduski et al., 2007, 2008). Our results also speak against a very strong eddy compensation mechanism that has been proposed to substantially reduce the effect of the wind changes on the upwelling (Dufour et al., 2013). The relatively uniform pattern of the first decade of the 2000s is even more puzzling, and also consistent with the interpretations of Landschützer et al. (2015), who suggested that this trend emerges from a complex interaction of a cooling trend in the Pacific sector, and a trend toward lower dissolved inorganic carbon concentration in the Atlantic sector, both possibly driven by the development of a more meandering westerly wind belt (cf. Figure 3 in Landschützer et al., 2015). Another plausible cause for

the observed variability is the remote impact from the Equatorial Pacific through atmospheric teleconnection patterns (Verdy et al., 2007), although the decadal signature of this driver has not yet been fully explored.

4. Conclusions

Here we analyze decadal trends in the SO based on observations of the sea surface partial pressure of CO_2 from a set of complementary pCO_2 mapping methods collated by the SOCOM project. While we find good agreement between methods regarding the sign of the basin average decadal trend and its variability, there is much less agreement regarding its strength. However, our analysis shows that the difference in strength is not due to an incorrect representation of the available observations but more likely a result of how the spatial and seasonal heterogeneity and the sparseness of the data is dealt with. A further reason are the different databases used for the buildup and testing of the methods. Spatially, six out of nine products tend to agree on the sign of the trend in both periods for the vast majority of the SO, but few of these trends are statistically significant at the level of the $1° \times 1°$ grid chosen for our investigation.

Overall, our analyses suggests that current observation-based interpolation methods are able to represent low-frequency variability of the SO carbon sink well. However, our results clearly highlight the need to further observe the SO to better understand the mechanisms driving the temporal mean and year-to-year variations in the air-sea CO_2 flux. This is a fundamental prerequisite to improve our ability to project the future fate of the SO sink, which is crucial given its significance for the global climate.

Acknowledgments
The main body of this work was supported by ETH Zurich. The contribution of K.B.R. was supported by NASA award NNX14AL85G. We thank all colleagues, staff, and funding bodies involved in the data collection and quality control of pCO_2, which form the basis of the mapping products within the SOCOM project. The Surface Ocean CO_2 Atlas (SOCAT) is an international effort, endorsed 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. We thank all colleagues of the Lamont-Doherty Earth Observatory (LDEO) for their efforts. SOCAT and LDEO data are publicly available via the NOAA National Centers for Environmental Information (NCEI) and can be publicly accessed via www.socat.info and https://www.nodc.noaa.gov/ocads/oceans/LDEO_Underway_Database/, respectively. For material, methods, a list of all participating methods and their documentation please visit the SOCOM project via http://www.bgc-jena.mpg.de/SOCOM/.

References

Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., … Xu, S. (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

Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., … Watson, A. J. (2014). An update to the Surface Ocean CO_2 Atlas (SOCAT version 2). *Earth System Science Data*, *6*(1), 69–90. https://doi.org/10.5194/essd-6-69-2014

Dee, D. P., Uppala, S., Simmons, A., Berrisford, P., Poli, P., Kobayashi, S., … Vitart, F. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597. https://doi.org/10.1002/qj.828

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

Dufour, C. O., Sommer, J. L., Gehlen, M., Orr, J. C., Molines, J.-M., Simeon, J., & Barnier, B. (2013). Eddy compensation and controls of the enhanced sea-to-air CO_2 flux during positive phases of the Southern Annular Mode. *Global Biogeochemical Cycles*, *27*, 950–961. https://doi.org/10.1002/gbc.20090

Fay, A. R., & McKinley, G. A. (2013). Global trends in surface ocean pCO_2 from in situ data. *Global Biogeochemical Cycles*, *27*, 1–17. https://doi.org/10.1002/gbc.20051

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

Fay, A. R., McKinley, G. A., & Lovenduski, N. S. (2014). Southern Ocean carbon trends: Sensitivity to methods. *Geophysical Research Letters*, *41*, 6833–6840. https://doi.org/10.1002/2014GL061324

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*(2), 862–886. https://doi.org/10.1175/JCLI-D-14-00117.1

GLOBALVIEW-CO2 (2013). Cooperative Atmospheric Data Integration Project-Carbon Dioxide, CD-ROM, NOAA ESRL, Boulder, Colorado. Retrieved from ftp.cmdl.noaa.gov, Path: ccg/co2/GLOBALVIEW

Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., … Takahashi, T. (2009). Oceanic sources, sinks, and transport of atmospheric CO_2. *Global Biogeochemical Cycles*, *23*, GB1005. https://doi.org/10.1029/2008GB003349

Iida, Y., Kojima, A., Takatani, Y., Nakano, T., Sugimoto, H., Midorikawa, T., & Ishii, M. (2015). Trends in pCO_2 and sea-air CO_2 flux over the global open oceans for the last two decades. *Jornal of Oceanography*, *71*, 637–661. https://doi.org/10.1007/s10872-015-0306-4

Jones, S. D., Le Quéré, C., Rödenbeck, C., Manning, A. C., & Olsen, A. (2015). A statistical gap-filling method to interpolate global monthly surface ocean carbon dioxide data. *Journal of Advances in Modeling Earth Systems*, *07*, 1942–2466. https://doi.org/10.1002/2014MS000416

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., … Joseph, D. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American meteorological Society*, *77*(3), 437–471.

Khatiwala, S., Tanhua, T., Mikaloff Fletcher, S., Gerber, M., Doney, S. C., Graven, H. D., … Sabine, C. L. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, *10*, 2169–2191. https://doi.org/10.5194/bg-10-2169-2013

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

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

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

Le Quéré, C., Andrew, R. M., Canadell, J. G., Sitch, S., Korsbakken, J. I., Peters, G. P., & Zaehle, S. (2016). Global Carbon Budget 2016. *Earth System Science Data*, *8*(2), 605–649. https://doi.org/10.5194/essd-8-605-2016

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

Lenton, A., Tilbrook, B., Law, R., Bakker, D., Doney, S., Gruber, N., … Takahashi, T. (2013). Sea-air CO_2 fluxes in the Southern Ocean for the period 1990–2009. *Biogeosciences*, *10*, 4037–4054. https://doi.org/10.5194/BGD-10-285-2013

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

Lovenduski, N. S., 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*, GB2026. https://doi.org/10.1029/2006GB002900

Majkut, J. D., Sarmiento, J. L., & Rodgers, K. B. (2014). A growing oceanic carbon uptake: Results from an inversion study of surface pCO_2 data. *Global Biochemical Cycles*, *28*, 335–351. https://doi.org/10.1002/2013GB004585

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

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., … Sarmiento, J. L. (2006). Inverse estimates of anthropogenic CO_2 uptake, transport, and storage by the ocean. *Global Biogeochemical Cycles*, *20*, GB2002. https://doi.org/10.1029/2005GB002530

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

Park, G.-H., Wanninkhof, R., Doney, S. C., Takahashi, T., Lee, K., Feely, R. A., … Lima, I. D. (2010). Variability of global net sea–air CO_2 fluxes over the last three decades using empirical relationships. *Tellus B*, *62*(5), 352–368. https://doi.org/10.1111/j.1600-0889.2010.00498.x

Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., … Yoshikawa-Inoue, H. (2013). A uniform, quality controlled Surface Ocean CO_2 Atlas (SOCAT). *Earth System Science Data*, *5*(1), 125–143. https://doi.org/10.5194/essd-5-125-2013

Rayner, N., Parker, D. E., Horton, E., Folland, C., Alexander, L., Rowell, D., … 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

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

Rödenbeck, C., Keeling, R., Bakker, D., 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*, 93–216. https://doi.org/10.5194/os-9-193-2013

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., … Rios, A. F. (2004). The oceanic sink for anthropogenic CO_2. *Science (New York, N.Y.)*, *305*(5682), 367–371. https://doi.org/10.1126/science.1097403

Sabine, C. L., Hankin, S., Koyuk, H., Bakker, D. C. E., Pfeil, B., Olsen, A., … Yoshikawa-Inoue, H. (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

Sarmiento, J., & Gruber, N. (2006). *Ocean biogeochemical dynamics* (503 pp.). Princeton, Woodstock: Princeton University Press.

Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., & Midgley, P. E. (2010). Meeting report of the intergovernmental panel on climate change expert meeting on assessing and combining multi model climate projections (Tech. Rep.) Bern, Switzerland: IPCC Working Group I Technical Support Unit, University of Bern.

Takahashi, T., Sutherland, S. C., & Kozyr, A. (2014). Global ocean surface water partial pressure of CO_2 database: Measurements performed during 1957–2013 (version 2013) (Tech. Rep.) Oak Ridge, TN. https://doi.org/10.3334/CDIAC/OTG.NDP088(V2013)

Verdy, A., Dutkiewicz, S., Follows, M. J., Marshall, J., & Czaja, A. (2007). Carbon dioxide and oxygen fluxes in the Southern Ocean: Mechanisms of interannual variability. *Global Biogeochemical Cycles*, *21*, GB2020. https://doi.org/10.1029/2006GB002916

Wanninkhof, R., Park, G.-H., Takahashi, T., Sweeney, C., Feely, R., Nojiri, Y., … Khatiwala, S. (2013). Global ocean carbon uptake: Magnitude, variability and trends. *Biogeosciences*, *10*, 1983–2000. https://doi.org/10.5194/bg-10-1983-2013

Wanninkhof, R., & Triñanes (2017). The impact of changing wind speeds on gas transfer and its effect on global air-sea CO_2 fluxes. *Global Biogeochemical Cycles*, *31*, 961–974. https://doi.org/10.1002/2016GB005592

Woolf, D., Land, P. E., Shutler, J. D., Goddijn-Murphy, L., & Donlon, C. J. (2016). On the calculation of air-sea fluxes of CO_2 in the presence of temperature and salinity gradients. *Journal of Geophysical Research: Oceans*, *121*, 1229–1248. https://doi.org/10.1002/2015JC011427

Xue, L., Gao, L., Cai, W.-J., Yu, W., & Wei, M. (2015). Response of sea surface fugacity of CO_2 to the SAM shift south of Tasmania: Regional differences. *Geophysical Research Letters*, *42*, 3973–3979. https://doi.org/10.1002/2015GL063926

York, D., Evensen, N. M., Martınez, M. L., & Delgado, J. D. B. (2004). Unified equations for the slope, intercept, and standard errors of the best straight line. *American Journal of Physics*, *72*(3), 367–375. https://doi.org/10.1119/1.1632486

Zeng, J., Nojiri, Y., Landschützer, P., Telszewski, M., & Nakaoka, S. (2014). A global surface ocean fCO_2 climatology based on a feed-forward neural network. *Journal of Atmospheric and Oceanic Technology*, *31*(8), 1838–1849. https://doi.org/10.1175/JTECH-D-13-00137.1