Integrated Analysis of Carbon Dioxide and Oxygen Concentrations as a Quality Control of Ocean Float Data

Yingxu Wu[1,2], Dorothee C.E. Bakker[3], Eric P. Achterberg[4], Amavi N. Silva[1], Daisy D. Pickup[1,7], Xiang Li[5], Sue Hartman[6], David Stappard[1], Di Qi[2]*, Toby Tyrrell[1]*

[1]Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, UK

[2]Polar and Marine Research Institute, Jimei University, Xiamen, China

[3]Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

[4]GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

[5]Department of Physics, George Washington University, Washington, DC, USA

[6]National Oceanography Centre, Southampton, UK

[7]now at: Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

First corresponding author: Toby Tyrrell (Toby.Tyrrell@soton.ac.uk)

Second corresponding author: Di Qi (qidi60@qq.com)

Abstract

The distributions of dissolved O_2 and CO_2 have not previously been systematically compared across the global surface ocean, despite their significance for life and climate. Here we analyze carbon dioxide and oxygen concentrations relative to saturation (equilibrium with the atmosphere) in surface waters, using two large datasets (ship-collected and float-collected data). When applied to a high-quality global ship-collected dataset, CO_2 and O_2 concentrations relative to saturation exhibit large seasonal and geographic variations. However, linear fits of CO_2 and O_2 deviations from saturation (ΔCO_2 against ΔO_2) yield y-intercepts close to zero, which suggests a requirement for data validity. We utilize this finding to investigate the accuracy of carbonate system data from biogeochemical-Argo floats. We find significant discrepancies in ΔCO_2-ΔO_2 y-intercepts compared to the global reference, implying overestimations of float-based CO_2 release in the Southern Ocean. We conclude that this technique can be applied to data from autonomous platforms for quality assessment.

Introduction

The dissolved gases carbon dioxide (CO_2) and oxygen (O_2) in seawater are of much biogeochemical interest[1]. Carbon dioxide is important because of its role as a greenhouse gas, with about one quarter of the anthropogenic CO_2 produced by fossil fuel combustion and land use changes being absorbed by the ocean[2]. The coupling of atmospheric CO_2 and O_2 is used to derive land/ocean carbon sink partitioning[3] and serve as a reference to verify ocean model results used in the global carbon budget[2]. Time-series observations at some specific locations (mostly in the northern hemisphere[4]), as well as distributed measurements of the partial pressure of CO_2 (pCO_2) at the global scale[5,6] show that surface seawater pCO_2 is rising at a similar rate to the mole fraction of CO_2 in the atmosphere, which has increased by more than 40% since pre-industrial times (from 280 to over 400 ppm or μmol mol^{-1}). Global change is also affecting oceanic O_2 concentrations; warming decreases oxygen solubility and enhances water column stratification, thereby reducing ventilation of subsurface waters with atmospheric oxygen and leading to a decline in oxygen in the global ocean[7-9]. Oxygen is biologically linked to CO_2, for instance during photosynthesis which simultaneously uses CO_2 and generates O_2.

There have been some attempts to jointly investigate dissolved O_2 and CO_2 in different ocean basins (e.g., ref. [1,10-15]). However, in these studies the two gases were usually treated differently, for instance O_2 as a concentration ([O_2]) or percent saturation and CO_2 as a partial pressure (pCO_2) (ref. [10,13,16,17]); in addition, oxygen values are sometimes reported relative to argon (Ar) (e.g., ref. [18-20]). Analyses in which O_2 and CO_2 are calculated in different units, or as percentages, cannot take straightforward advantage of the stoichiometric relationships (i.e., Redfield ratios) between carbon, oxygen, and nutrients[21,22]. An improved O_2-CO_2 analysis method was proposed by Torgersen and Branco[23] and Vachon et al.[24] to compare deviations of O_2 and CO_2 concentrations away from saturation, or in other words disequilibria (discrepancies compared to equilibrium with atmospheric values). This approach was shown to provide insights into river and lake ecosystems, and has the potential to be applied more broadly to marine systems. Investigating co-variations of O_2 and CO_2 concentrations can help improve understanding of the drivers of surface ocean carbon dynamics.

In this study we extend the application of the O_2-CO_2 approach to the global surface ocean and name the

approach CORS (Carbon and Oxygen Relative to Saturation). We treat O_2 and CO_2 identically and compare dissolved concentrations of O_2 and CO_2 ([CO_2] and [O_2]) in surface seawater to saturation values (values at which the net air-sea gas exchange rate is zero). The saturation values for O_2 and CO_2 are strongly temperature-dependent, as was already shown by the first plot of [O_2] against temperature over much of the global surface ocean in the early 1980s using GEOSECS data[25] (Supplementary Figure 1). The global database has been greatly expanded in recent decades, providing wider spatial and temporal coverage, culminating in the Global Ocean Data Analysis Project (GLODAPv2.2020[26-28]; used throughout this study, for simplicity, it is referred to as GLODAPv2 hereafter), which is by far the largest high-quality observational dataset of both carbon and oxygen. This expanded dataset has not previously been used to compare [O_2] and [CO_2] to each other and to saturating values.

Furthermore, a potential application is to compare patterns in CORS plots from GLODAPv2 with those from float data from the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM). Equipped with biogeochemical sensors (e.g., oxygen, nitrate, pH, and bio-optical sensors), ~200 autonomous biogeochemical Argo floats were deployed by the SOCCOM project[29] and have enabled a better understanding of carbon and oxygen cycles in the Southern Ocean[30,31]. They strikingly found that the high-latitude Southern Ocean (i.e., Antarctic-Southern Zone, ASZ) has released much more CO_2 to the atmosphere than previously estimated[31-33], which attracted community concerns on the sensor bias and data quality control (QC) of pH and associated carbonate parameters[34]. Funding has recently been announced (Global Ocean Biogeochemistry GO-BGC Array project) for the construction and deployment of 500 floats (as a contribution towards an anticipated eventual fleet of 1000 floats) to provide float coverage similar to that provided by the SOCCOM project but across the global ocean. A consensus is urgently required on how to calibrate and validate float carbonate data to ensure the highest accuracy and comparability among different studies and datasets[34]. Given that the float O_2 data is likely to be more accurate than pH and the calculated carbonate data (ref. [29,31,35]; see also descriptions in Methods), O_2 in the context of CORS plots could provide a strong constraint for detecting questionable float CO_2 data, if compared to GLODAPv2.

To advance knowledge of oceanic oxygen and carbon cycling and address the above concerns, our study had two main objectives: (1) to construct CORS plots from the GLODAPv2 database, to be used later as a reference to compare against. These plots show the CORS analysis to be capable of identifying regions and

periods where processes have driven both O_2 and CO_2 away from their equilibrium with the atmosphere. (2) to apply this method to the SOCCOM dataset and compare the resulting plots to the GLODAP reference, in order to investigate the potential of CORS as a tool for interpreting and validating data collected by autonomous platforms.

Results

Overall patterns in CORS plots from GLODAPv2 data.

To first order, both [CO_2] and [O_2] from GLODAPv2 follow the solubility relationship with temperature (decreasing values with increasing temperature) (Fig. 1), as found previously for [O_2] in GEOSECS data (Supplementary Figure 1). However, deviations occur in certain regions and seasons (Supplementary Figures 2-6). Deviations of O_2 from its equilibrium with the atmosphere are usually of the opposite sign to the corresponding CO_2 deviations (Supplementary Figure 2).

In both hemispheres, the distributions of [CO_2] and [O_2] show strong seasonal variations: [CO_2] and [O_2] deviate furthest from their temperature-dependent saturation values in spring and summer while staying close to saturation in autumn and winter (Fig. 1). We do not discuss further the Indian Ocean because deviations of [CO_2] and [O_2] from saturation are less pronounced there than in other ocean basins. Individual CORS plots for each basin are presented as supporting information (Supplementary Figures 3-6).

In spring, supersaturation of O_2 usually accompanies CO_2 undersaturation. Strong supersaturation of CO_2 (together with undersaturation of O_2) is observed in parts of the northeast and eastern equatorial Pacific for water temperatures close to 10 and 18°C (Supplementary Figure 2a). However, other Pacific data exhibit CO_2 undersaturation and accompanying O_2 supersaturation (Fig. 2a). In the Atlantic Ocean, 73% of all spring data are undersaturated in CO_2 while supersaturated in O_2 (Fig. 2a).

In summer, the undersaturation of CO_2 is less pronounced in the Atlantic Ocean (Figs. 1c and 2b), whereas in the Pacific Ocean it is more or less similar to that in spring. In summer, some simultaneous CO_2 and O_2 undersaturations are observed in the Southern Ocean at latitudes polewards of 60°S where ice melt occurs at coastal regions (Figs. 1c, d): 20% of the summer Southern Ocean data show undersaturation in both CO_2 and O_2.

In autumn and winter, there is less data across the global oceans but it appears that both gases stay closer to saturation as biological activity weakens and air-sea gas exchange strengthens. There are striking opposite changes to CO_2 and O_2 in the Southern Ocean (Figs. 1 and 2), where [CO_2] is elevated (on occasion to as high as 30 μmol kg^{-1}) and [O_2] is depleted (sometimes to as low as 260 μmol kg^{-1}) (Fig. 1). Overall, both gases deviate more strongly from saturation in winter than in autumn in the Southern Ocean.

Processes causing deviations in CORS plots from GLODAPv2 data.

Figures 1 and 2 show, for the global surface ocean, the coupling of CO_2 deviations and O_2 deviations from saturation across geographic and seasonal scales, with four specific features (F1 to F4 in Fig. 2) warranting further investigation: (F1) CO_2 undersaturation in conjunction with O_2 supersaturation in the high-latitude Atlantic and Pacific Oceans in spring; (F2) CO_2 supersaturation paired with O_2 undersaturation in the eastern equatorial Pacific and California coast in spring and summer; and (F3-F4) supersaturation of CO_2 together with undersaturation of O_2 in the Southern Ocean in winter and to a lesser extent in spring and autumn. Processes known to simultaneously affect ΔCO_2 and ΔO_2 include warming/cooling, ice melting, respiration and photosynthesis, and upwelling. The impacts of these processes on ΔCO_2 and ΔO_2 are shown in the inset to Fig. 2d (see Methods - Predicted effects of different processes – for explanation of the inset figure).

With additional plots we explore the possible causes of these features. Figure 3a shows a CORS plot of Atlantic and Pacific spring data, colored by in-situ nitrate concentration. The data falling in the fourth quadrant (negative ΔCO_2 and positive ΔO_2) are associated with depleted nitrate concentrations and are located primarily in the Irminger Basin in the North Atlantic and the Oyashio region in the western subarctic Pacific Ocean (Supplementary Figure 2a, b), regions where intense spring blooms are observed[36-39]. The data patterns are generally consistent with phytoplankton blooms (photosynthesis) as the driver of the F1 deviations (although the lack of correlation (Supplementary Table 1) between CORS and NO_3 is surprising). Data in quadrant two of Fig. 3a (F2) are mainly from off the northern California coast, a region where seasonal coastal upwelling is known to bring subsurface waters (depleted in O_2 and enriched in CO_2 and nutrients from decomposition of organic matter) to the surface ocean[40]. CORS and NO_3 are strongly correlated in these data (Supplementary Table 1).

Figure 3b shows the relationship between ΔCO_2, ΔO_2, and the NO_3 anomaly in the Southern Ocean in winter, where the NO_3 anomaly is the difference of surface in-situ nitrate concentration from its annual

mean value in the surface Southern Ocean based on GLODAPv2. The nitrate anomaly is strongly correlated with ΔCO_2 and ΔO_2 (Supplementary Table 1). The winter data in the Southern Ocean (Fig. 2d) imply that respiration or, more likely, upwelling of 'old' water into which organic matter has been respired, is responsible for F3 and F4. The calculation of ΔCO_2 (Equation 2) is made relative to atmospheric CO_2 at the time of measurement, even for water that has recently upwelled and never previously had contact with anthropogenic ΔCO_2. For this reason, data from recently upwelled water will tend to plot lower on CORS plots than it would if its $[CO_2]$ value was compared to the $[CO_2]$ value in equilibrium with pre-industrial atmospheric CO_2. In addition, decreases in $[CO_2]$ because of $CaCO_3$ dissolution[41,42] (which does not affect $[O_2]$) are likely to contribute to the lower-than-expected slope. Figure 3c (and Supplementary Figure 7) shows that Southern Ocean surface waters with the largest deviations (those furthest from the origin of the CORS plots) in winter are those which have recently upwelled (neutral density > 27.8 kg m^{-3}, ref. [43-46]). We therefore conclude that the large excursions in the Southern Ocean in autumn and winter are driven by upwelling of deep waters that have previously been altered by decomposition of sinking organic matter.

Near-zero y-intercepts in CORS plots from GLODAPv2 data.

We investigated CORS plots to look for common features in them, when generated from the high-quality data in GLODAPv2. For quiescent regions not experiencing intense biogeochemical activity, gas exchange is the dominant control for both dissolved gases, keeping them close to equilibrium with the atmosphere. In such regions, we would expect the centroid of the data in the CORS plots to then be close to the origin (i.e. $\overline{\Delta O_2}$ and $\overline{\Delta CO_2} \approx 0$ μmol kg^{-1}), and this is what is seen when CORS plots are constructed from the HOT and BATS time-series data (Supplementary Figure 8; $\overline{\Delta O_2}$= 1.9 μmol kg^{-1} and $\overline{\Delta CO_2}$= -0.2 μmol kg^{-1} at HOT; $\overline{\Delta O_2}$= 3.2 μmol kg^{-1} and $\overline{\Delta CO_2}$= -0.3 μmol kg^{-1} at BATS) from sub-tropical gyres. However, in less quiescent regions, the centroid can be shifted away from the origin (as seen for instance in the Southern Ocean in winter – Figure 3b – where upwelling leads to a displaced centroid: $\overline{\Delta O_2} = -26.9$ μmol kg^{-1} and $\overline{\Delta CO_2} = 1.4$ μmol kg^{-1}). Distance of the centroid from the origin cannot, therefore, be considered a reliable indicator of data quality.

Several other statistical properties can be calculated (see for instance Vachon et al.[24]), of which we found the y-intercept value (i.e., value of ΔCO_2 when ΔO_2 is zero) to be the most useful. The application of CORS to the GLODAPv2 dataset shows that, when strongly influenced by a dominant biogeochemical process, the departures of O_2 and CO_2 from atmospheric equilibrium are coupled and the best-fit lines of ΔO_2 and ΔCO_2

still tend to intersect close to the coordinate origin (Fig. 4). Across the global oceans, we found y-intercepts close to zero (range -1.10 to -0.16 μmol kg^{-1} in different basins (Fig. 4), with value for the global dataset of -0.18 μmol kg^{-1}). Due to the effect of ice melt in the summer Southern Ocean, the Southern Ocean y-intercept is significantly lowered compared to other ocean basins (Fig. 4e). The relative uniformity of y-intercept values suggests their usefulness as indicators of data quality.

CORS plots from all float data, regardless of QC flag.

Below we show that CORS plots are capable of distinguishing 'questionable' or 'bad' float data from QCed 'good' data. As an illustration, we examined data from floats F9096 and F9099 deployed in the high-latitude Southern Ocean, using which Williams et al.[33] found significantly higher sea surface pCO_2 and air-sea CO_2 efflux in wintertime. Unlike Williams et al.[33] in Figure 5 we have plotted both data flagged as 'questionable' or 'bad' and data flagged as 'good'. The figure shows some abnormally high (up to 20 μmol kg^{-1}) and low (down to -20 μmol kg^{-1}) ΔCO_2 values, neither of which are coupled with equivalent ΔO_2 values. The resulting CORS plots are significantly different from the general CORS pattern across the global ocean (Fig. 4). All of these abnormal data have been flagged 'questionable' or 'bad' by SOCCOM's QC procedure but CORS plots also reveal the data to be problematic.

CORS plots using only float data flagged as 'good'.

Below we show that CORS plots are also useful for analyzing and examining 'good' float data. In order to be comparable with the GLODAPv2 database, we first selected 12 biogeochemical Argo floats (Supplementary Figure 9) around the Drake Passage or south of Tasmania to get as many overlaps as possible with the shipboard dataset[28]. When CORS plots were produced for the float data, we found (Fig. 6) a pattern that is overall rather similar to that from GLODAP data in the Southern Ocean (Figs. 3b and 4f), consistent with upwelling of CO_2-rich deep waters. However, a discrepancy was found with regards to the y-axis intercepts of the best-fit regressions of ΔCO_2 and ΔO_2: in contrast to the global GLODAP y-intercepts, which are all similar to each other (range of less than 1.0 between minimum and maximum), the float-derived y-intercepts are highly variable (range of -3.81 to +0.99 μmol kg^{-1}; Figs. 4 and 6, Table 1). We adopted a subsampling strategy (see Methods) to treat the GLODAP and float datasets identically, with the result showing that half of the selected floats have y-intercepts greater than the GLODAP-derived value of -1.10 μmol kg^{-1} (Fig. 6, Table 1). Among these floats, F9096 and F12545 deviate more strongly from the GLODAP pattern along the positive y-axis direction, whereas some other floats, e.g., F9275 and F9646 deviate along the opposite direction (Fig. 6). The average difference in y-intercepts (calculated as

GLODAPv2 minus float values) is -0.12 µmol kg^{-1}, implying that, overall, float y-intercepts are more positive (or less negative) than GLODAP ones (Table 1).

Applying CORS to the whole SOCCOM dataset south of 55°S (48 floats with QCed O_2 and CO_2 data, Supplementary Table 2), we also found that y-intercepts were on average greater than the GLODAP-derived ones, with an average difference (GLODAPv2 minus float) of -0.36 µmol kg^{-1} (Supplementary Table 2). Given that oxygen sensors are accepted as more established, reliable, and accurate than pH sensors from which the CO_2 values were calculated[29], the anomalous float y-intercepts suggest offsets most likely due to pH-related biases, such as the uniform crossover correction assuming a fixed offset of pH from 1500 m depth to surface[29,33,34]. While unusual y-intercept values are suggestive of data quality issues, they are not necessarily definitive proof. It is also possible, we believe, that, in some locations, local processes produce real CORS patterns that differ from those normally seen. For instance, surface water near to where rivers enter the sea, or in regions of strong mixing with deeper waters, could potentially exhibit persistent unusual CORS patterns. Data displaying unusual behaviour on CORS plots should therefore not be immediately discounted but should instead be flagged as requiring further investigation before it can be accepted as valid. Detailed investigation of this issue is beyond the scope of this manuscript, but we note that a correction of -0.36 µmol kg^{-1} to [CO_2] corresponds to a correction to pCO_2 of -5.8 µatm under typical Southern Ocean conditions. Our finding is in line with some recent studies[47-49] based on different approaches (airborne observations of atmospheric CO_2 gradients, uncrewed surface vehicle observations of circumnavigation of Antarctica, and reconstructed estimates of winter observations and CO_2 fluxes) that indicate the possible overestimation of CO_2 outgassing from SOCCOM float data.

Discussion and Conclusions

An improved analysis technique (CORS) for paired O_2-CO_2 data has been presented here, treating both gases identically and accounting for the atmospheric pressure effect on both gas saturation concentrations. The CORS technique was applied to the large, high-quality, global dataset GLODAPv2, as well as to the SOCCOM float dataset. CORS plots provide a detailed insight into the identity and intensity of processes impacting CO_2 and O_2. Although both gas concentrations were seen to be often close to the temperature-determined equilibrium value in GLODAP2 data, several noteworthy deviations from

equilibrium with the atmosphere and the possible processes driving them were discussed: in spring, phytoplankton blooms (most notably in the Irminger Basin of the North Atlantic and in the Oyashio region of the western subarctic Pacific Ocean) drive undersaturation of CO_2 and coincident supersaturation of O_2, whereas upwelling in the California coast and equatorial Pacific drives the opposite; in autumn and winter, upwelling in the Southern Ocean produces supersaturation of CO_2 and undersaturation of O_2. CORS plots can serve as a useful tool for detection of processes (e.g., blooms and upwelling) that simultaneously affect both oxygen and carbon.

Recent developments in sensors mounted on gliders, floats, and moorings have become increasingly important because they provide opportunities for sampling in remote regions and in inclement weather where traditional shipboard measurements are difficult and expensive to obtain[31,32,50-54]. Along with the basic hydrological parameters (i.e., temperature and salinity), biogeochemical variables such as nutrients, oxygen, and CO_2 (or pH, from which CO_2 can be estimated) are now being measured[30-33,55]. ΔCO_2 and ΔO_2 can therefore be obtained from measurements made autonomously, and CORS plots generated from them. Even though oxygen returns to gas exchange equilibrium more rapidly than does carbon dioxide[56,57], ΔCO_2 and ΔO_2 have been shown here to exhibit coupled (simultaneous and proportional) changes in CORS plots from GLODAPv2 data. Moreover, oxygen data from sensors is generally more accurate and reliable than CO_2 derived pH measured on Argo floats[29,35]. Because of this it makes sense to exploit the O_2 data to improve the CO_2 data. Here we have shown that CORS plots can often identify questionable data (data shown to be questionable by other QC methods) immediately. In addition, our results suggest that CORS plots can also reveal issues with supposed ‘good’ data (i.e., quality issues not picked up by other QC methods). This is because systematic errors in either [CO_2] or [O_2] tend to stand out in CORS plots, and to lead to anomalous y-intercept values relative to the GLODAPv2 reference (Table 1). Our approach provides a more straightforward way to assess and potentially improve CO_2 data quality by comparison to other float measurements. CORS plots allow a check on sensor performance, which is important for instruments on these unmanned platforms which operate without servicing or recalibration.

As the oceanographic community becomes increasingly reliant on data collected from autonomous platforms[50,58], techniques such as CORS will be beneficial for diagnosing data quality, and for immediate detection of questionable data.

Methods

The surface ocean is defined[41,59] as shallower than 30 m at latitudes greater than 30°, and shallower than 20 m at latitudes less than 30°. The Southern Ocean is defined as south of 50°S. Boreal spring is taken as from April to June, and austral spring from October to December, and so on for the other seasons (following global scale studies[60]).

Dataset descriptions.

GLODAP dataset: Data for this study were obtained from GLODAPv2.2020[26-28] (denoted 'GLODAPv2' in this manuscript), which includes data from 946 cruises conducted during the period 1972-2019. Only open ocean data (seafloor depth > 200 m) were included. We excluded data from the Arctic Ocean (>65°N) because of data scarcity and strong perturbations from river inputs[61]. The observed CO_2 concentration was calculated using the MATLAB version[62] of CO2SYS, from in-situ temperature, salinity, DIC, TA, phosphate, and silicate in the GLODAPv2 database. In this study, the dissociation constants for carbonic acid and sulfate were taken from Lueker et al.[63] and Dickson[64], respectively, and the total borate-salinity relationship from Lee et al.[65]. We used data only when both O_2 and carbonate system measurements are available and when the quality control of data is flagged as 'good'.

The accuracies of measured O_2, DIC, and TA from GLODAPv2 are stated as 1% (≈ 3 μmol kg^{-1}), 4 μmol kg^{-1}, and 4 μmol kg^{-1}, respectively[28]. The uncertainty of calculated CO_2 concentration is dominated by the uncertainties from DIC and TA[66], and is assessed using an add-on to the CO2SYS program that calculates uncertainty propagation[66]. The propagated uncertainty of CO_2, taking into account the uncertainties in the input variables as well as in the equilibrium constants, is 0.4 μmol kg^{-1}. We also evaluated the uncertainty ourselves using a Monte Carlo analysis (following Wu et al.[67]), which produced a similar uncertainty of 0.5 μmol kg^{-1}.

SOCCOM float dataset: The SOCCOM project (https://soccom.princeton.edu/) has deployed ~200 biogeochemical profiling floats in the Southern Ocean since 2014. The floats are mounted with a combination of biogeochemical sensors including ones for measuring water column pH, oxygen, and nitrate[29]. Carbonate system parameters including pCO_2 and others are first calculated from sensor-measured temperature, salinity, pH, LIAR algorithm-estimated TA, and silicate and phosphate concentrations[33,68,69]. The nutrient data are derived from a matched GLODAP database as a function of potential density (see details in Williams et al.[33]); ignoring silicate and phosphate concentrations has

anyway only a negligible effect on carbonate system calculation[33]. The CO_2 concentration [CO_2] is then calculated from pCO_2 and CO_2 solubility using Henry's Law ([CO_2] = $K_H \times pCO_2$). The pH-dependent bias correction and quality control[29,33] has been applied to the accessible data from https://soccompu.princeton.edu/www/index.html. The quality control of pH data is based on the crossover analysis for deep waters between float and shipboard (including available Southern Ocean dataset and SOCCOM deployment cruises) measurements (see also ref[33,69]). The empirical algorithm for estimating in-situ pH as a function of temperature, salinity, pressure, and O_2 is determined for shipboard bottle measurements at 1000-2000 m depth, which is then applied to float-measured temperature, salinity, pressure, and O_2. By comparing the two pH values at 1500 m depth, an offset in pH is applied to the entire float profile. The measured oxygen and pH have reported uncertainties of 1% and 0.01 respectively, and the estimated TA and pCO_2 have reported uncertainties of 5.6 µmol kg^{-1} and 2.7% respectively[33]. Only data with a quality-control flag of 'good' were used.

The propagated uncertainty of float [CO_2] calculated following Orr et al.[66] is 0.8 µmol kg^{-1}, which is twice the uncertainty of [CO_2] from GLODAPv2. The float oxygen sensors have been shown to perform robustly and with good stability (100% good data return[29]). The oxygen data can be calibrated every time a float surfaces because it can measure oxygen above the sea surface; this calibration is then used to adjust the entire profile[29]. From the crossover comparison to GLODAPv2, the float-measured oxygen data was seen to be closely correlated with GLODAPv2 data and to follow a 1:1 relationship. In contrast, the float-measured pH data exhibited large offsets from adjacent GLODAPv2 data and there was a significant departure from a 1:1 relationship (e.g., fitted line deviated from 1:1 line by 0.03 at pH = 8.05; ref. [29]).

Converting CO_2 to pCO_2 is obtained by multiplying pCO_2 by the Henry's constant for CO_2 (K_H), based on an average sea surface temperature of 1°C in the Southern Ocean (i.e., $K_H \approx 0.06$).

Calculation of saturation values for O_2 and CO_2 and their uncertainties.

$$\Delta O_2 = [O_{2,obs}] - [O_{2,sat}] \quad (1)$$

$$\Delta CO_2 = [CO_{2,obs}] - [CO_{2,sat}] \quad (2)$$

where the subscript 'obs' indicates the observed concentration, and 'sat' indicates the saturation concentration (in equilibrium with the atmosphere).

The saturation concentration for O_2 was calculated using the equation introduced by Garcia & Gordon[70,71]

for the solubility of O_2. It is noteworthy that Garcia & Gordon[70,71] determined the O_2 saturation value at an assumed atmospheric pressure of 1 atm, which means that their approach (Equation 3) needs to be modified to account for local in-situ sea level pressure (SLP) using a parallel equation (Equation 4):

$$O_{2,sat}^{1atm} = K \times pO_2^{1atm} = K \times xO_{2,air} \times (P_{1atm} - P_{sw}), \quad (3)$$

$$O_{2,sat}^{SLP} = K \times pO_2^{SLP} = K \times xO_{2,air} \times (P_{SLP} - P_{sw}), \quad (4)$$

where '1atm' and 'SLP' denotes two different pressures, K is the solubility of O_2, $O_{2,sat}^{1atm}$ is the result of the calculation based on Garcia & Gordon[70,71] methodology, P_{sw} is the water vapor pressure calculated from surface ocean temperature and salinity[72].

Substituting (3) into (4):

$$O_{2,sat}^{SLP} = O_{2,sat}^{1atm} \times (P_{SLP} – P_{sw})/(P_{1atm} - P_{sw}) \quad (5)$$

where $O_{2,sat}^{SLP}$ is the O_2 saturation value corrected for local sea level pressure and P_{sw}. P_{SLP} is from National Centers for Environmental Prediction/National Center for Atmospheric Research (NCRP/NCAR) reanalysis data at the time of the measurement (https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.surface.html).

In order to account for the impacts of bubble injection on O_2 saturation in the surface ocean, we applied a saturation anomaly of 0.75%[56] to O_2 saturation (i.e., the saturation of O_2 in the surface ocean is here calculated as 100.75% of the value from the saturation equations). The bubble injection effect on the more soluble CO_2 is negligible (less than 0.1%, ref. [73]) and therefore it is not applied to the CO_2 calculations.

The saturation concentration for CO_2 was calculated using Henry's Law ($[CO_{2,sat}] = K_H \times pCO_{2,eq}$), where $pCO_{2,eq}$ refers to the partial pressure of CO_2 in seawater when it is in equilibrium with atmospheric CO_2. This was calculated as: $pCO_{2,eq} = xCO_{2,air} \times (P_{SLP} - P_{sw})$, where $xCO_{2,air}$ is the mole fraction (ppm) of CO_2 in dry air. The values of $xCO_{2,air}$ are subject to spatiotemporal variabilities; we used the monthly mean atmospheric xCO_2 values for each ocean basin for each year from the NOAA/ESRL/Global Monitoring Division (ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/). We used data from the following monitoring sites: BMW (Tudor Hill, Bermuda) and ASC (Ascension Island) in the north and south Atlantic Ocean, respectively; SEY (Mahe Island) in the Indian Ocean; MLO (Mauna Loa, Hawaii) and SMO (Tutuila) in the north and south Pacific Ocean, respectively; and PSA (Palmer Station, Antarctica) in the Southern Ocean (>50°S). P_{SLP} and P_{sw} are described in Equations 3-5. The solubility (K_H) of CO_2 was calculated following Weiss[74].

Predicted effects of different processes.

ΔCO_2 and ΔO_2 exhibit some co-variation (Figs. 1 and 2) and so predicted joint effects of different processes were calculated (inset to Fig. 2d), based on global average sea surface conditions (unless specified otherwise) calculated from GLODAPv2 data: salinity of 34.6, temperature of 15°C, TA of 2300 µmol kg^{-1}, and atmospheric pCO_2 of 380 µatm (for the year 2005), which yields saturation concentrations of CO_2 and O_2 of 14.3 µmol kg^{-1} and 248.5 µmol kg^{-1} respectively, and DIC of 2072 µmol kg^{-1}.

To predict the effects of warming and cooling, we calculated saturation gas concentrations along a temperature gradient, and then compared the saturation value at 15°C (T_0) to that at another temperature (T_1) using Equation 1:

$$\Delta Gas = [Gas_{sat}]^{T_0} - [Gas_{sat}]^{T_1} \quad (6)$$

Instantaneous warming and cooling would have an immediate impact on the CORS values (ΔCO_2 and ΔO_2) because changes in temperature alter gas solubility. Warming decreases the gas solubility (equilibrium value), so warming increases both ΔCO_2 and ΔO_2, whereas cooling decreases them. The calculated molar ratio between changes in [$CO_{2,sat}$] and changes in [$O_{2,sat}$] for warming is 0.086 and for cooling is 0.091.

To predict the effect of ice melt, we assumed that ice contains so little dissolved CO_2 and O_2 that melting adds insignificant amounts[75,76], so [DIC] and [O_2] are subjected to the same degree of dilution during ice melting. A degree of dilution gradient (e.g., diluted by 5%, 10%, 20%, and 30%) was then assumed for salinity, TA, DIC, and O_2 at a temperature of -1.8°C, which is the average value for regions subject to ice melt. Each dilution step yielded a new carbonate system, for each of which [CO_2] was recalculated using CO2SYS. By comparing each of [CO_2] and [O_2] to their original values, the ratio between the changes induced by ice melt was calculated as 0.125.

To predict the effects of photosynthesis and respiration, DIC changes were made proportional to changes in O_2 of ±25, ±50, ±75, and ±100 µmol kg^{-1}. The corresponding DIC changes were calculated by multiplying the O_2 changes by the Redfield ratio[21] of DIC/O_2 = -117/170. So, for instance, DIC changed by -17.2 µmol kg^{-1} when O_2 changed by +25 µmol kg^{-1} due to photosynthesis. Using CO2SYS[62], [CO_2] was then calculated to change by -1.28 µmol kg^{-1} for this example. The same logic was applied to changes in [CO_2] and [O_2] due to respiration. The ratio of [CO_2] change to [O_2] change is not perfectly linear; the ratio (slope) of a line fitted to the results is -0.044 for photosynthesis and -0.067 for respiration. Since temperature

dominates solubility and carbonate system dissociation constants, and because spring blooms at high latitudes occur in cold water, we also made the same calculation at a temperature of 5°C, resulting in -0.070 for photosynthesis and -0.116 for respiration (shown by the grey shading in the inset to Fig. 2d). Although the Redfield ratio refers to changes in DIC and [O_2], our choice of plotting [CO_2] rather than DIC versus [O_2] has advantages: 1) the y-axis value on the CORS plot relates directly to the tendency for air-sea CO_2 exchange to occur; 2) assessing [CO_2] relative to saturation is quite straightforward whereas assessing [DIC] relative to its saturation value requires an additional step involving alkalinity.

Since there is spatial variation in the parameters (i.e., salinity, temperature, TA, and DIC) used to calculate the slopes for photosynthesis and respiration in different ocean basins, the theoretical slopes in Fig. 2a and c were calculated based on the average condition in each specific ocean basin.

The formation and dissolution of $CaCO_3$ affects [CO_2] but not [O_2] and is therefore only changing the ΔCO_2 axis. It is also noteworthy that in addition to the processes above, upwelling and entrainment of subsurface waters (characterized by CO_2-rich and O_2-depleted due to remineralization) also account for the departures from equilibrium of both gases, where these effects tend to drive CORS towards the second quadrant.

Subsampling treatment of GLODAP dataset.

Because the GLODAP and SOCCOM float datasets have different amounts of data, we adopted a subsampling approach in order to treat the two datasets equally. The procedure is described using the following scenario as an imagined example:

GLODAP Southern Ocean data from all seasons: N = 2500, y-intercept = -1.5;

Float dataset from all seasons: N = 300, y-intercept = -3.0;

To test statistically whether the second y-intercept is significantly lower than the first, we set up a null hypothesis (H0) that the float-derived y-intercept is less than or equal to the GLODAP-derived y-intercept. We repeatedly (10,000 times) took random subsamples (size N = 300) from the GLODAP Southern Ocean data, calculated the y-intercepts of the fitted lines, and then calculated the frequency with which y_float <= y_GLODAP (subsample). If frequency < 5% then y_float is significantly greater than y_GLODAP. Furthermore, we calculated the value of a term (y_diff) for the minimum significant difference in y-intercepts by subtracting y_float from y_GLODAP.

Data availability

GLODAPv2020 dataset was downloaded from the Ocean Carbon Data System (OCADS, https://www.nodc.noaa.gov/ocads/oceans/). The time-series data in BATS and HOT stations were downloaded from https://www.ncei.noaa.gov/access/ocean-carbon-data-system/oceans/time_series_moorings.html. The SOCCOM float data (Matlab formatted version, accessed on 10 November 2020) was obtained from https://soccom.princeton.edu/. The monthly mean atmospheric xCO_2 values for each observing site were obtained from ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/flask/.

Code availability

Matlab code for the analyses is available upon request to Y. Wu.

References

1. Brewer, P. G. & Peltzer, E. T. Limits to Marine Life. **324**, 347-348, doi:10.1126/science.1170756 (2009).
2. Friedlingstein, P. *et al.* Global Carbon Budget 2020. *Earth System Science Data* **12**, 3269-3340, doi:10.5194/essd-12-3269-2020 (2020).
3. Manning, A. & Keeling, R. F. Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network. *Tellus B: Chemical and Physical Meteorology* **58**, 95-116, doi:10.1111/j.1600-0889.2006.00175.x (2006).
4. Bates, N. *et al.* A time-series view of changing ocean chemistry due to ocean uptake of anthropogenic CO_2 and ocean acidification. *Oceanography* **27**, 126-141 (2014).
5. Sutton, A. J. *et al.* Natural variability and anthropogenic change in equatorial Pacific surface ocean pCO_2 and pH. *Global Biogeochemical Cycles* **28**, 131-145, doi:10.1002/2013GB004679 (2014).
6. Takahashi, T. *et al.* Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep Sea Research II* **56**, 554-577 (2009).
7. Breitburg, D. *et al.* Declining oxygen in the global ocean and coastal waters. *Science* **359**, eaam7240, doi:10.1126/science.aam7240 (2018).
8. Schmidtko, S., Stramma, L. & Visbeck, M. Decline in global oceanic oxygen content during the past five decades. *Nature* **542**, 335, doi:10.1038/nature21399 (2017).
9. Stramma, L., Johnson, G. C., Sprintall, J. & Mohrholz, V. Expanding oxygen-minimum zones in the tropical oceans. *Science* **320**, 655-658, doi:10.1126/science.1153847 (2008).
10. Carrillo, C. J., Smith, R. C. & Karl, D. M. Processes regulating oxygen and carbon dioxide in surface waters west of the Antarctic Peninsula. *Marine Chemistry* **84**, 161-179, doi:10.1016/j.marchem.2003.07.004 (2004).
11. Johnson, K. S. Simultaneous measurements of nitrate, oxygen, and carbon dioxide on oceanographic

moorings: Observing the Redfield ratio in real time. *Limnology and Oceanography* **55**, 615 (2010).

12. Körtzinger, A., Send, U., Wallace, D. W., Karstensen, J. & DeGrandpre, M. Seasonal cycle of O_2 and pCO_2 in the central Labrador Sea: Atmospheric, biological, and physical implications. *Global Biogeochemical Cycles* **22**, GB1014, doi:10.1029/2007GB003029 (2008).
13. Ogundare, M. O., Fransson, A., Chierici, M., Joubert, W. R. & Roychoudhury, A. N. Variability of sea-air carbon dioxide flux in autumn across the Weddell Gyre and offshore Dronning Maud Land in the Southern Ocean. *Frontiers in Marine Science* **7**, doi:10.3389/fmars.2020.614263 (2021).
14. Tortell, P. D. *et al.* Metabolic balance of coastal Antarctic waters revealed by autonomous pCO_2 and ΔO_2/Ar measurements. *Geophysical Research Letters* **41**, 6803-6810, doi:10.1002/2014GL061266 (2014).
15. Tortell, P. D., Bittig, H. C., Körtzinger, A., Jones, E. M. & Hoppema, M. Biological and physical controls on N_2, O_2, and CO_2 distributions in contrasting Southern Ocean surface waters. *Global Biogeochemical Cycles* **29**, 994-1013, doi:10.1002/2014GB004975 (2015).
16. Moreau, S. *et al.* Influence of microbial community composition and metabolism on airsea ΔpCO_2 variation off the western Antarctic Peninsula. *Marine Ecology Progress Series* **446**, 45-59 (2012).
17. Mu, L., Stammerjohn, S. E., Lowry, K. E. & Yager, P. L. Spatial variability of surface pCO_2 and air-sea CO_2 flux in the Amundsen Sea Polynya, Antarctica. *Elementa: Science of the Anthropocene* **3**, doi:10.12952/journal.elementa.000036 (2014).
18. Craig, H. & Hayward, T. Oxygen supersaturation in the ocean: biological versus physical contributions. *Science* **235**, 199-202, doi:10.1126/science.235.4785.199 (1987).
19. Guéguen, C. & Tortell, P. D. High-resolution measurement of Southern Ocean CO_2 and O_2/Ar by membrane inlet mass spectrometry. *Marine Chemistry* **108**, 184-194, doi:10.1016/j.marchem.2007.11.007 (2008).
20. Tortell, P. D. & Long, M. C. Spatial and temporal variability of biogenic gases during the Southern Ocean spring bloom. *Geophysical Research Letters* **36**, L01603, doi:10.1029/2008GL035819 (2009).
21. Anderson, L. A. & Sarmiento, J. L. Redfield ratios of remineralization determined by nutrient data analysis. *Global Biogeochemical Cycles* **8**, 65-80, doi:10.1029/93GB03318 (1994).
22. Redfield, A. C. The influence of organisms on the composition of sea-water. *The Sea*, 26-77 (1963).
23. Torgersen, T. & Branco, B. Carbon and oxygen dynamics of shallow aquatic systems: Process vectors and bacterial productivity. *Journal of Geophysical Research* **112**, doi:10.1029/2007JG000401 (2007).
24. Vachon, D. *et al.* Paired O_2–CO_2 measurements provide emergent insights into aquatic ecosystem function. *Limnology and Oceanography Letters* **5**, 287-294, doi:10.1002/lol2.10135 (2020).
25. Broecker, W. S. & Peng, T.-H. *Tracers in the Sea*. (Lamont-Doherty Geological Observatory, Columbia University, 1982).
26. Key, R. M. *et al.* Global Ocean Data Analysis Project, Version 2 (GLODAPv2). *ORNL/CDIAC-162, NDP-093*, doi:10.3334/CDIAC/OTG.NDP093_GLODAPv2 (2015).
27. Olsen, A. *et al.* The Global Ocean Data Analysis Project version 2 (GLODAPv2) – An internally consistent data product for the world ocean. *Earth System Science Data* **8**, 297-323, doi:10.5194/essd-8-297-2016 (2016).
28. Olsen, A. *et al.* An updated version of the global interior ocean biogeochemical data product, GLODAPv2.2020. *Earth System Science Data* **12**, 3653-3678, doi:10.5194/essd-12-3653-2020 (2020).
29. Johnson, K. S. *et al.* Biogeochemical sensor performance in the SOCCOM profiling float array. *Journal of Geophysical Research: Oceans* **122**, 6416-6436, doi:10.1002/2017JC012838 (2017).
30. Bushinsky, S. M., Gray, A. R., Johnson, K. S. & Sarmiento, J. L. Oxygen in the Southern Ocean from

Argo floats: Determination of processes driving air-sea fluxes. *Journal of Geophysical Research: Oceans* **122**, 8661-8682, doi:10.1002/2017JC012923 (2017).

31. Gray, A. R. *et al.* Autonomous biogeochemical floats detect significant carbon dioxide outgassing in the high-latitude Southern Ocean. *Geophysical Research Letters* **45**, 9049-9057, doi:10.1029/2018GL078013 (2018).

32. Bushinsky, S. M. *et al.* Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations. *Global Biogeochem Cycles* **33**, 1370-1388, doi:10.1029/2019GB006176 (2019).

33. Williams, N. L. *et al.* Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. *Global Biogeochemical Cycles* **31**, 591-604, doi:10.1002/2016GB005541 (2017).

34. Álvarez, M. *et al.* Global ocean spectrophotometric pH assessment: consistent inconsistencies. *Environmental Science & Technology* **54**, 10977-10988, doi:10.1021/acs.est.9b06932 (2020).

35. Mignot, A., D'Ortenzio, F., Taillandier, V., Cossarini, G. & Salon, S. Quantifying Observational Errors in Biogeochemical-Argo Oxygen, Nitrate, and Chlorophyll a Concentrations. *Geophysical Research Letters* **46**, 4330-4337, doi:10.1029/2018GL080541 (2019).36. Henson, S. A., Dunne, J. P. & Sarmiento, J. L. Decadal variability in North Atlantic phytoplankton blooms. *Journal of Geophysical Research: Oceans* **114**, C04013, doi:10.1029/2008JC005139 (2009).

37. Mahadevan, A., D'Asaro, E., Lee, C. & Perry, M. J. Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. *Science* **337**, 54-58, doi:10.1126/science.1218740 (2012).

38. Saito, H., Tsuda, A. & Kasai, H. Nutrient and plankton dynamics in the Oyashio region of the western subarctic Pacific Ocean. *Deep Sea Research Part II: Topical Studies in Oceanography* **49**, 5463-5486, doi:10.1016/S0967-0645(02)00204-7 (2002).

39. Shiomoto, A. Chlorophyll-a and primary production during spring in the oceanic region of the Oyashio Water, the north-western Pacific. *Journal of the Marine Biological Association of the United Kingdom* **80**, 343-354 (2000).

40. García-Reyes, M. & Largier, J. L. Seasonality of coastal upwelling off central and northern California: New insights, including temporal and spatial variability. *Journal of Geophysical Research: Oceans* **117**, C03028, doi:10.1029/2011JC007629 (2012).

41. Fry, C. H., Tyrrell, T., Hain, M. P., Bates, N. R. & Achterberg, E. P. Analysis of global surface ocean alkalinity to determine controlling processes. *Marine Chemistry* **174**, 46-57, doi:10.1016/j.marchem.2015.05.003 (2015).

42. Krumhardt, K. M., Long, M. C., Lindsay, K. & Levy, M. N. Southern Ocean calcification controls the global distribution of alkalinity. *global Biogeochem Cycles* **34**, e2020GB006727, doi:10.1029/2020GB006727 (2020).

43. Lumpkin, R. & Speer, K. Global ocean meridional overturning. *Journal of Physical Oceanography* **37**, 2550-2562, doi:10.1175/JPO3130.1 (2007).

44. Marshall, J. & Speer, K. Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience* **5**, 171-180 (2012).

45. Talley, L. D. Closure of the global overturning circulation through the Indian, Pacific, and Southern Oceans: Schematics and transports. *Oceanography* **26**, 80-97, doi:10.5670/oceanog.2013.07 (2013).

46. Talley, L. D., Reid, J. L. & Robbins, P. E. Data-based meridional overturning streamfunctions for the global ocean. *Journal of Climate* **16**, 3213-3226 (2003).

47. Long, M. C. *et al.* Strong Southern Ocean carbon uptake evident in airborne observations. *Science* **374**,

1275-1280, doi:10.1126/science.abi4355 (2021).

48. Mackay, N. & Watson, A. Winter air-sea CO_2 fluxes constructed from summer observations of the Polar Southern Ocean suggest weak outgassing. *Journal of Geophysical Research: Oceans*, e2020JC016600, doi:10.1029/2020JC016600 (2021).

49. Sutton, A. J., Williams, N. L. & Tilbrook, B. Constraining Southern Ocean CO_2 Flux Uncertainty Using Uncrewed Surface Vehicle Observations. *Geophysical Research Letters* **48**, e2020GL091748, doi:10.1029/2020GL091748 (2021).

50. Bushinsky, S. M., Takeshita, Y. & Williams, N. L. Observing changes in ocean carbonate chemistry: Our autonomous future. *Current Climate Change Reports* **5**, 207-220, doi:10.1007/s40641-019-00129-8 (2019).

51. Possenti, L. *et al.* Norwegian Sea net community production estimated from O_2 and prototype CO_2 optode measurements on a Seaglider. *Ocean Science Discussions* **2020**, 1-35, doi:10.5194/os-2020-72 (2020).

52. Wu, Y. *et al.* High-frequency time-series autonomous observations of sea surface pCO_2 and pH. *Limnology and Oceanography* **66**, 588-606, doi:10.1002/lno.11625 (2021).

53. Claustre, H., Johnson, K. S. & Takeshita, Y. Observing the Global Ocean with Biogeochemical-Argo. *Annual review of marine science* **12**, doi:10.1146/annurev-marine-010419-010956 (2020).

54. Riser, S. C. *et al.* Fifteen years of ocean observations with the global Argo array. *Nature Climate Change* **6**, 145-153, doi:10.1038/nclimate2872 (2016).

55. Takeshita, Y., Johnson, K. S., Martz, T. R., Plant, J. N. & Sarmiento, J. L. Assessment of Autonomous pH Measurements for Determining Surface Seawater Partial Pressure of CO_2. *Journal of Geophysical Research: Oceans* **123**, 4003-4013, doi:10.1029/2017JC013387 (2018).

56. Sarmiento, J. L. & Gruber, N. *Ocean Biogeochemical Dynamics*. (Princeton University Press, Princeton, New Jersey, USA, 2006).

57. Jones, D. C., Ito, T., Takano, Y. & Hsu, W.-C. Spatial and seasonal variability of the air-sea equilibration timescale of carbon dioxide. *Global Biogeochemical Cycles* **28**, 1163-1178, doi:10.1002/2014GB004813 (2014).

58. Chai, F. *et al.* Monitoring ocean biogeochemistry with autonomous platforms. *Nature Reviews Earth & Environment* **1**, 315-326, doi:10.1038/s43017-020-0053-y (2020).

59. Lee, K. *et al.* Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. *Geophysical Research Letters* **33**, L19605, doi:10.1029/2006GL027207 (2006).

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

61. Dittmar, T. & Kattner, G. The biogeochemistry of the river and shelf ecosystem of the Arctic Ocean: A review. *Marine Chemistry* **83**, 103-120, doi:10.1016/S0304-4203(03)00105-1 (2003).

62. Van Heuven, S., D. Pierrot, J.W.B. Rae, E. Lewis & Wallace, D. W. R. MATLAB program developed for CO_2 system calculations (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, 2011).

63. Lueker, T. J., Dickson, A. G. & Keeling, C. D. Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K1 and K2: validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. *Marine Chemistry* **70**, 105-119, doi:10.1016/S0304-4203(00)00022-0 (2000).

64. Dickson, A. G. Standard potential of the reaction: AgCl (s) + $12H_2$ (g) = Ag (s) + HCl (aq), and the

standard acidity constant of the ion HSO_4^- in synthetic sea water from 273.15 to 318.15 K. *The Journal of Chemical Thermodynamics* **22**, 113-127 (1990).

65. Lee, K. *et al.* The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. *Geochimica et Cosmochimica Acta* **74**, 1801-1811, doi:10.1016/j.gca.2009.12.027 (2010).

66. Orr, J. C., Epitalon, J.-M., Dickson, A. G. & Gattuso, J.-P. Routine uncertainty propagation for the marine carbon dioxide system. *Marine Chemistry* **207**, 84-107, doi:10.1016/j.marchem.2018.10.006 (2018).

67. Wu, Y., Hain, M. P., Humphreys, M. P., Hartman, S. & Tyrrell, T. What drives the latitudinal gradient in open-ocean surface dissolved inorganic carbon concentration? *Biogeosciences* **16**, 2661-2681, doi:10.5194/bg-16-2661-2019 (2019).

68. Carter, B. R. *et al.* Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. *Limnology and Oceanography: Methods* **16**, 119-131, doi:10.1002/lom3.10232 (2018).

69. Williams, N. L. *et al.* Empirical algorithms to estimate water column pH in the Southern Ocean. *Geophysical Research Letters* **43**, 3415-3422, doi:10.1002/2016GL068539 (2016).

70. Garcia, H. & Gordon, L. Erratum: Oxygen solubility in seawater: Better fitting equations. *Limnology and Oceanography* **38**, 656 (1993).

71. Garcia, H. E. & Gordon, L. I. Oxygen solubility in seawater: Better fitting equations. *Limnology and Oceanography* **37**, 1307-1312, doi:10.4319/lo.1992.37.6.1307 (1992).

72. Weiss, R. F. & Price, B. A. Nitrous oxide solubility in water and seawater. *Marine Chemistry* **8**, 347-359, doi:10.1016/0304-4203(80)90024-9 (1980).

73. Woolf, D. K. & Thorpe, S. A. Bubbles and the air-sea exchange of gases in near-saturation conditions. *Journal of Marine Research* **49**, 435-466, doi:10.1357/002224091784995765 (1991).

74. Weiss, R. F. Carbon dioxide in water and seawater: the solubility of a non-ideal gas. *Marine chemistry* **2**, 203-215, doi:10.1016/0304-4203(74)90015-2 (1974).

75. Loose, B., McGillis, W. R., Schlosser, P., Perovich, D. & Takahashi, T. Effects of freezing, growth, and ice cover on gas transport processes in laboratory seawater experiments. *Geophysical Research Letters* **36**, L05603, doi:10.1029/2008GL036318 (2009).

76. Nomura, D., Yoshikawa-Inoue, H. & Toyota, T. The effect of sea-ice growth on air–sea CO_2 flux in a tank experiment. *Tellus B: Chemical and Physical Meteorology* **58**, 418-426, doi:10.1111/j.1600-0889.2006.00204.x (2006).

Figure 1. Sea surface dissolved CO_2 and O_2 against sea surface temperature in the global dataset GLODAPv2. Each row is a different season; [CO_2] data are shown in (a, c, e, g) and [O_2] data in (b, d, f, h). The black dashed curves indicate the saturation values of [CO_2] or [O_2] (i.e., concentrations that would be in equilibrium with the atmosphere). The saturation curves for [CO_2] were calculated with respect to the atmospheric pCO_2 of 380 µatm in year 2005 and fitted. For this figure only, [CO_2] values measured in other years were adjusted to year 2005 following Wu et al.[67] to be consistent with the saturation values calculated. Colors indicate different ocean basins: Atlantic (magenta), Pacific (dark yellow), Indian (green) and Southern Ocean (blue). Dotted ovals with labels F1-F4 highlight major features, discussed in the Results section.

Figure 2. CORS plots: Carbon dioxide and oxygen concentrations relative to saturation in the global surface ocean in four seasons. Note the different axis scales for CO_2 and O_2. The inset in (d) shows the predicted effects of different processes (see Methods) on ΔCO_2 and ΔO_2: warming (W), cooling (C), ice melt (M), photosynthesis (P) and respiration (R), calcium carbonate precipitation (CP) and dissolution (CD). Grey shading shows the range of the P and R slopes for temperatures between 5°C and 15°C. The inset is proportional to the 4 subplots for directly comparing slope of the processes in the inset with those in the subplots.

Figure 3. Color-coded CORS plots for specific regions and seasons. (a) data from the Atlantic and Pacific Oceans in spring, colored by the concentration of in-situ nitrate; (b) data from the Southern Ocean in winter, colored by the nitrate anomaly (see text); and (c) data from the Southern Ocean in winter, colored by neutral density. Circles with solid black edges in (c) denote surface waters whereas circles without edges denote subsurface water (deeper than 30 m). Panel c uniquely contains subsurface as well as surface data. The black dashed lines in (a) and (b) are the best-fit straight-line regressions of all Pacific data (all four quadrants) and Southern Ocean data, respectively. The black solid line in (a) is the best-fit straight-line regression of data from the Atlantic, in the fourth quadrant. r is the associated Pearson correlation coefficient; n is the number of data points. The red dashed line in (a) is the expected slope due to respiration in the Pacific, and the red solid line in (a) is the expected slope due to photosynthesis in the

Atlantic. The red dashed line in (c) is the expected slope due to respiration in the Southern Ocean. The Subantarctic Mode Water and Antarctic Intermediate Water (SAMW/AAIW) in (c) are defined as water masses with neutral density ranging from 26.8 to 27.5 kg m^{-3}; Circumpolar Deep Water (CDW) is defined as neutral density ranging from 27.5 to 28.2 kg m^{-3}; and Antarctic Bottom Water (AABW) as neutral density greater than 28.2 kg m^{-3}.

Figure 4. Relationships between ΔCO_2 and ΔO_2 in the global ocean basins. The black dashed lines are the least-squares best-fit lines of data; unc denotes the uncertainty in y-intercept with 95% confidence level; r is the associated Pearson correlation coefficient; n is the number of data points.

Figure 5. CORS plots from data collected by SOCCOM float F9096 and F9099 in the high-latitude Southern Ocean. Circle with solid edge denotes ‘good’ flagged data, whereas cross denotes ‘questionable’ flagged data.

Figure 6. CORS plots from data collected by 12 floats in the Southern Ocean. The blue lines are the least-squares best-fit lines of data, green lines are the 95% confidence bounds for the fitted coefficients (Table 1); unc denotes the uncertainty in y-intercept with 95% confidence level; r is the associated Pearson correlation coefficient; n is the number of data points. The label on the top right of each subplot denotes the UW float ID number.

Table 1. Statistical analysis of best-fit lines to the CORS plots produced using data collected by 12 floats in the Southern Ocean. The regions are defined by ocean fronts (Supplementary Figure 9) following Gray et al[31]: Subantarctic Zone (SAZ), Polar-Frontal Zone (PFZ), and Antarctic-Southern Zone (ASZ). The fifth column (y_diff) is the difference in y-intercepts between GLODAP and float data. The sixth column converts the offsets in y-intercept (μmol kg^{-1}) to differences in pCO_2 (μatm) for an average sea surface temperature of 1°C.

Float ID number	Region	Fitted line: y = m × x + c (with 95% confidence bounds) m	c	y_diff*	pCO_2 difference (μatm) corresponding to y_diff
F9646	SAZ	-0.040 (-0.051, -0.029)	-1.99 (-2.09, -1.89)	0.74	12
F9666	SAZ	-0.097 (-0.108, -0.085)	-1.73 (-1.80, -1.65)	0.40	6
F0569	ASZ	0.003 (-0.004, 0.009)	-0.77 (-0.85, -0.70)	-0.50	-8
F12545	PFZ	-0.127 (-0.134, -0.120)	0.06 (-0.03, 0.14)	-1.30	-21
F9652	PFZ/ASZ	-0.144 (-0.152, -0.136)	-0.73 (-0.84, -0.61)	-0.54	-9
F9275	ASZ	-0.104 (-0.112, -0.095)	-3.81 (-4.22, -3.40)	2.47	40
F9096	ASZ	-0.106 (-0.118, -0.095)	0.99 (0.92, 1.07)	-2.27	-36
F9099	ASZ	-0.101 (-0.105, -0.097)	-1.37 (-1.52, -1.22)	0.12	2
F12575	SAZ	-0.044 (-0.051, -0.037)	-1.73 (-1.77, -1.70)	0.51	8
F9650	SAZ	-0.044 (-0.057, -0.032)	-0.73 (-0.79, -0.66)	-0.56	-9
F0690	SAZ	-0.070 (-0.075, -0.065)	-1.32 (-1.35, -1.29)	0.04	1
F12727	ASZ	-0.050 (-0.055, -0.046)	-0.73 (-0.85, -0.61)	-0.58	-9

*GLODAP-derived y-intercepts minus float-derived y-intercepts; negative (positive) values mean that the float y-intercept is greater (lower) than the GLODAP y-intercept.

Acknowledgements

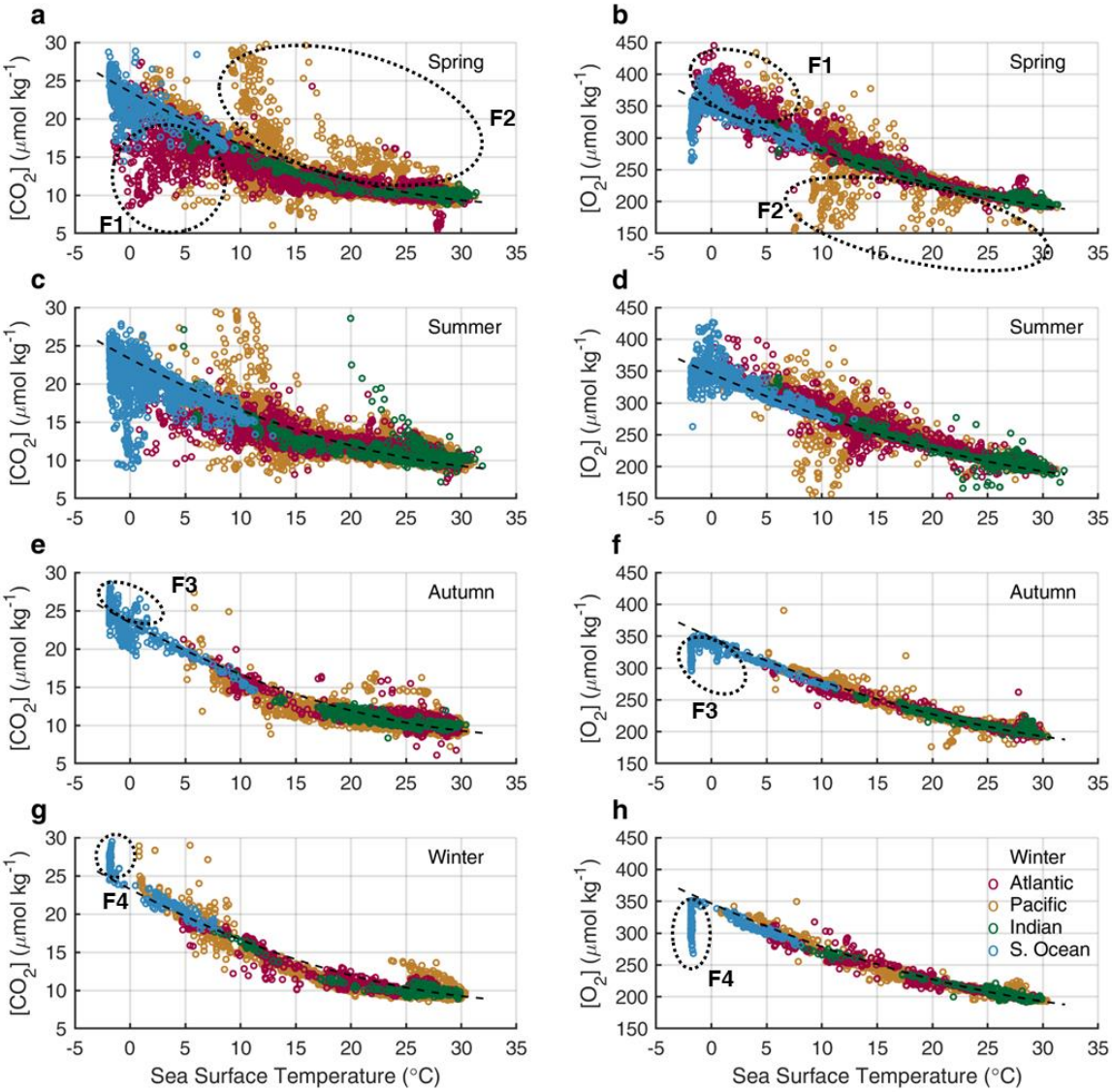
This work was jointly funded by the Swire Educational Trust, National Key Research and Development Program of China (2019YFE0114800, 2019YFC1509100), National Natural Science Foundation of China (42106222), Natural Science Foundation of Fujian Province, China (2020J05075), and Key Deployment Project of Centre for Ocean Mega-Research of Science, CAS (Grant No. COMS2020Q12). We acknowledge the work by the data providers and scientists carrying out data synthesis that led to the production of the datasets used here. We thank Wei-Jun Cai at University of Delaware for helpful and constructive discussions.

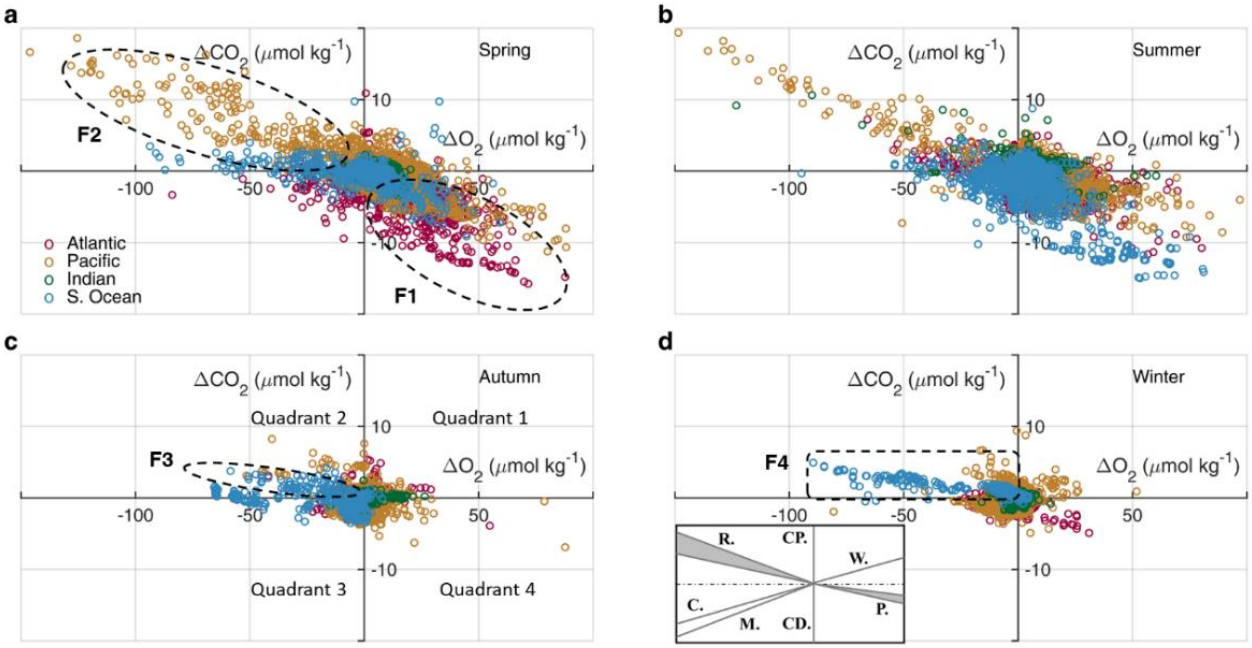
Author contributions

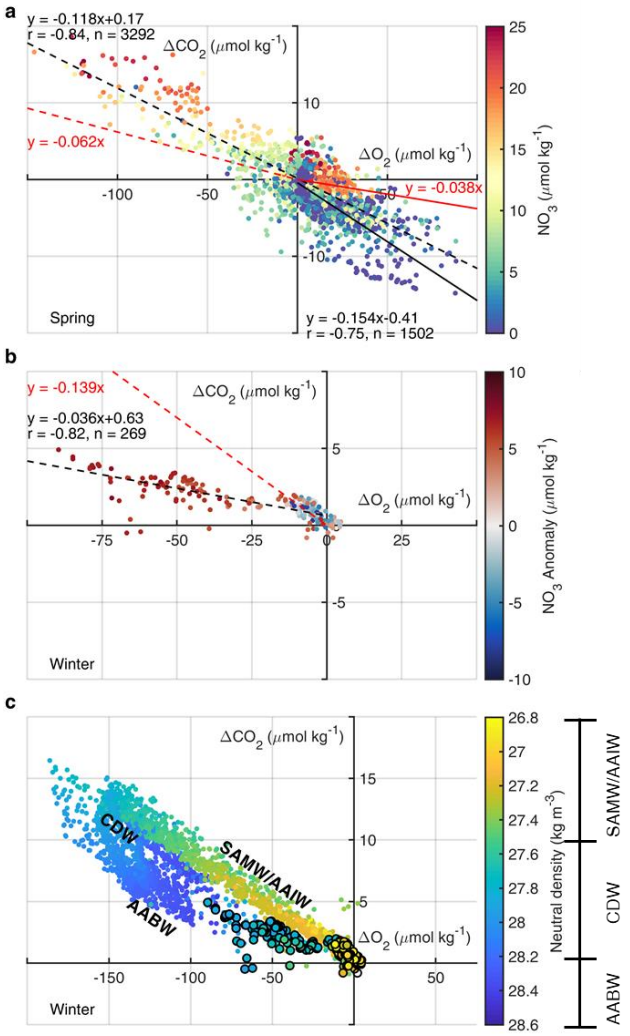
Y. W. and T. T. developed the theoretical formalism and conceived the original idea. Y. W. performed the analytic calculations, created the computational framework, and led the writing. T. T. supervised the project. Y. W., T. T., D. P., and S. H. contributed to data collection efforts. Y. W., D. Q., X. L., and T. T. are responsible for the statistical analyses. All authors provided constructive suggestions for further analyses and contributed to the writing of the manuscript.

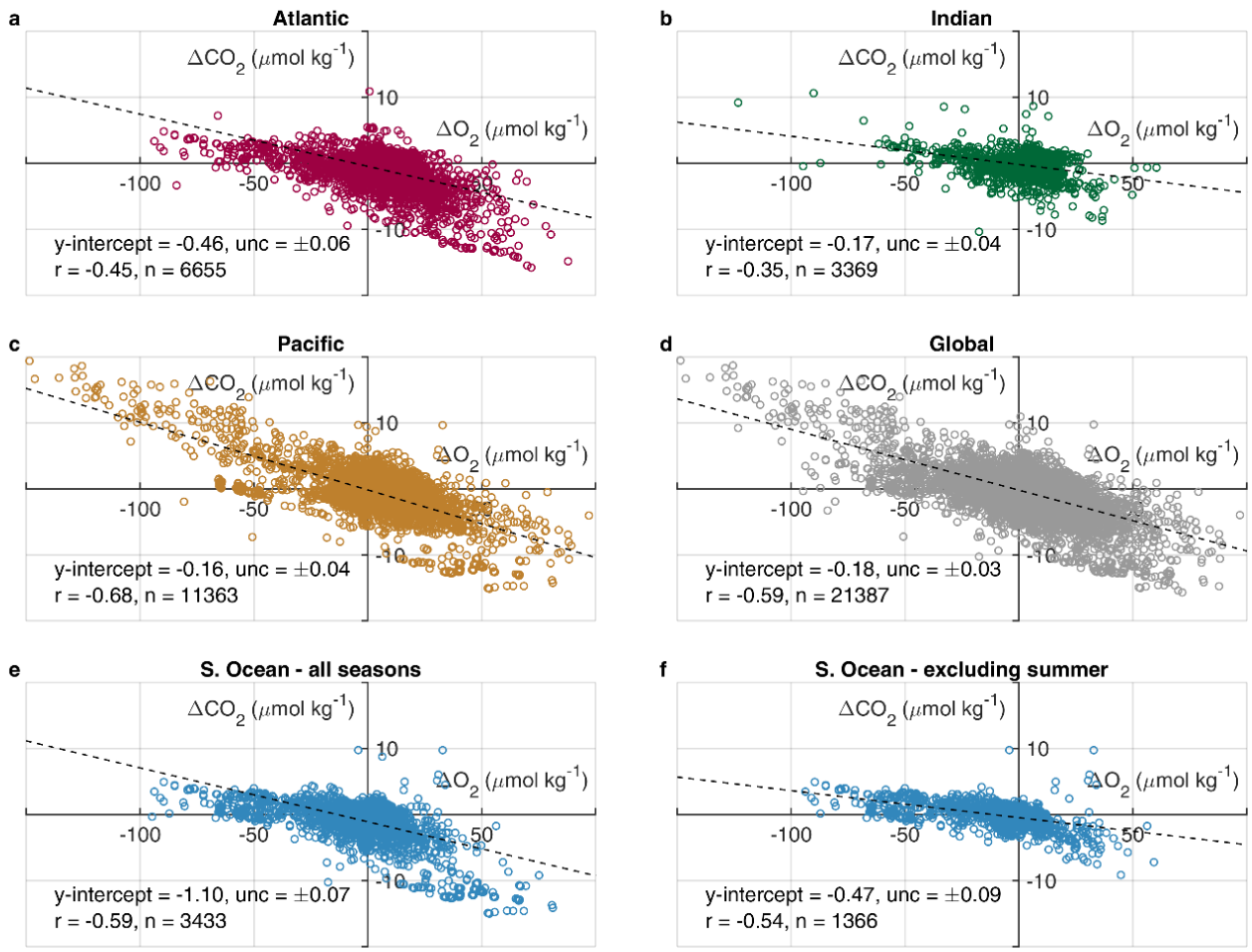
Competing interests

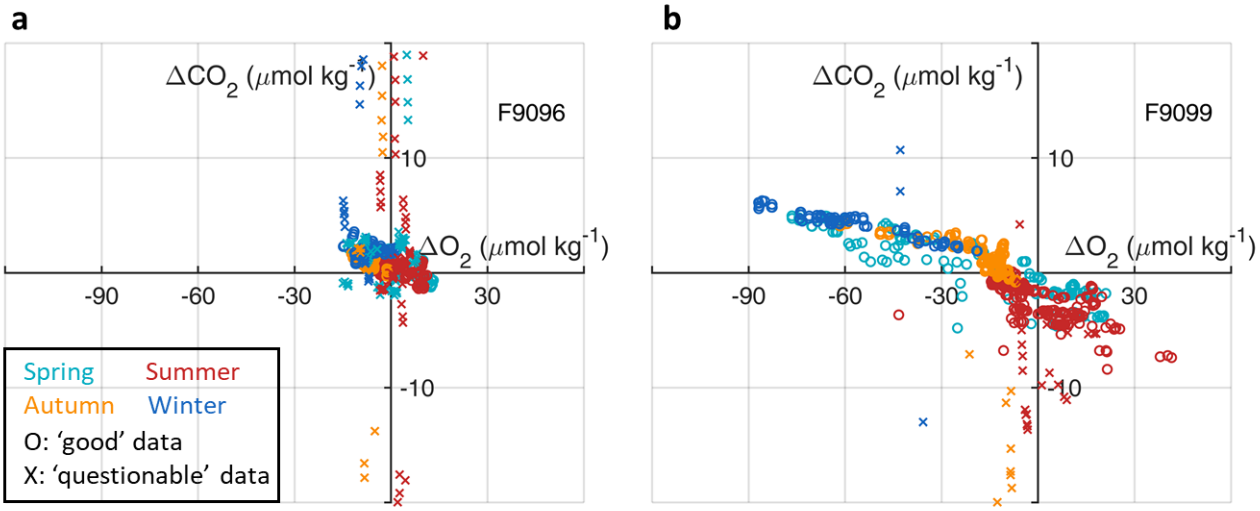
The authors declare no competing interests for both financial and non-financial sides.

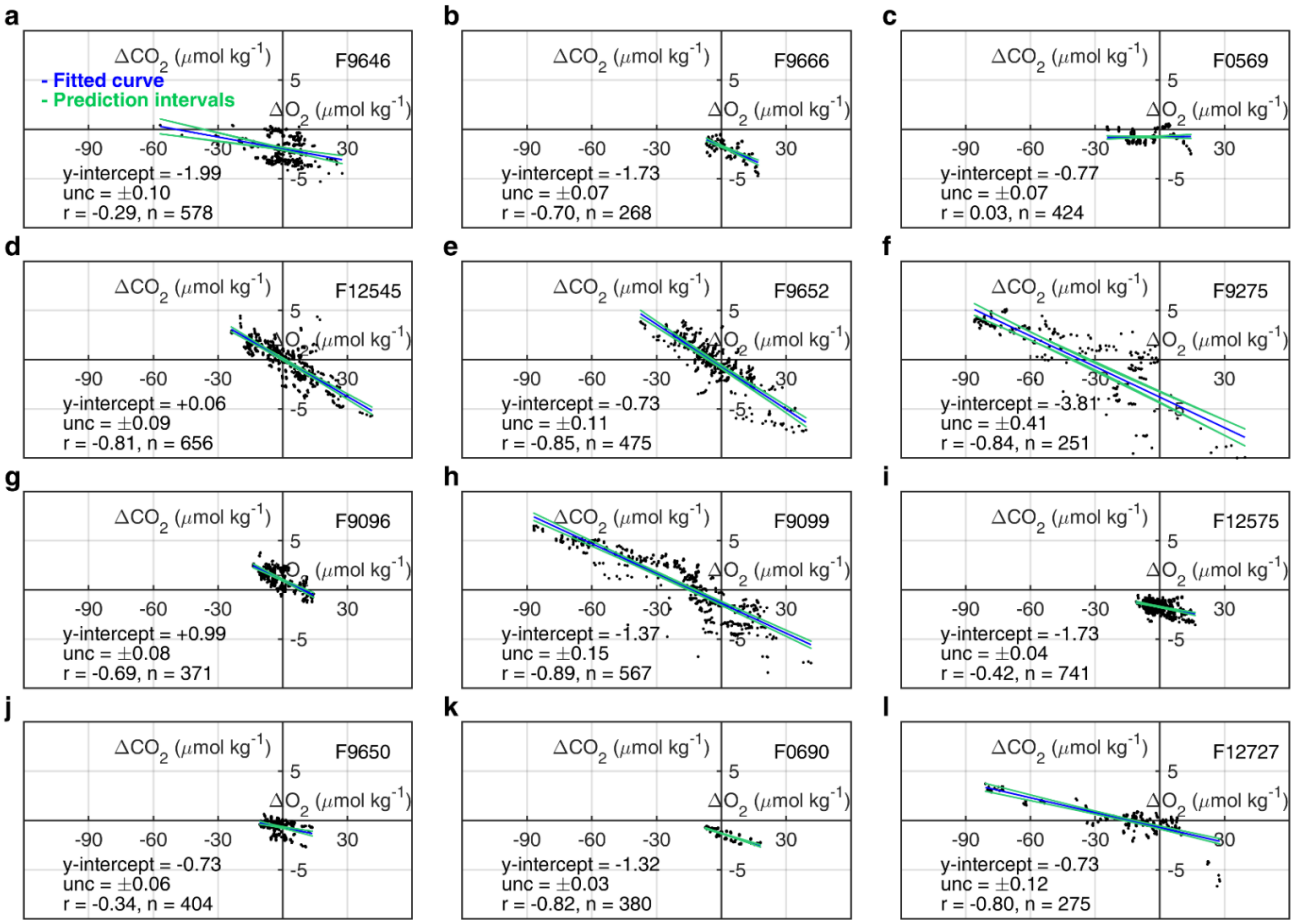

a
b
c
d
e
f
g
h
Spring
Summer
Autumn
Winter
F1
F2
F3
F4
[CO2] (μmol kg-1)
[O2] (μmol kg-1)
Sea Surface Temperature (°C)
Atlantic
Pacific
Indian
S. Ocean


a
Spring
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
F2
F1
Atlantic
Pacific
Indian
S. Ocean
b
Summer
c
Autumn
Quadrant 1
Quadrant 2
Quadrant 3
Quadrant 4
F3
d
Winter
F4
R.
CP.
W.
C.
M.
CD.
P.


a
y = -0.118x+0.17
r = -0.84, n = 3292
ΔCO2 (μmol kg-1)
y = -0.062x
ΔO2 (μmol kg-1)
y = -0.038x
y = -0.154x-0.41
r = -0.75, n = 1502
Spring
NO3 (μmol kg-1)
b
y = -0.139x
y = -0.036x+0.63
r = -0.82, n = 269
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
Winter
NO3 Anomaly (μmol kg-1)
c
ΔCO2 (μmol kg-1)
CDW
SAMW/AAIW
AABW
ΔO2 (μmol kg-1)
Winter
Neutral density (kg m-3)
SAMW/AAIW
CDW
AABW


a
Atlantic
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
50
y-intercept = -0.46, unc = ±0.06
r = -0.45, n = 6655
b
Indian
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
50
y-intercept = -0.17, unc = ±0.04
r = -0.35, n = 3369
c
Pacific
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
y-intercept = -0.16, unc = ±0.04
r = -0.68, n = 11363
d
Global
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
y-intercept = -0.18, unc = ±0.03
r = -0.59, n = 21387
e
S. Ocean - all seasons
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
50
y-intercept = -1.10, unc = ±0.07
r = -0.59, n = 3433
f
S. Ocean - excluding summer
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)
10
-10
-100
-50
50
y-intercept = -0.47, unc = ±0.09
r = -0.54, n = 1366


a
ΔCO2 (μmol kg-1)
F9096
10
ΔO2 (μmol kg-1)
-90
-60
-30
30
-10
Spring
Summer
Autumn
Winter
O: 'good' data
X: 'questionable' data
b
ΔCO2 (μmol kg-1)
F9099
10
ΔO2 (μmol kg-1)
-90
-60
-30
30
-10


a
F9646
- Fitted curve
- Prediction intervals
y-intercept = -1.99
unc = ±0.10
r = -0.29, n = 578
b
F9666
y-intercept = -1.73
unc = ±0.07
r = -0.70, n = 268
c
F0569
y-intercept = -0.77
unc = ±0.07
r = 0.03, n = 424
d
F12545
y-intercept = +0.06
unc = ±0.09
r = -0.81, n = 656
e
F9652
y-intercept = -0.73
unc = ±0.11
r = -0.85, n = 475
f
F9275
y-intercept = -3.81
unc = ±0.41
r = -0.84, n = 251
g
F9096
y-intercept = +0.99
unc = ±0.08
r = -0.69, n = 371
h
F9099
y-intercept = -1.37
unc = ±0.15
r = -0.89, n = 567
i
F12575
y-intercept = -1.73
unc = ±0.04
r = -0.42, n = 741
j
F9650
y-intercept = -0.73
unc = ±0.06
r = -0.34, n = 404
k
F0690
y-intercept = -1.32
unc = ±0.03
r = -0.82, n = 380
l
F12727
y-intercept = -0.73
unc = ±0.12
r = -0.80, n = 275
ΔCO2 (μmol kg-1)
ΔO2 (μmol kg-1)