JGR Atmospheres

RESEARCH ARTICLE
10.1029/2019JD031479

Summertime Atmospheric Boundary Layer Gradients of O_2 and CO_2 over the Southern Ocean

Eric J. Morgan[1], Britton B. Stephens[2], Matthew C. Long[2], Ralph F. Keeling[1], Jonathan D. Bent[2], Kathryn McKain[3,4], Colm Sweeney[4], Martín S. Hoecker-Martínez[5], and Eric A. Kort[5]

[1]Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, USA, [2]National Center for Atmospheric Research, Boulder, CO, USA, [3]Cooperative Institute for Research in Environmental Sciences, Boulder, CO, USA, [4]National Oceanic and Atmospheric Administration, Boulder, CO, USA, [5]Department of Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA

Key Points:
- We present atmospheric boundary layer gradients of oxygen and carbon dioxide over the Southern Ocean
- Gradients are driven primarily by air-sea fluxes due to biological production
- The ratio of these gradients can be used to approximate the flux ratio of the study region

Correspondence to:
E. J. Morgan,
ejmorgan@ucsd.edu

Citation:
Morgan, E., Stephens, B. B., Long, M. C., Keeling, R. F. F., Bent, J. D., McKain, K., et al. (2019). Summertime Atmospheric Boundary Layer Gradients of O_2 and CO_2 over the Southern Ocean. *Journal of Geophysical Research Atmospheres*, *124*, 13,439–13,456. https://doi.org/10.1029/2019JD031479

Received 7 AUG 2019
Accepted 10 NOV 2019
Accepted article online 18 NOV 2019
Published online 4 DEC 2019

Abstract We present airborne observations of the vertical gradient of atmospheric oxygen ($\delta(O_2/N_2)$) and carbon dioxide (CO_2) through the atmospheric boundary layer (BL) over the Drake Passage region of the Southern Ocean, during the O_2/N_2 Ratio and CO_2 Airborne Southern Ocean Study, from 15 January to 29 February 2016. Gradients were predominately anticorrelated, with excesses of $\delta(O_2/N_2)$ and depletions of CO_2 found within the boundary layer, relative to a mean reference height of 1.7 km. Through analysis of the molar ratio of the gradients (GR), the behavior of other trace gases measured in situ, and modeling experiments with the Community Earth System Model, we found that the main driver of gradients was air-sea exchange of O_2 and CO_2 driven by biological processes, more so than solubility effects. An exception to this was in the eastern Drake Passage, where positive GRs were occasionally observed, likely due to the dominance of thermal forcing on the air-sea flux of both species. GRs were more spatially consistent than the magnitudes of the gradients, suggesting that GRs can provide integrated process constraints over broad spatial scales. Based on the model simulation within a domain bounded by 45°S, 75°S, 100°W, and 45°W, we show that the sampling density of the campaign was such that the observed mean GR (± standard error), −4.0± 0.8 mol O_2 per mol CO_2, was a reasonable proxy for both the mean GR and the mean molar ratio of air-sea fluxes of O_2 and CO_2 during the O_2/N_2 Ratio and CO_2 Airborne Southern Ocean Study.

Plain Language Summary Using an aircraft, we measured changes in atmospheric oxygen and carbon dioxide over the surface of the Southern Ocean. We show that these changes can be used to estimate the relative amounts of oxygen and carbon dioxide that the Southern Ocean exchanges in summer.

1. Introduction

The Southern Ocean is a globally important sink for anthropogenic CO_2 (Khatiwala et al., 2009). Several recent studies suggest that the Southern Ocean is absorbing more atmospheric CO_2 than in previous decades (Ito et al., 2015; Landschützer et al., 2015; Munro et al., 2015). However, CO_2 air-sea exchange in this region shows considerable interannual and interdecadal variability, is spatially heterogeneous (Takahashi et al., 2009; Lenton et al., 2013; Gruber et al., 2019), and is not well constrained by observations (Monteiro et al., 2015), adding uncertainty to current understanding and future predictions of the sink. Simulations of the predicted air-sea flux of CO_2 from Earth system models involved in the Coupled Model Intercomparison Project version 5 found a large disparity in magnitude and phase of the seasonal cycle in the simulated air-sea flux of CO_2 (Anav et al., 2013; Mongwe et al., 2018). The need for additional observational constraints motivated the O_2/N_2 Ratio and CO_2 Airborne Southern Ocean Study (ORCAS) (Stephens et al., 2018), an airborne campaign centered over the Drake Passage during a 6-week period in late austral summer of 2016. In this study we present an analysis of the atmospheric O_2 and CO_2 data from this campaign, to better understand the balance of marine thermal and biological forcing on CO_2 air-sea fluxes.

Air-sea exchange of CO_2 occurs whenever there is a difference in the partial pressure of CO_2 between the ocean and the atmosphere. There are four main drivers of disequilibrium in the Southern Ocean: (1) the increase in atmospheric CO_2, mainly as a result of fossil fuel burning; (2) net photosynthesis and respiration; (3) changes in solubility driven by variations in water temperature; and 4) ocean and atmospheric transport,

which can bring unequilibrated air and water parcels to the air-sea interface (Lauderdale et al., 2016; Bopp & Le, 2009; IPCC, 2013; Gruber et al., 2019). Complementary observations of O_2 can be a useful diagnostic for distinguishing the relative importance of these main drivers for the net air-sea CO_2 flux, because the global oxygen cycle has many commonalities with the global carbon cycle. (Keeling and Shertz, 1992; Keeling et al., 1993; Keeling et al., 1996). There are key differences in how the main processes that drive CO_2 exchange affect O_2 exchange. For instance, net photosynthesis and respiration cause opposing exchanges of O_2 and CO_2, while solubility-driven changes produce O_2 and CO_2 exchanges of the same sign.

The relative exchange of O_2 and CO_2 can be expressed as a flux or molar exchange ratio, F_{O_2}/F_{CO_2}, where both F_{O_2} and F_{CO_2} are in units of flux density (e.g., mol·m^{-2}·s^{-1}). F_{O_2}/F_{CO_2} varies depending on the process driving the flux; for example, terrestrial exchange typically produces a value close to −1.1, and combustion of liquid fuels a value of −1.4 (Keeling and Manning, 2014). In contrast, the F_{O_2}/F_{CO_2} of air-sea exchange is highly variable, due to the different equilibration times of both gases, and can be very positive or very negative.

In the marine atmospheric boundary layer (BL), we expect fluxes to drive anomalies of O_2 and CO_2 relative to the overlying troposphere, creating an observable vertical gradient, Δ. Under ideal conditions, the gradient ratio (GR) of the two species, $\Delta O_2/\Delta CO_2$, should approximate F_{O_2}/F_{CO_2}, provided the background is uniform and slowly varying. In such a situation the GR can inform on flux processes; for instance, under ideal conditions a positive GR over the ocean would indicate the predominance of thermally driven air-sea fluxes on the variation of both species, as temperature changes increase or decrease the solubility of both gases.

Interpreting GRs in practice requires consideration of atmospheric mixing and the temporal and spatial scales involved in comparing air from the free troposphere with the BL. Interpretations of GRs may not require a precise constraint on atmospheric transport, however, since both species will be equally affected, removing the effect on the GR. In the Southern Ocean, where the CO_2 flux signals are small, and meteorology and prior flux fields are not well-constrained by observations, the GRs of O_2 and CO_2 have the potential to provide an additional constraint on the processes controlling CO_2 exchange.

In this work we combine ORCAS airborne measurements of gradients and GRs of atmospheric O_2 and CO_2, atmospheric transport simulations with the Stochastic Time-Inverted Lagrangian Transport (STILT) model, and a simulation of the Community Earth System Model (CESM) to assess the temporal and spatial scales over which air-sea fluxes and BL gradients were coupled. While we do not estimate regional fluxes of CO_2 or O_2 directly from the observations, we determine a regional domain over which the GR approximates F_{O_2}/F_{CO_2}. We then use these data sets to determine how much temporal coverage is needed to resolve the summertime mean gradients in the Drake Passage region, discuss the extent to which gradients across the BL in a remote area of the Southern Ocean were controlled by air-sea fluxes, and discuss whether fluxes during late summer were dominated by thermal or biological processes.

1.1. Box Model Illustration

To illustrate how ocean processes can combine to produce varying F_{O_2}/F_{CO_2}, we create a simple scenario with the mixed layer box model described by Keeling et al. (1993), where the annual cycle in biological and thermal forcing within the ocean mixed layer are represented by sinusoidal functions, from which air-sea fluxes are derived. The ocean mixed layer depth (MLD) is fixed, as is the gas exchange velocity. A realization of this box model is shown in Figure 1, with the amplitude and phasing of the thermal and biological forcing of O_2 and CO_2 tuned to match a single harmonic fit to the summed fluxes of O_2 and CO_2 for the Drake Passage, from a simulation of the CESM (see below). Since the thermal and biological forcing are in phase and of the same sign for O_2, the amplitude of the total flux is larger than for that of CO_2, where the thermal and biological forcing largely compensate one another. F_{CO_2} is also out of phase with F_{O_2} due to the ~20-fold slower mixed layer equilibration rate of CO_2 compared with O_2. During the late summer months, F_{O_2} approaches 0 more rapidly than F_{CO_2}, and F_{O_2}/F_{CO_2} becomes less negative as a result.

The box model also allows for an examination of the sensitivity of F_{O_2}/F_{CO_2} to key parameters. Deepening the MLD or lowering the wind speed drives the flux ratio to be more negative than in the normal case, because changes in the mixed layer concentration of O_2 are more sensitive to these parameters than the concentration of CO_2 is. A twofold increase in the ratio of thermal to biological forcing changes the sign of the F_{O_2}/F_{CO_2} flux ratio from negative to positive, as this changes the sign of F_{CO_2} while increasing F_{O_2}. A

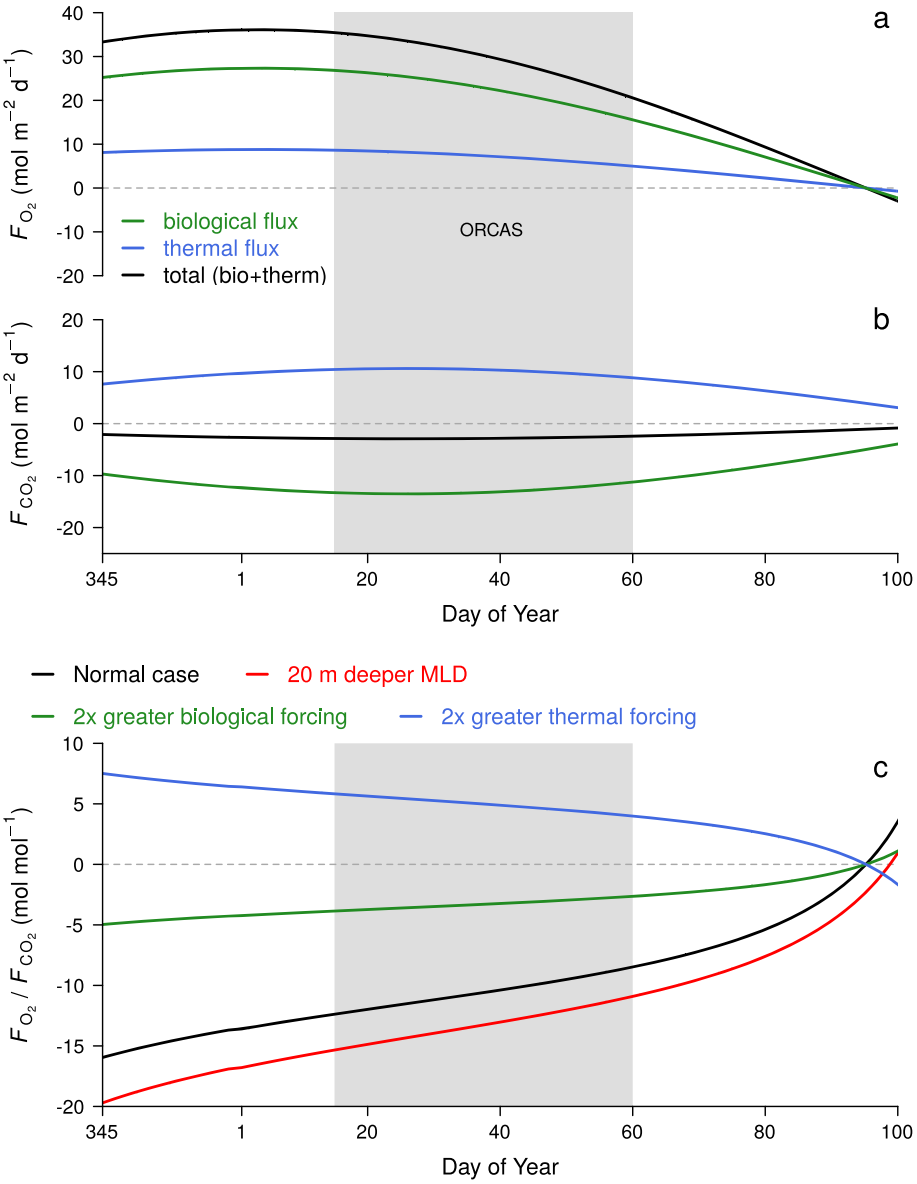

Figure 1. Realizations of a simple box model, tuned to match a one harmonic fit to the seasonal cycle of F_{O_2} (a) and F_{CO_2} (b) for the Drake Passage region. F_{O_2}/F_{CO_2} is shown in (c), with several cases in which variables are perturbed.

measurement of F_{O_2}/F_{CO_2} is thus diagnostic of the ratio of thermal to biological forcing, assuming the other physical parameters, such as gas exchange velocity and MLDs, are constrained.

2. Methods

2.1. The ORCAS Campaign

ORCAS was a 6-week airborne campaign based out of Punta Arenas, Chile, from 15 January to 29 February 2016; for an overview of the campaign, see Stephens et al. (2018). Nineteen research flights were conducted with the National Center for Atmospheric Research (NCAR) and the National Science Foundation (NSF) Gulfstream V (GV) aircraft (UCAR/NCAR - Earth Observing Laboratory, 2005). Eight flights were conducted primarily over the Drake Passage and the Antarctic Peninsula, two flights were over the Patagonian Shelf, and three were along the western coast of Chile. The remaining six flights were over southern Chile, when high winds required the relocation of the aircraft. Flight tracks were designed to characterize as much of the region as possible, both vertically and horizontally. During flight, the aircraft ascended and descended

whenever feasible. While a single descent or ascent is known as a vertical profile maneuver, or simply "profile," in this work we treat any ascent following a descent as a single profile (see section 2.6 for details). Excluding test flights and ferry flights, there were 74 instances in which the aircraft descended to below 300 m south of 35°S over the ocean; data from these profiles are the basis of this work.

2.2. Airborne Measurements of Atmospheric Oxygen

We report changes in the abundance of atmospheric O_2 in terms of relative changes in the O_2/N_2 ratio, in per meg (Keeling and Shertz, 1992; Keeling et al., 1998), that is,

$$\delta(O_2/N_2) = \left(\frac{(O_2/N_2)_{sample}}{(O_2/N_2)_{reference}} - 1 \right) \times 10^6 \quad (1)$$

where the *sample* and *reference* subscripts refer to sample air and a reference value, respectively.

In situ airborne observations of $\delta(O_2/N_2)$ were made with the NCAR Airborne Oxygen Instrument (AO2), a rapid-response instrument that employs a vacuum ultraviolet absorption technique (Stephens et al., 2003, see also https://www.eol.ucar.edu/instruments/airborne-oxygen-instrument). A radio frequency-powered xenon lamp emits UV light at a wavelength of 147 nm, and the absorption of this light is measured with a photocathode. The instrument switches between sample air and a reference tank every 2.5 s; the difference between successive jogs of these two gas streams is a measure of sample gas O_2 content. AO2 has a typical average precision of ± 2.5 per meg (1σ) for a given 5 s measurement interval. All values are reported on the Scripps O_2 Scale, as defined on 1 December 2017.

In addition to these in situ measurements of $\delta(O_2/N_2)$, flask samples were collected with the NCAR/Scripps Medusa Flask Sampler (Bent, 2014, see also https://www.eol.ucar.edu/instruments/ncar-scripps-airborne-flask-sampler). Flasks were analyzed at Scripps Institution of Oceanography for $\delta(O_2/N_2)$ on a Micromass IsoPrime isotope ratio mass spectrometer (IRMS), and for CO_2 on a LI-COR LI-6252 nondispersive infrared analyzer. The IRMS analytical approach was first described by Bender et al. (1994), and further developed by Keeling et al. (2004).

The AO2 and Medusa sample air streams were drawn through a fin-shaped HIAPER Modular Inlet (HIMIL; see https://www.eol.ucar.edu/instruments/hiaper-modular-inlet), consisting of aft facing 6.4-mm OD, 4.6-mm ID tubes extending 46 cm from the aircraft supported by a pylon with an aerodynamic cross-section. The stainless steel tubing extended 5 mm outside of the fin. Both the AO2 and Medusa sample air streams were dried cryogenically with electropolished stainless steel traps, which were partially immersed in a dry ice slurry, sufficient to cool the traps to a temperature of ∼ −78.5°C for at least 10 hr.

Comparisons of measured $\delta(O_2/N_2)$ between AO2 and Medusa show linear drift on hourly time scales in AO2 that is not observed in Medusa, a full discussion of which is beyond the scope of this paper. AO2 is sensitive to even > 1 ppm (parts per million) of water vapor (due to interactions with the windows in a high UV environment), and we attribute this drift to drying of tank regulators and the sample gas line. Since flasks do not share this sensitivity, we correct AO2 on a per flight basis, using linear time-of-flight fits to the difference between the two instruments. These correction were −0.8 per meg/hr on average for ORCAS, not larger than −3 per meg/hr. These corrections do not affect the BL gradients used here, since we are only considering relative changes in AO2 data over the ∼10 min it took to profile between ∼2 km and 0.15 km (see section 2.6). An example profile of $\delta(O_2/N_2)$ from AO2 is shown in Figure 2.

Due to the large pressure gradients surrounding aircraft inlets, the potential for fractionation of O_2 relative to N_2 is high (Steinbach, 2010; Langenfelds, 2002; Bent, 2014; Ishidoya et al., 2014). Since we also measure $\delta(Ar/N_2)$ in our flask samples, this allowed us to test our data for artifacts. Since fractionation caused by mass-dependent processes should scale according to the ratio of the mass differences of Ar/N_2 and O_2/N_2, any (mass-dependent) artifact in $\delta(O_2/N_2)$ should be 3 times as large in $\delta(Ar/N_2)$. Furthermore, we expect that in the absence of artifacts, low-altitude $\delta(Ar/N_2)$ measurements should have a spread comparable to surface flasks records within the same region.

We found no statistically significant increase or decrease between 0.15 and 2 km in $\delta(Ar/N_2)$, using a two-sided t test of the linear model ($t = -0.986$, $p = 0.33$, DF = 121), and the standard deviation of all samples from this altitude range was 10.6 per meg, which is within the range of expected latitudinal and synoptic variability (Battle et al., 2003; Keeling et al., 2004; Blaine, 2005; Bent, 2014). This supports the interpretation

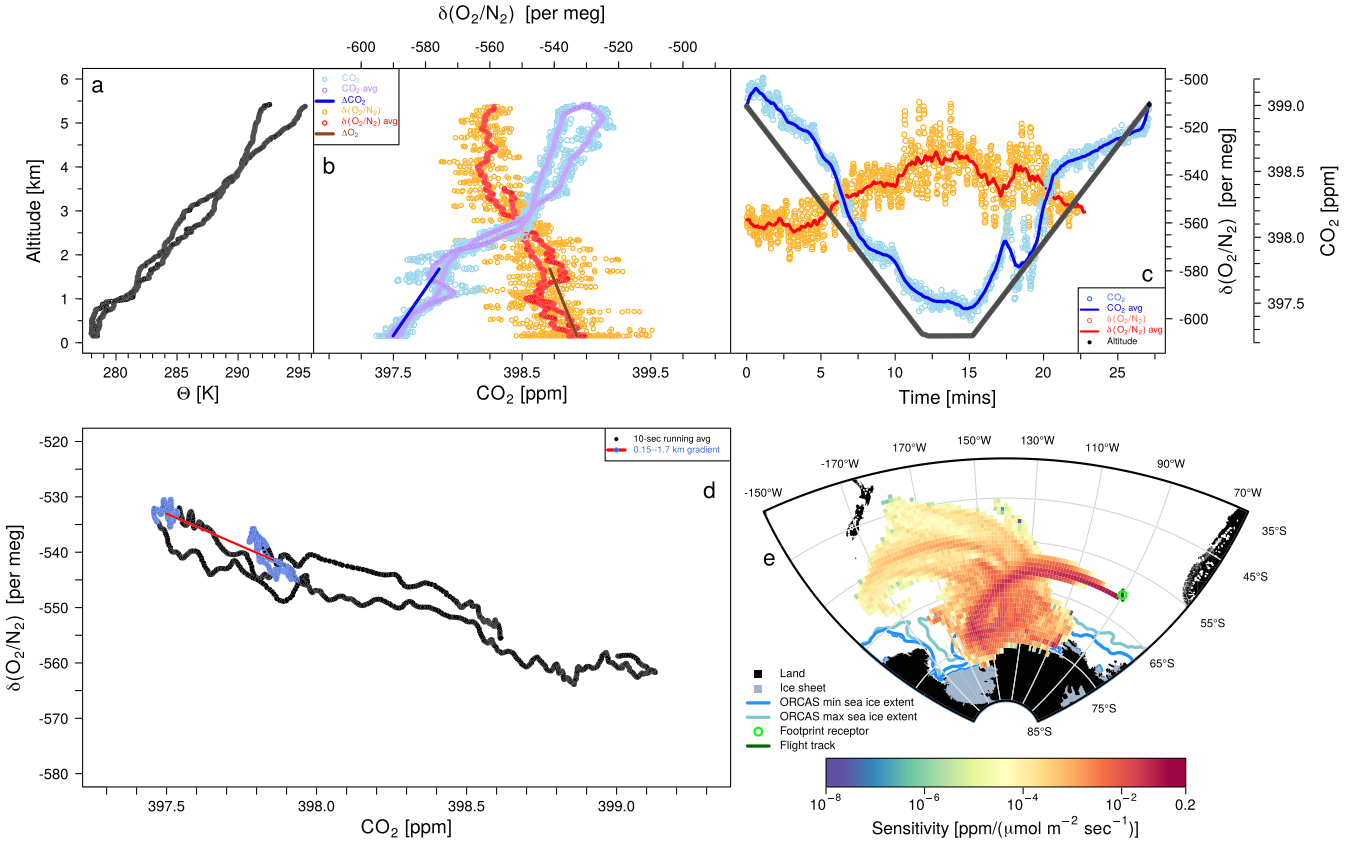

Figure 2. An example of observed vertical profile to the west of the Drake Passage, showing (a) potential temperature (Θ), (b) $\delta(O_2/N_2)$ and CO_2, and (c) $\delta(O_2/N_2)$ and CO_2, overlain with the altitude of the aircraft, as a function of time. The average Δ value is depicted in (b) and (d) for both species as a line. In (b)–(d), 60-s running averages of $\delta(O_2/N_2)$ and CO_2 are shown as lines. In (d), data from the same time period but with $\delta(O_2/N_2)$ as a function of CO_2. Data used in the calculation of the GR are colored blue, and the GR is depicted as the slope of the red line. In (e), a 7-day STILT footprint is shown for a single receptor point (indicated by a black circle on the map) at the bottom of the profile. The shading indicates the relative strength of the sensitivity of the receptor point to surface fluxes.

that the gradients in $\delta(O_2/N_2)$ as measured from flasks are robust with respect to inlet fractionation. However, we do note that for tropospheric air above 5 km, $\delta(Ar/N_2)$ did show signs of inlet fractionation, and that all of the data are ~30 per meg more positive in $\delta(Ar/N_2)$ than both Scripps Institution of Oceanography surface flask data and previous campaigns conducted with the same aircraft but a different sample inlet (Bent, 2014). From this we conclude that our relative $\delta(O_2/N_2)$ gradients are unaffected, but conclusions drawn from the absolute value of $\delta(O_2/N_2)$ and $\delta(Ar/N_2)$ data from ORCAS need to be considered with respect to this artifact.

2.3. Airborne CO_2 Measurements

AO2 also measures CO_2 with a single-cell nondispersive infrared analyzer, the LI-COR LI-840. Two other in situ instruments also measured CO_2 on ORCAS: the Harvard Quantum Cascade Laser Spectrometer (QCLS) (Santoni et al., 2014), and the NOAA Picarro G2401-m cavity ringdown spectrometer (CRDS) (Karion et al., 2013). CO_2 was measured in flask air collected by the Medusa Air Sampler (see section 2.2). The NOAA Picarro (see http://data.eol.ucar.edu/datafile/nph-get/490.008/NOAA-Picarro_ORCA_readme.pdf) was selected as the primary source of in situ CO_2 data, because it had the best stability, the best short-term precision of all the measurements made, and the best sample coverage, due to the fewer number of in-flight calibrations required for this instrument.

Sample air was introduced through rear-facing, 3.2-mm outer diameter stainless steel tubing mounted inside a standard diffusing type HIMIL, via a pump located downstream of the analyzer. The cell pressure was lowered from its standard setpoint of 140 to 80 torr, to allow for operation across the full ambient pressure range of the campaign without upstream pressure control. A correction for cell pressure variability was made during post-processing of the data. Measurements were made every 1.2 s. The instrument was operated

without drying the sample air stream; a campaign-specific empirical water correction function was generated through laboratory tests before and after ORCAS. The response of the analyzer was calibrated in the lab against six tanks with known concentrations spanning 388–460 ppm CO_2 before and after the ORCAS campaign. Two high-pressure reference cylinders were measured for 3.5 min every other hour during flight and immediately before and after each flight. These in situ calibration measurements were used to apply a time-dependent offset correction to the measurements from each flight. Average total analytical uncertainty was 0.07 ppm for 1-s measurements and 0.06 ppm for 10-s average measurements. CO_2 dry air mole fractions are reported in ppm, on the World Meteorological Organization X2007 scale. An example profile of CO_2 from the NOAA Picarro is shown in Figure 2.

2.4. STILT Model

To assess the influence of recent surface fluxes on CO_2 and $\delta(O_2/N_2)$ gradients, atmospheric footprints were calculated with the STILT model (Lin et al., 2003). The footprint is the sensitivity of a chosen point (known as the "receptor," the point from which particles are released) to upwind surface fluxes; that is, how much the mole fraction of a trace gas would be changed by a given unit of flux. Footprint units are in change in mixing ratio per unit of flux density, for example, ppm / ($\mu mol \cdot m^{-2} \cdot s^{-1}$).

Footprints were determined approximately every 2 min along the flight track. For each receptor, 4,096 particles were released backward for a total of 7 days with a 10-min time step. In some cases we present footprints that are summed over a shorter time interval than 7 days. As STILT is a Lagrangian model, it has no computational grid; footprints were gridded at hourly intervals and $0.94 \times 1.25°$ (latitude × longitude) to match the resolution of the CESM grid. STILT was driven with meteorological fields from the National Center for Environmental Prediction's (NCEP) Global Data Assimilation System (GDAS) product (Kanamitsu, 1989) with a spatial resolution of $0.5 \times 0.5°$, 55 hybrid sigma-pressure levels, and a temporal resolution of 3 hr.

2.5. CESM

Simulated CO_2 and O_2 surface fluxes and atmospheric mole fractions of CO_2 and O_2 were taken from the results of a run of the CESM (Hurrell et al., 2013). This was a fully coupled integration of the Parallel Ocean Program version 2 model, a land surface model (Community Land Model, version 4.5), and the Community Atmosphere Model version 5 (CAM5). Air temperature and wind fields were nudged to Goddard Earth Observing System version 5 (GEOS-5) initialized forecasts. CO_2 fluxes from fossil fuel combustion (CO_2^{ff}), air-sea exchange (CO_2^{ocn}), and the terrestrial biosphere (CO_2^{lnd}) were carried as separate tracers, which were summed to produce an atmospheric CO_2 tracer (CO_2^{mod}) that combined all of these influences and could be compared to observed CO_2:

$$CO_2{}^{mod} = CO_2{}^{ff} + CO_2{}^{lnd} + CO_2{}^{ocn} \tag{2}$$

with all variables in units of ppm.

O_2 was treated as a trace gas. Since only marine surface fluxes of O_2 were simulated, CO_2^{ff} and CO_2^{lnd} were multiplied by their average F_{O_2}/F_{CO_2} values, −1.4 and −1.1, respectively (Keeling and Manning, 2014). This enabled us to create an atmospheric $\delta(O_2/N_2)$-like product that could be compared directly to the observations (Stephens et al., 1998):

$$\delta\left(O_2/N_2\right)^{mod} = \frac{-1.4\left(CO_2{}^{ff}\right) - 1.1\left(CO_2{}^{lnd}\right) + O_2{}^{ocn}}{X_{O_2}} - \frac{N_2{}^{ocn}}{X_{N_2}} \tag{3}$$

where O_2^{ocn} is the atmospheric tracer for O_2 from air-sea exchange, and N_2^{ocn} is the tracer for atmospheric variability of N_2 due to abiotic air-sea fluxes of N_2. The latter was taken from a preindustrial simulation of CESM and averaged into an atmospheric climatology, in order to estimate the small contribution of atmospheric N_2 variations on the O_2/N_2 ratio. X_{N_2} is the standard mole fraction of N_2 in dry air, 0.78084. All tracer variables are in units of ppm, but yield $\delta(O_2/N_2)^{mod}$ in per meg. To convert to units of ppm equivalents (ppm eq), $\delta(O_2/N_2)^{mod}$ is multiplied by X_{O_2} (see equation (6)).

2.6. Determination of Atmospheric Vertical Gradients

We define the atmospheric vertical gradient (Δ) of $\delta(O_2/N_2)$ and CO_2 simply as

$$\Delta O_2 = \delta(O_2/N_2)_B - \delta(O_2/N_2)_T \tag{4}$$

$$\Delta CO_2 = (CO_2)_B - (CO_2)_T \quad (5)$$

where the subscripts B and T denote the average value of each species in the atmospheric BL near the ocean surface and at a reference height in the free troposphere, respectively. By the sign convention of Δ, a positive value indicates a higher value in the BL.

To compute gradient ratios (GR) in molar units, and thus make it comparable to F_{O_2}/F_{CO_2}, we defined the GR of $\delta(O_2/N_2)$ and CO_2 as

$$GR = \frac{\Delta O_2 \cdot X_{O_2}}{\Delta CO_2} \quad (6)$$

where ΔO_2 is in per meg and ΔCO_2 is in ppm. X_{O_2} is the mole fraction of oxygen in the atmosphere, 0.2095 as defined by the Scripps O_2 Scale, and serves to convert from per meg to ppm equivalents (ppm eq) (Keeling et al., 1998), yielding a GR in mol/mol. To compute the "average GR," we in all instances take the ratio of average Δ values, rather than average multiple GRs themselves.

Vertical gradients across the top of the BL were determined by taking the average mole fraction of CO_2 or CO_2^{mod} and the average $\delta(O_2/N_2)$ or $\delta(O_2/N_2)^{mod}$ value over 1.37–1.98 km, producing $\delta(O_2/N_2)_T$ in equation (4) and $(CO_2)_T$ in equation (5), and subtracting that value from the average between the surface and 0.30 km above sea level (producing $\delta(O_2/N_2)_B$ in equation (4) and $(CO_2)_B$ in equation (5)). All GRs are given in units of ppm eq/ppm, that is, mol/mol.

All altitudes were taken from the aircraft GPS. A typical minimum altitude for a profile was 0.15 km. Ascents after descents were averaged and considered to be a single profile. We tested potential reference heights of 1–3, 1.37–1.98, 4.5–6.5, and 5.94–6.24 km. Additionally, we considered a dynamic reference height of 0.10–2.1 km above the BL as estimated from observed potential temperature (Θ) or the modeled BL height from CESM. For a discussion of the merits of different reference heights, and our selection of 1.37–1.98 km, see section 3.3. Note that all reported ΔCO_2, ΔO_2, and GRs use 0.15–0.30 km for the average BL value and 1.37–1.98 km as the average reference value, unless otherwise stated.

We excluded those profiles from all analyses with an absolute ΔCO_2 less than 0.1 ppm, except where explicitly noted. We chose this basis for filtering profiles because ΔCO_2 values close to 0 typically produce GRs that are very negative or very positive. This filtering step resulted in five profiles being excluded from the East Drake Passage spatial domain, two from the Chilean Margin, and two from the Patagonian Shelf (for definitions of these domains, see section 3.2).

3. Results

3.1. Example of Observed Gradient and Gradient Ratio

An example profile is presented in Figure 2. During a western survey (Research Flight 3) over the Drake Passage and Bellingshausen Sea on 21 January 2016, the aircraft conducted a profile from 6.1 to 0.28 km at 60.1°S, 90.2°W. We determined the height of the BL to be approximately 0.63 km, as this was the altitude below which observed potential temperature (Θ) was nearly constant. $\delta(O_2/N_2)$ and CO_2 were anticorrelated throughout the maneuver as $\delta(O_2/N_2)$ increased and CO_2 decreased toward the surface. ΔCO_2 (± propagated standard error) was −0.36 ± 0.006 ppm and ΔO_2 was 8.4 ± 0.4 per meg. The GR was −4.9 ± 0.2 mol O_2 per mol of CO_2.

A 7-day footprint for a receptor point at the bottom of this profile (Figure 2e), as determined through the STILT model, indicated that the measurements were influenced by cyclonic circulation and that the dominant upwind surface influence regions were the southern South Pacific Ocean, extending westward almost to the coast of New Zealand, and the Amundsen Sea. Sensitivity to surface fluxes remained relatively high backward in time from the receptor location, due to a lack of transport of particles out of the BL (data not shown), meaning that the BL gradients were influenced by flux processes over a fairly large region.

3.2. Observed Vertical Gradients of CO_2 and $\delta(O_2/N_2)$

Profiles showed that CO_2 typically decreased toward the surface; 94% of all ΔCO_2 values over the ocean south of 35°S were negative. The campaign average ΔCO_2 for all profiles south of 35°S (plus or minus standard error) was −0.51 ± 0.06 ppm. Based on the work of Munro et al. (2015), which showed that in the Drake

Table 1
Average Observed ΔCO_2, ΔO_2, and GR by Domain (± Standard Error), With Standard Deviations

Location	n[a]	ΔCO_2 (ppm)	ΔO_2 (per meg)	GR (mol/mol)
All points south of 35°S	61	−0.51 ± 0.06	9.8 ± 1.5	−4.0 ± 0.8
1σ		0.47	11.8	6.2
East Drake Passage	11	−0.070 ± 0.1	−3.3 ± 2.2	9.9 ± 17.5
1σ		0.38	7.31	58.2
West Drake Passage	8	−0.44 ± 0.1	10.2 ± 3.9	−4.9 ± 2.2
1σ		0.30	11.1	6.2
Bellingshausen Sea	14	−0.34 ± 0.06	10.0 ± 3.1	−6.2 ± 2.2
1σ		0.22	11.6	8.1
Chilean Margin	10	−0.83 ± 0.2	14.1 ± 2.1	−3.6 ± 0.9
1σ		0.6	6.8	3.1
Patagonian Shelf	12	−0.93 ± 0.1	21.4 ± 2.2	−4.8 ± 0.7
1σ		0.34	7.8	2.5

Number of profiles included in average. Note that profiles are only included if they have data for both CO_2 and $\delta(O_2/N_2)$.

Passage the ocean uptake of CO_2 was higher south of the Antarctic Polar Front than north of it, we hypothesized that ΔCO_2 would exhibit a trend with latitude, but no relationship was observed between either ΔCO_2 or ΔO_2 and latitude south of 56°S, using a two-sided t test of the linear model (for ΔCO_2: $t = 0.194$, $p = 0.85$, DF = 39; for ΔO_2: $t = -0.566$, $p = 0.58$, DF = 35). The lack of a latitudinal gradient was further confirmed after adjusting the latitude of each profile by weighting it for surface influence from STILT footprints, integrated over both 1 and 7 days. We suggest that the rapid meridional mixing by synoptic systems in this region, such as those evident in Figure 2e, effectively obscured any systematic within-region latitudinal variations in flux ratios from being detectable in the atmosphere.

We separated the ORCAS data into five spatial domains, to test the simple hypothesis that areas of the Southern Ocean that had higher/lower fluxes predicted by the CESM simulation would, on average, have higher/lower ΔCO_2 or ΔO_2, and potentially different GRs. Average observed ΔCO_2 and ΔO_2 for these

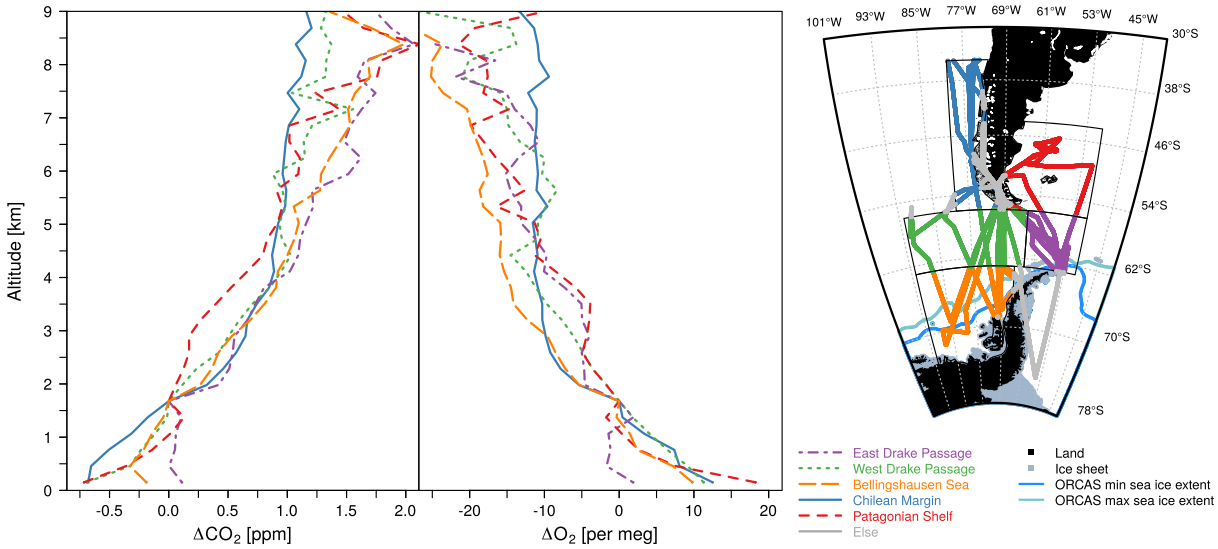

Figure 3. Average ΔO_2 and ΔCO_2 profiles for different spatial domains. Each profile shows the change in CO_2 or $\delta(O_2/N_2)$ relative to the mean value between 1.4 and 2.0 km. Profiles are separated and colored by five domains: a region on the western coast of Chile (blue), the Patagonian shelf (red), the eastern Drake Passage (purple), the western Drake Passage (green), and the Bellingshausen Sea (orange). Data from flight tracks that were not used are shown in gray. The boundaries for each domain are shown as black lines. The spatial domains show differences in the degree of $\delta(O_2/N_2)$ enhancement and/or CO_2 depletion between 2 km and the surface. A positive value indicates an enhancement in the boundary layer relative to the overlying air.

Table 2
p Values From a Pairwise Mann-Whitney U Test of ΔCO_2 by Domain

	West Drake Passage	Bellingshausen Sea	Chilean Margin	Patagonian Shelf
East Drake Passage	0.055	0.079	0.0069*	0.00022*
West Drake Passage	—	0.29	0.18	0.018*
Bellingshausen Sea	—	—	0.023*	0.00012*
Chilean Margin	—	—	—	0.31

Statistically significant value, using $p < 0.05$ as a threshold.

regions are presented in Table 1 and Figure 3. The location and magnitude of individual observed ΔCO_2, ΔO_2, and GRs are shown in Figure 4.

Comparisons between populations of ΔCO_2 from different spatial domains are shown in Figure 5. The East Drake Passage was distinct from the other domains, in that it was the only region where positive values (higher near the surface) of ΔCO_2 were regularly detected, although the average ΔCO_2 was still slightly negative. This observation, combined with a pairwise, nonpaired Mann-Whitney *U* test between all domains indicated that they fell into three groups: the (1) East Drake Passage, (2) West Drake Passage and Bellingshausen Sea, and (3) Chilean Margin and Patagonian Shelf (Table 2). The largest and most negative gradients were seen in the two regions closest to South America (the Chilean Margin and Patagonian Shelf), possibly a result of increased nutrient supply and productivity in these regions. While the East Drake Passage, West Drake Passage, and Bellingshausen Sea are fairly remote from any terrestrial fluxes or fossil fuel burning, the Patagonian Shelf (and to a lesser extent the Chilean Margin) region often include these complicating influences due to transport from South America (Stephens et al., 2018). During ORCAS, however, terrestrial signals flowing eastward off the continent were likely lofted above the marine BL, as the BL was generally cooler than the warmer continental air, a supposition which is supported by the sign of ΔCO_2 (see section 4.1), and the low levels of trace gases produced by terrestrial or industrial processes. For instance, the maximum value of CO observed at the reference altitude for profiles over the Patagonian Shelf region was 49.3 ppb (parts per billion), and the maximum value of CO observed at the surface was 47.7 ppb, values which are indistinguishable from the background as measured at Southern Ocean flask sampling sites (Petron et al>., 2018).

A somewhat similar spatial pattern (with opposing sign) was seen for $\delta(O_2/N_2)$. The campaign average ΔO_2 south of 35°S was 9.8 ± 1.5 per meg. A pairwise, nonpaired Mann-Whitney *U* test illustrated that the differences between domains were similar to those found for ΔCO_2 (Table 3), though the differences were generally larger and more statistically significant. A relatively larger range of ΔO_2 was seen in the Bellingshausen Sea, which meant that the differences between this domain and others was smaller and less significant than that of ΔCO_2. The East Drake Passage was the only domain in which negative ΔO_2 values were routinely seen. Larger, positive ΔO_2 values were seen over the Patagonian Shelf and Chilean Margin. The Mann-Whitney *U* test indicated that these differences were unlikely to arise by chance.

The average GR for the entire campaign south of 35°S was -4.0 ± 0.8, plus or minus the propagated standard error. We found similar average GRs in all domains (note that by "average GR," we mean ratio of the average ΔO_2 and ΔCO_2). Due to the spread in the GR as a result of many small Δ values, no statistical difference was found at the $p < 0.05$ level. When ΔCO_2 values are small and centered around 0, the sign and magnitude of the GR can vary markedly, such as in the East Drake Passage (Figure 6).

Table 3
p Values From a Pairwise Mann-Whitney U Test of ΔO_2 by Domain

	West Drake Passage	Bellingshausen Sea	Chilean Margin	Patagonian Shelf
East Drake Passage	0.0040*	0.018*	0.00034*	3×10^{-5}*
West Drake Passage	—	0.92	0.18	0.0082*
Bellingshausen Sea	—	—	0.26	0.0082*
Chilean Margin	—	—	—	0.051

Statistically significant value, using $p < 0.05$ as a threshold.

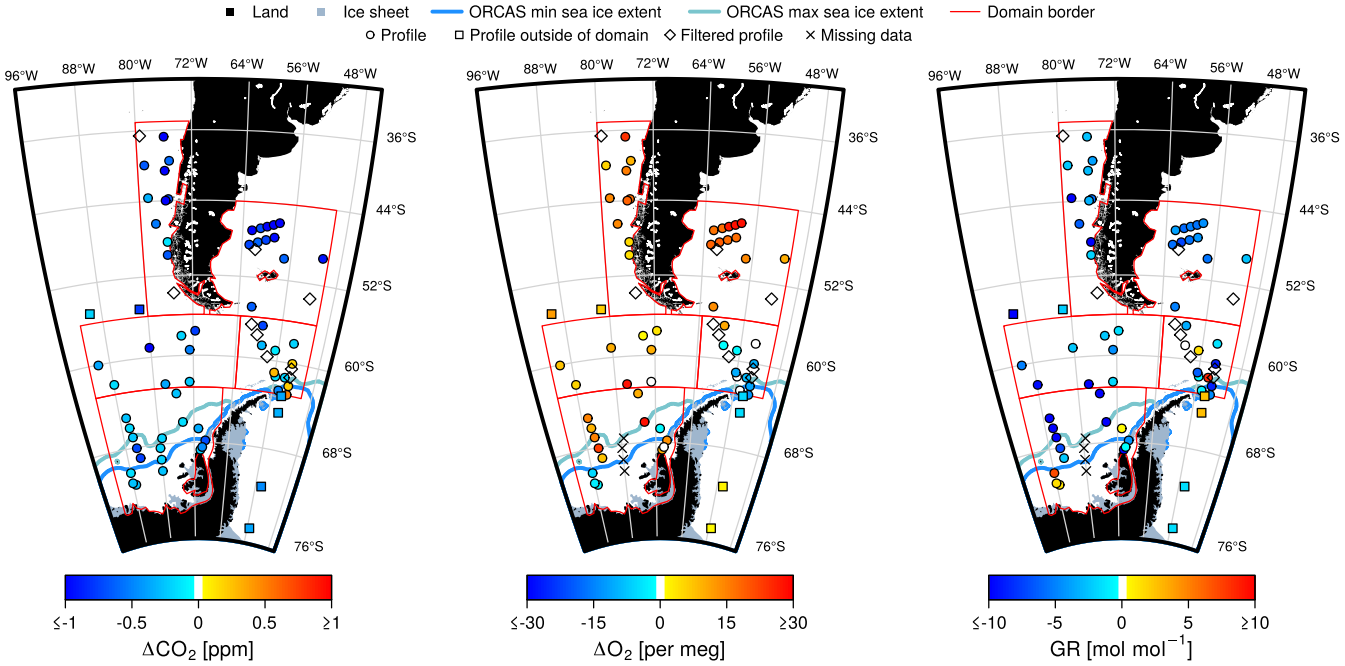

Figure 4. Locations of observed ΔO_2, ΔCO_2, and GRs during ORCAS, shaded by magnitude. Profiles included in the averages for all columns presented in Table 1 (except "south of 35°S") are depicted as circles, while those excluded from these averages are depicted as squares. Profiles which were filtered for having $|\Delta CO_2| < 0.1$ ppm are shown as empty diamonds, and profiles with no data are shown as x's.

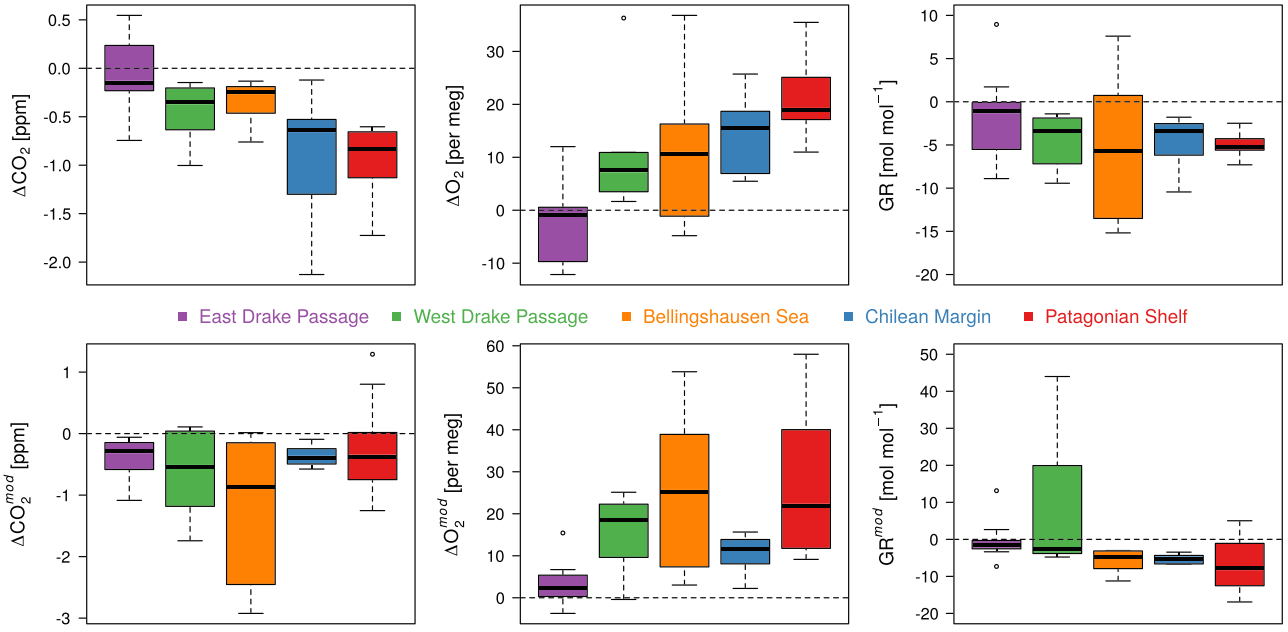

Figure 5. (top row) Boxplots of observed ΔCO_2, ΔO_2, and GRs for the different spatial domains presented in Figure 3. The box is bounded by the upper and lower quartiles, and the black horizontal line indicates the median. The whiskers show either the range (if smaller than 1.5 times the interquartile range, IQR), or the minimum and maximum data point that does not exceed 1.5 * IQR. Data beyond these values are shown as individual points. (bottom row) same as top row, except data are from the CESM simulation. Note the different *y* axes.

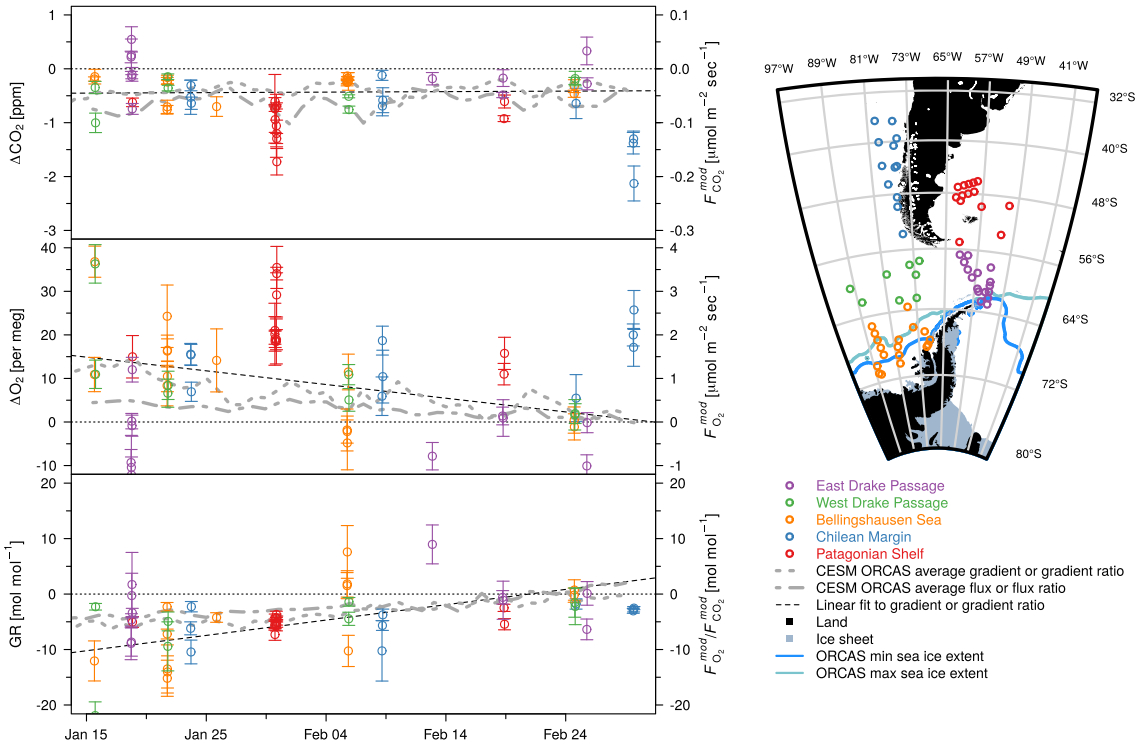

Figure 6. Time series of observed ΔCO_2, ΔO_2, and GRs during ORCAS. The map shows the location of each profile, colored by domain. The long dash gray lines in the time series plots show the average gradient and gradient ratio from the CESM simulation for the spatial domain shown in the map. The short dash gray lines show the average flux and flux ratio from the CESM simulation for the same domain. Dashed black lines are linear fits to the observations.

While there was no significant temporal trend in ΔCO_2 over the course of the campaign, a trend was discernible for the ΔO_2 and the GR, as shown in Figure 6. We conclude that this trend was caused by the seasonal cycle in air-sea fluxes of O_2, as summertime fluxes should decline from their peak sometime in January over the 6-week period. This seasonal trend was also predicted by the model. The presence of a trend was assessed using a two-sided *t*-test of the linear model after excluding profiles on the final flight day, which were more northerly and close to land (for ΔCO_2: $t = 0.24$, $p = 0.80$, DF = 60; for ΔO_2: $t = -2.8$, $p = 0.0069$, DF = 56; for GR $t = 3.6$, $p = 0.0007$, DF = 56). The differences between regions were larger than this seasonal variation.

3.3. Spatial Representativeness and Reference Altitude Selection

To assess the temporal and spatial scales integrated by the vertical gradient, Δ, we first quantified the sensitivity of Δ to the choice of the upper reference altitude, by varying the upper altitude while keeping the lower altitude range fixed. Increasing the upper reference altitude generally increased the magnitude of the gradient difference: ΔO_2 became more positive by 4.9 per meg/km on average and ΔCO_2 became more negative by −0.34 ppm/km on average. Increasing the reference altitude also reduced the variability in ΔO_2, ΔCO_2, and GR. GRs were more negative when calculated with a lower reference altitude than with a higher; for instance, the ratio of the entire campaign gradient from the upper troposphere (7 km) to 0.15 km was -2.5 ± 0.1 and the GR from the upper troposphere to 1.7 km was -2.3 ± 0.1. The final selection of 1.4–2.0 km as a reference height was made because this range was typically above the BL, but not so much higher as to mask the local and regional flux influences with influences from further north.

To explore which areas of the Southern Ocean influenced ΔO_2 and ΔCO_2, we computed average three-day footprints for each profile using the STILT model. The average footprint of a single profile is the average of all three-day footprints along the portion of the flight track which fell within the reference altitude range or within the surface altitude range. We then averaged these individual average profile footprints for each domain; these footprints are shown in Figure 7. The predominant mode of atmospheric transport as predicted by the STILT model was from the west and cyclonic. The domain averaged footprints showed surface influence from a region that stretched westward to 160°W. Using the same model output, Stephens et al.

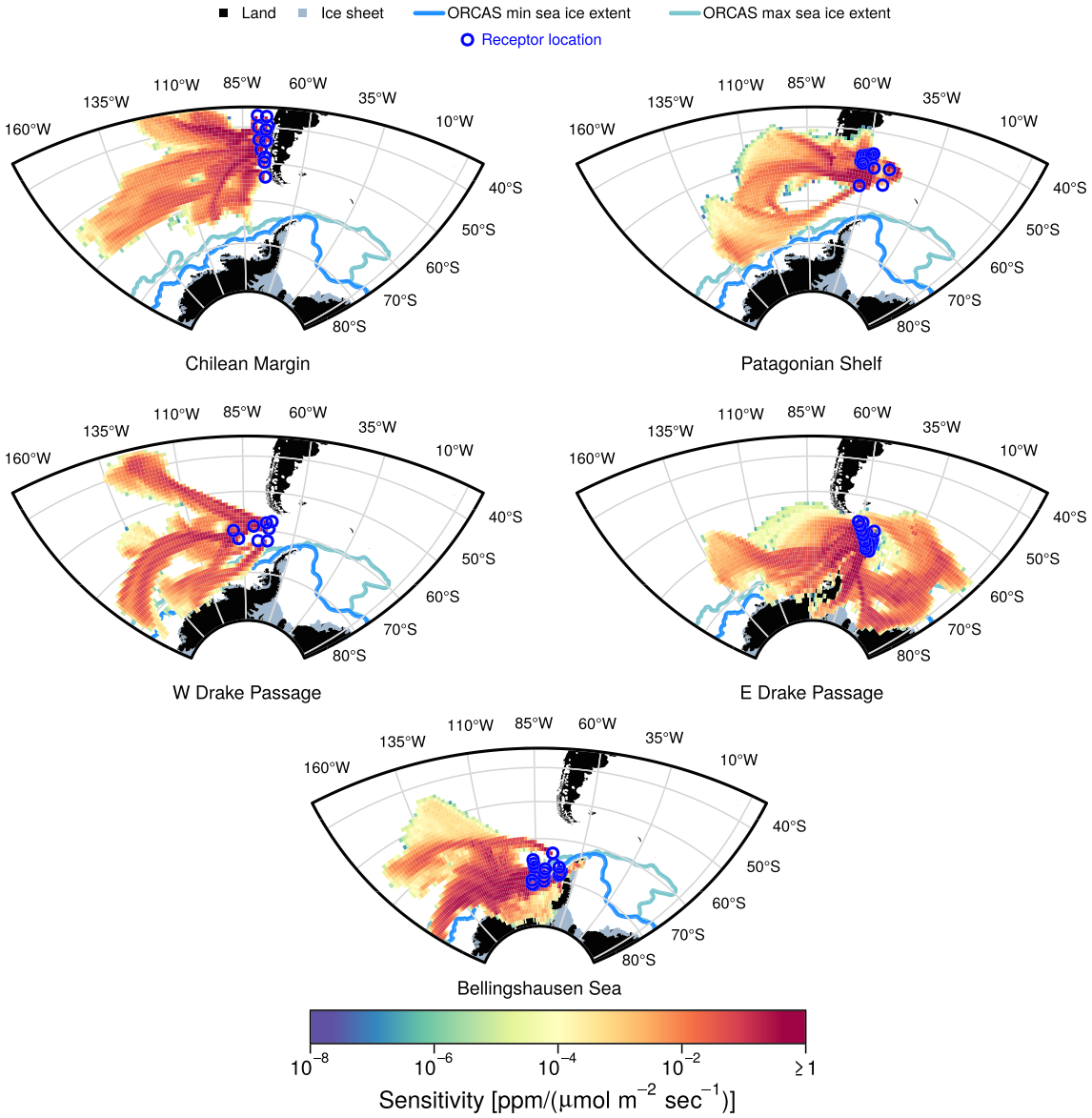

Figure 7. Average 3-day gridded STILT footprints for profiles contained within each domain. The BL receptor points for each profile are shown as blue circles. The average footprint shows the region of the surface which was expected to most influence the BL gradients and GRs.

(2018) found that 7-day footprints showed influence from most of the Southern Ocean and parts of the South Pacific, though these influences were small.

As expected, increasing the reference height decreased the influence of local surface fluxes on ΔO_2 and ΔCO_2, because the sensitivity declined quickly with increasing altitude. The average sensitivity to surface fluxes for dropped to 0 above 2.6 km. A higher reference height was also associated with airmasses that had traveled over a greater distance. For instance, the average distance traveled by a particle in 1 day in the STILT simulation for receptor points at the 1.4–2.0 km reference height was 514±213 km compared to 326±166 km for receptor points corresponding to the bottom of each profile.

3.4. The Relationship Between Surface Fluxes and Gradients Within CESM

In order to better understand the degree of spatial and temporal correspondence between BL gradients and surface fluxes, we conducted several analyses of the CESM simulation. In the model, the spatial distribution of the atmospheric gradients and the surface fluxes exhibited similar patterns (Figure 8). The spatial pattern was more similar in areas with larger CO_2 fluxes than it was in areas with weaker CO_2 fluxes.

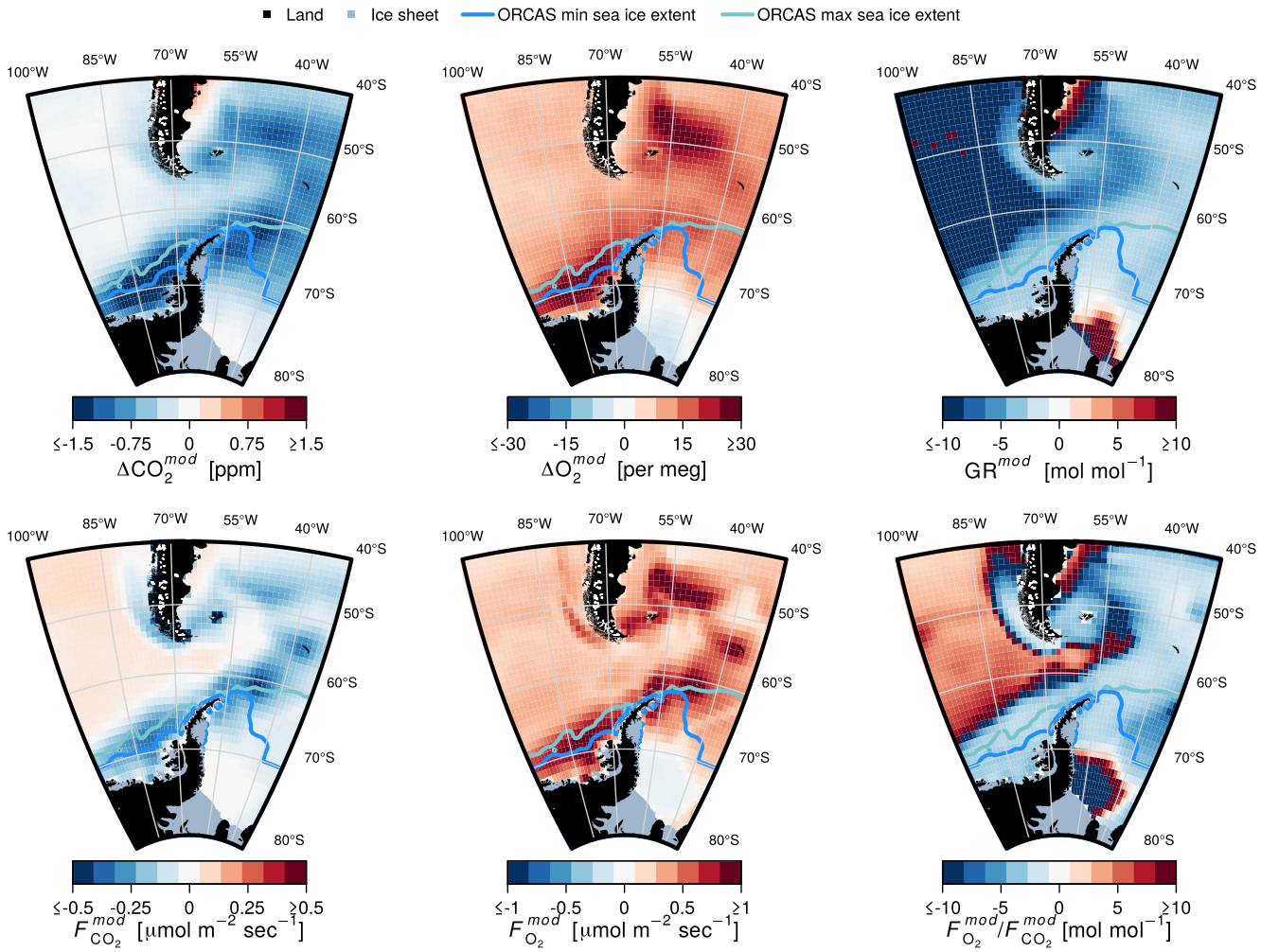

Figure 8. CESM surface fluxes and flux ratios compared to CESM gradients and GRs, averaged over the ORCAS campaign period (15 January to 29 February 2016). All data are from CESM. Gradients, fluxes, and their respective ratios show similar spatial patterns within the model.

Generally, the Bellingshausen Sea, the continental shelf surrounding southern Chile and Argentina, and the Antarctic Peninsula all showed negative $F^{mod}_{O_2}/F^{mod}_{CO_2}$, a pattern which was preserved in the GR. We interpret regions in the model with positive $F^{mod}_{O_2}/F^{mod}_{CO_2}$ as being dominated by thermal forcing, and regions with negative $F^{mod}_{O_2}/F^{mod}_{CO_2}$ as being dominated by biological forcing, if the fluxes are not close to 0. Away from these productive regions with higher fluxes, the CESM simulation predicted weaker fluxes and positive $F^{mod}_{O_2}/F^{mod}_{CO_2}$, but still a negative GR, reflecting the large scale influences on the gradients. The boundaries between regions with positive $F^{mod}_{O_2}/F^{mod}_{CO_2}$ and negative $F^{mod}_{O_2}/F^{mod}_{CO_2}$ produced highly variable $F^{mod}_{O_2}/F^{mod}_{CO_2}$ and GRs. The reason for this can be seen in the following example: If an air parcel moves first over an ocean region dominated by biological forcing, it will pick up positive O_2 and negative CO_2 anomalies, relative to the background. If it is then advected over a region with air-sea exchange dominated by thermal outgassing, ΔO_2 will continue to increase but ΔCO_2 will move back toward 0, with a resulting GR that can be a very positive or very negative number, depending on the sign of ΔCO_2.

4. Discussion

4.1. Observed and Simulated Terrestrial and Fossil Fuel Gradients

Both observation and model results suggested that the contribution of fossil fuel burning and the land biosphere to ΔO_2 and ΔCO_2 was relatively small. Observed profiles of CO, a trace species produced primarily from combustion processes, showed an average ΔCO ($\pm 1\,\sigma$) south of 35°S of -0.13 ± 2.1 ppb, implying that few airmasses influenced by these processes were encountered.

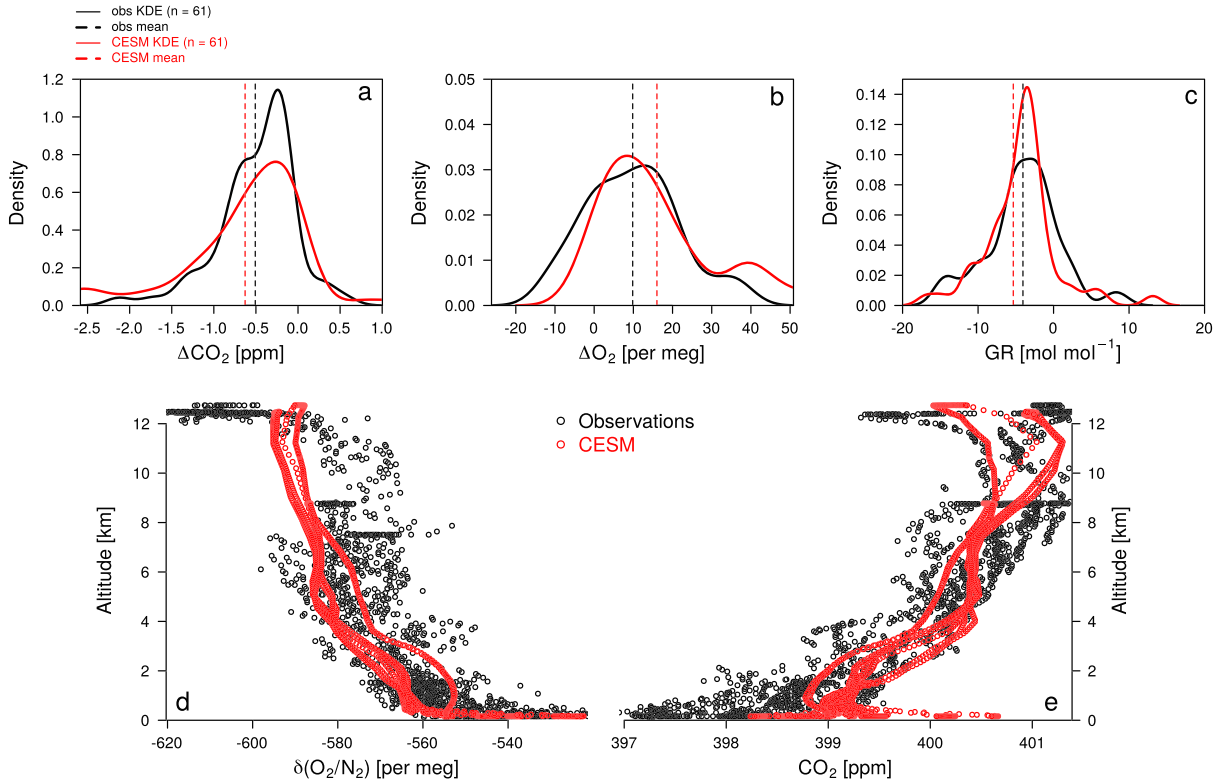

Figure 9. Kernel density estimates (KDE) for observed and modeled ΔCO_2 (a), ΔO_2 (b), and GRs (c) for all ORCAS profiles. The KDE is an estimate of the probability density function for a given variable. The mean gradients and ratio of the mean gradients are given as dashed vertical lines. Also shown are observed and modeled profiles from Research Flight 7 for $\delta(O_2/N_2)$ (d) and CO_2 (e).

In the CESM simulation, a small contribution of terrestrial processes and fossil fuel burning was predicted. Using only the tracer for ocean fluxes, CO_2^{ocn}, resulted in an average ΔCO_2^{ocn} south of 35°S of −0.58 ppm, only 0.04 ppm less negative than using the combined tracers for ocean, land, and fossil fuel contributions (ΔCO_2^{mod}). For $\delta(O_2/N_2)^{ocn}$ the average value of ΔO_2^{ocn} was 16.7, more positive than ΔO_2^{mod} by only 0.7 per meg. The resulting GR for ocean processes (GR^{ocn}) was indistinguishable from the GR^{mod} including terrestrial processes and fossil fuel burning (−5.3 ± 0.9 and −5.3 ± 1.0 mol/mol, respectively). The exception to this was over the Patagonian Shelf, where ΔCO_2^{ocn} was −0.58 ppm, more negative by 0.3 ppm than ΔCO_2^{mod}, but ΔO_2^{ocn} was less positive at −9.7 per meg than ΔO_2^{mod} by 10 per meg. This resulted in a GR^{mod} that was less negative (−20 ± 15 for all model process and −10 ± 2 for only ocean processes). This conforms to expectations, since out of the five regions considered, the Patagonian Shelf region is the most prominently downwind of South America out of the spatial domains, and hence we expect a more complex blend of processes. We do not use this to confirm the interpretation of our observations per se, especially when considering the fact that the model seems to overestimate marine fluxes (see section 4.3), but instead to illustrate how a small contribution from nonmarine fluxes might impact the GR.

4.2. Data-Model Comparisons

We do not necessarily expect that the CESM simulation would agree well with individual profiles given its resolution (see bottom panels in Figure 9), stochastic mesoscale ocean features, and inevitable marine and atmospheric transport errors. We nevertheless show that the model does a reasonable job of representing the sign and magnitude of the gradients and GRs, by averaging over temporal and spatial scales large enough to remove most synoptic variability, that is, the regional mean gradient. Kernel density estimates (KDE; a nonparametric method for estimating the probability density function) for the distribution of all observed dips are shown in Figure 9, overlain with the corresponding KDE for matching model data. For ΔCO_2, the distributions are quite similar and from this we conclude that CESM accurately predicts the mean ΔCO_2 south of 35°S. A comparison of modeled to observed ΔCO_2 along ORCAS flight tracks shows that while the mean ΔCO_2^{mod} (plus or minus standard error) in CESM was slightly more negative at −0.63 ± 0.09 ppm (compared to −0.51 ± 0.006 ppm for the observations), the difference between the two means is likely to be

due to chance (Mann-Whitney U test, nonpaired, $p = 0.69$). The model performs more poorly in simulating $\delta(O_2/N_2)$, however; it predicts a mean ΔO_2^{mod} of 16.0 ± 1.8 per meg (plus or minus standard error) compared to 9.8 ± 1.4 per meg for the observations, a finding that appears statistically significant (Mann-Whitney U test, nonpaired, $p = 0.036$).

The average GR^{mod} (matched to the observations) is more negative than that of the observations themselves (Figure 9), -5.3 ± 1.0 versus -4.0 ± 0.8 for the observations (plus or minus the propagated standard error), but the modeled and observed GRs are not statistically different when compared directly (Mann-Whitney U test, nonpaired, $p = 0.93$). That the ΔO_2 values are significantly different but the GRs are not is partially a result of the small CO_2 disagreement being in the same direction as O_2, and because the observations and the model sampled along the flight track both have noise associated with subsampling a heterogeneous domain. A more valid comparison may be to the domain mean from CESM as discussed below in section 4.4.

CESM ΔO_2^{mod} and ΔCO_2^{mod} were in poorest agreement with the observations in the East Drake Passage, where $F_{O_2}^{mod}/F_{CO_2}^{mod}$ are predominantly positive, and the West Drake Passage and Bellingshausen Sea, where $F_{O_2}^{mod}/F_{CO_2}^{mod}$ are predominantly negative (compare top and bottom panels in Figure 5). For both ΔO_2 and ΔCO_2 the CESM predictions are much larger than observed, and often of the wrong sign for ΔCO_2. The disagreement with observed ΔO_2 is larger than for ΔCO_2. Since we infer these two areas are the most strongly dominated by solubility-driven air-sea fluxes (East Drake Passage) and biologically driven air-sea fluxes (Bellingshausen Sea), this raises the possibility that the model may consistently overestimate both thermal and biological forcing, as discussed by Stephens et al. (2018). During austral summer thermally and biologically driven fluxes are typically of the same sign for O_2, and hence such errors would compound in ΔO_2^{mod}, whereas they would counteract for ΔCO_2^{mod}. Finally, we note that CESM correctly predicted a seasonal trend toward less negative GRs, driven by O_2 outgassing fluxes decreasing more rapidly than CO_2 uptake, most likely a result of the slower equilibration time of CO_2 (Figure 6).

4.3. Interpretation of Atmospheric Gradients

The observed trend toward a less negative GR with increasing altitude is consistent with an expected larger fossil fuel influence in the upper atmosphere (Stephens et al., 2018), which tends to pull the GR toward a less negative value than we would expect to observe if biological air-sea fluxes alone were driving the gradients, though this is complicated by the influence of solubility-driven fluxes at lower latitudes, the mixing of Northern Hemisphere air at higher altitudes, the transport time between the two hemispheres, and the seasonal change in the flux (Stephens et al., 2013). Based on the STILT footprints, ORCAS upper troposphere air and BL air integrated the influence of surface fluxes from different spatial and temporal scales, which could also result in a difference in GR when using higher reference heights.

Stephens et al. (2003) conducted shipboard measurements of atmospheric O_2 and CO_2 along a transect from Palmer Station to Punta Arenas during austral spring, and found variations of -4.8 mol O_2 per mol CO_2, which they attributed to upwind biologically driven air-sea fluxes of O_2 (ingassing) and CO_2 (outgassing). Although this implies a gradient ratio similar to that observed by ORCAS, the different circumstances make comparison difficult. This GR was a meridional gradient, measured over a different temporal scale, at a different time of year, and suspected to be driven by net respiration, rather than net production. We these caveats we note that it is very similar in magnitude to our mean regional GR, -4.0 ± 0.8 mol O_2 per mol CO_2, during a time when biological fluxes should dominate the BL gradient.

Since atmospheric footprints showed sensitivities heavily weighted toward the ocean, and most observed GRs were more negative than -1.1 (with excesses of O_2 and depletions of CO_2 occurring within the BL) we conclude that the dominant process which created ΔCO_2 was air-sea exchange driven by net production. The spatial and temporal variation of the ΔO_2, ΔCO_2, and GR showed a fair amount of correspondence to the spatial and temporal variability in CESM predictions of F_{O_2}, F_{CO_2}, and F_{O_2}/F_{CO_2}, respectively. Further interpretation of ΔO_2, ΔCO_2, and the GR, however, is complicated by the synoptic mixing of the atmosphere.

The model results support an important caveat, that the marine BL excesses and their ratios should not be interpreted as directly serving as a proxy for the instantaneous and co-located flux ratio, nor even the flux ratio as integrated over several days, except in cases where the air-sea fluxes are quite strong, and the BL is shallow. The simulation also showed that the BL gradient was more representative of the surface flux in the middle of regions with similar fluxes—rather than the edge—where the upwind region has a surface flux which is similar to the flux of the receptor point. In these cases, BL gradients and their GRs can

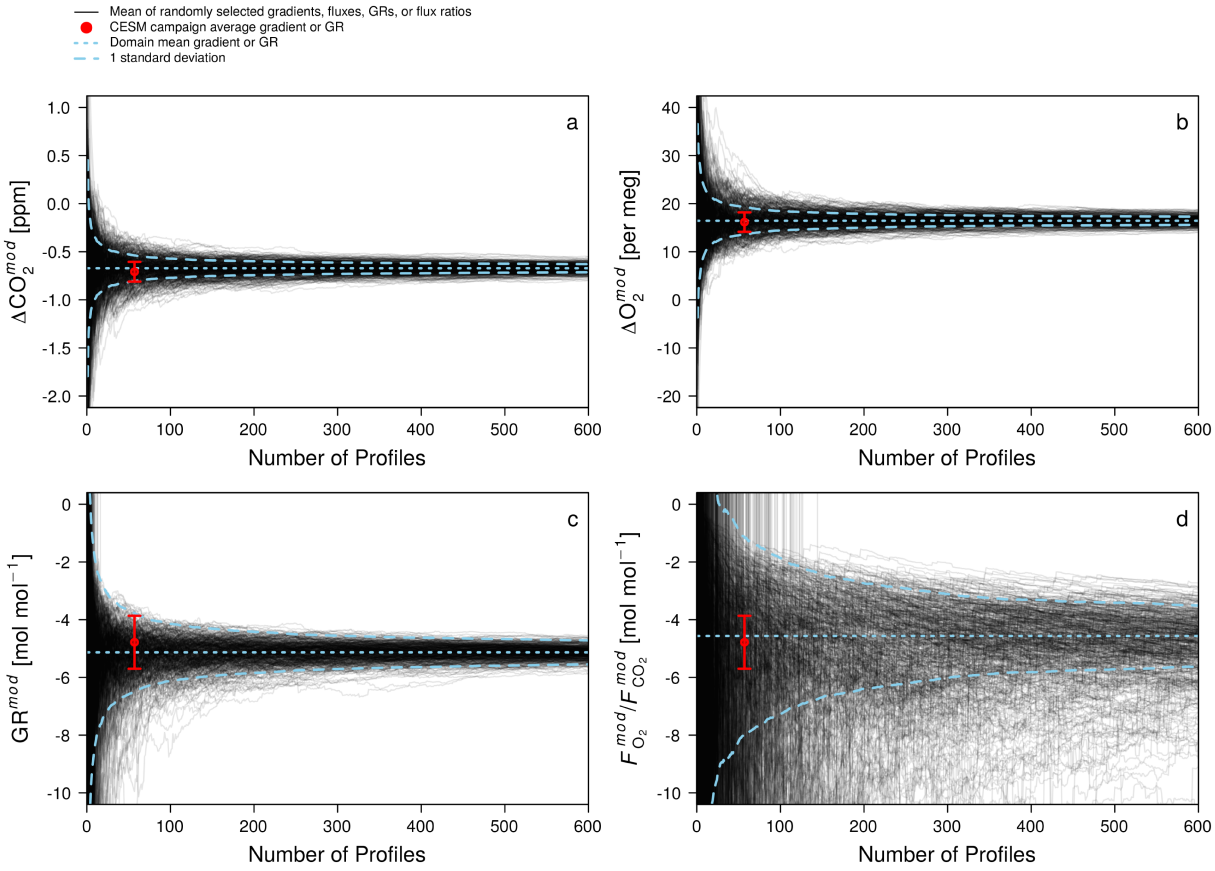

Figure 10. The cumulative mean gradient (gray lines) of (a) ΔCO_2^{mod}, (b) ΔO_2^{mod}, (c) the gradient ratio (GR^{mod}), and (d) the flux ratio, $F_{O_2}^{mod}/F_{CO_2}^{mod}$, as a function of the number of randomly sampled marine grid cells for a domain bounded by 45°S, 75°S, 100°W, and 45°W. Each random sample represents a single vertical gradient or, in the case of (d), the $F_{O_2}^{mod}/F_{CO_2}^{mod}$ of a single grid cell. The red circle on each plot shows the average gradient (a and b) or ratio (c and d) of the average gradients interpolated to observed flight tracks. Error bars on the red points show the standard error or propagated standard error (gradient ratio only). The whole domain mean is depicted by a dotted blue line, and the standard deviation or propagated standard deviation of the 500 iterations for each cumulative profile number is shown as a dashed line. All data are from the CESM simulation.

provide a reasonable approximation of the sign and relative strength of local surface fluxes. Atmospheric transport causes additional noise in the gradients by rapidly mixing anomalies from regions with different flux processes, so repeat observations are necessary to properly capture the large scale flux signal. For larger regions that encompass varied flux regimes, such as the domain covered by ORCAS, repeated profiles in the same location may be of more use in retrieving local flux information directly from the observed vertical gradient, though this would of course limit the spatial extent of a campaign. Despite the limitations on interpreting individual profiles locally, averages of many profiles provided a valuable test of regional average model fluxes.

4.4. Spatial and Temporal Coverage of ORCAS Flight Tracks

To what extent did the ORCAS flights capture the mean gradient within the area covered by the campaign flight tracks? To answer this question we performed a Monte Carlo-like simulation with the CESM atmospheric fields, the results of which are shown in Figure 10. We selected a domain bounded by 45°S, 75°S, 100°W, and 45°W. This was selected for convenience, although it has roughly the same boundaries as the average 1-day footprint of the whole campaign. 500 iterations were run in which the domain was sampled 600 times without replacement, depicted in the figure as cumulative means. The domain had an hourly time resolution, allowing for 1,110,167 potential profiles. The mean simulated ΔO_2^{mod}, ΔCO_2^{mod}, and GR^{mod} of this domain were −0.67 ± 0.001 ppm, 16.4 ± 0.02 per meg, and −5.0 ± 0.01, respectively. When these large spatial means are compared with the modeled output that exactly matches the gradients sampled in the observations, the agreement is very good: −0.71 ± 0.02 ppm, 16.2 ± 2 per meg, and −4.8 ± 0.9. The exact degree of agreement for the GR^{mod} was sensitive to the domain chosen within ± 1 mol/mol. This suggests

the ORCAS campaign had sufficiently extensive spatial coverage to capture the average ΔO_2, ΔCO_2, and GR of the region. Furthermore, the domain mean GR^{mod} and the profile-corresponding GR^{mod} both agreed well with the $F^{mod}_{O_2}/F^{mod}_{CO_2}$ of the average model fluxes over the domain, -4.6 ± 0.02 mol/mol.

The standard deviation of the domain mean ΔO_2, ΔCO_2, and GR (blue dashes in Figure 10) and the standard error on the CESM flight-track interpolated values (red whiskers in Figure 10) are comparable to the standard error we get for the observed mean ΔO_2, ΔCO_2, and GR, suggesting that the magnitude of model variability is generally representative of that we observed.

5. Summary and Conclusions

Middle- to late-summer vertical profiles of $\delta(O_2/N_2)$ and CO_2 from the ORCAS campaign over the Southern Ocean generally showed increases of O_2 and decreases of CO_2 in the BL, relative to air from a mean reference height of 1.7 km. Based on the stoichiometry of the BL enhancements/depletions, these BL gradients (ΔO_2 and ΔCO_2) were driven primarily by biological production in surface waters. In the CESM simulation, the spatial pattern of the magnitude and sign of ΔO^{mod}_2 and ΔCO^{mod}_2 values were similar to the patterns and signs of modeled air-sea fluxes, $F^{mod}_{O_2}$ and $F^{mod}_{CO_2}$. In areas with high air-sea fluxes, GR^{mod} served as a good proxy for the flux ratio, $F^{mod}_{O_2}/F^{mod}_{CO_2}$. Analysis of the model output indicated that repeat sampling of specific regions could accurately capture the 6-week mean ΔO_2, ΔCO_2, and GR, as well as the mean F_{O_2}/F_{CO_2}. While there were not enough profiles to capture the time evolution of all the distinct flux regions within the ORCAS domain over the length of the campaign, there were enough to characterize the campaign mean ΔO_2 and ΔCO_2 for a region of interest and to characterize the temporal evolution of the whole domain. The GR of the mean campaign gradients ($\pm$ propagated standard error) was -4.0 ± 0.8 mol O_2 per mol CO_2, which should approximate the F_{O_2}/F_{CO_2} of the Drake Passage region in January and February of 2016.

Both in situ measurements and the model simulation indicated that contributions to ΔO_2 and ΔCO_2 from fossil fuel combustion and terrestrial biological fluxes were small. The GR exhibited a seasonal trend during the 6-week ORCAS field campaign. This trend is consistent with CESM predictions of O_2 outgassing fluxes decreasing more rapidly than CO_2 uptake over this period of time. However, CESM closely matched observed ΔCO_2 but predicted ΔO_2 values that were too large, which resulted in GRs that were too negative, suggesting it overestimates both thermal and biological forcing.

Given the large spread in model predictions of the phasing and amplitude of the seasonal cycle of the air-sea flux of CO_2 in the Southern Ocean (Anav et al., 2013; Mongwe et al., 2018), the observed BL gradients and gradient ratios from the ORCAS campaign could be used as a constraint on the seasonality of air-sea fluxes of CO_2 and O_2 for the Drake Passage and adjacent regions of the Southern Ocean. Such an observational metric could be useful in improving model skill in simulating the Southern Ocean carbon sink.

Acknowledgments
All data are freely and publicly available from the NCAR EOL webpage (https://www.eol.ucar.edu/field_projects/orcas). In this work we have used the 10-s merge product (Stephens, 2017). This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the National Science Foundation (NSF) under Cooperative Agreement 1852977. The authors thank the pilots and crew of the NCAR/NSF GV, the NCAR project managers, field support staff, and logistics experts. ORCAS science activities were supported by NSF Polar Programs Grants 1501993, 1501997, 1501292, 1502301, and 1543457; NSF Atmospheric Chemistry Grant 1535364; NSF Atmospheric and Geospace Grant 1104642; and NASA Grant NNH14Z- DA001N-RRNES. The CESM project is supported primarily by NSF. Computing and data storage resources, including the Cheyenne supercomputer, were provided by the Computational and Information Systems Laboratory (CISL) at NCAR (Computational and Information Systems Laboratory, 2017). At Scripps Institution of Oceanography, Bill Paplawsky, Sara Afshar, Adam Cox, and Tim Lueker conducted or assisted with flask shipment, preparation, and analyses. Andy Watt at NCAR EOL provided substantial support in preparing the AO2 instrument for the field and operated AO2 and Medusa on research flights. We thank Tim Newberger for his field support of the NOAA Picarro instrument. Finally, the authors wish to thank and acknowledge the efforts and expertise of the full ORCAS science team.

References

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

Battle, M., Bender, M., Hendricks, M. B., Ho, D. T., Mika, R., Mckinley, G., et al. (2003). Measurements and models of the atmospheric Ar/N_2 ratio. *Geophysical Research Letters*, *30*(15), 1786. https://doi.org/10.1029/2003GL017411

Bender, M., Tans, P., Ellis, J., Orchardo, J., & Habfast, K. (1994). A high precision isotope ratio mass spectrometry method for measuring the O_2/N_2 ratio of air. *Geochimica et Cosmochimica Acta*, *58*(21), 4751–4758.

Bent, J. (2014), Airborne oxygen measurements over the Southern Ocean as an integrated constraint of seasonal biogeochemical processes, PhD thesis, University of California, San Diego.

Blaine, T. (2005), Continuous measurements of atmospheric Ar/N_2 as a tracer of air-sea heat flux: Models, methods, and data, PhD thesis, University of California, San Diego.

Bopp, L., & Le Quéré, C. (2009). Ocean carbon cycle. *Washington DC American Geophysical Union Geophysical Monograph Series*, *187*, 181–195. https://doi.org/10.1029/2008GM000780

Computational and Information Systems Laboratory (2017). *Cheyenne: HPE/SGI ICE XA System (Climate Simulation Laboratory)*. Boulder, CO: National Center for Atmospheric Research. https://doi.org/10.5065/D6RX99HX

Gruber, N., Landschutzer, P., & Lovenduski, N. (2019). The variable Southern Ocean carbon sink. *Annual Review of Marine Science*, *11*(1), 16.1–16.28. https://doi.org/10.1146/annurev-marine-121916-063407

Hurrell, J., Holland, M., Gent, P., Ghan, S., Kay, J., Kushner, P., et al. (2013). The Community Earth System Model: A framework for collaborative research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. https://doi.org/10.1175/BAMS-D-12-00121.1

IPCC (2013), Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Integovernmental Panel on Climate Change.

Ishidoya, S., Tsuboi, K., Matsueda, H., Murayama, S., Taguchi, S., Sawa, Y., et al. (2014). New atmospheric O_2/N_2 ratio measurements over the western North Pacific using a cargo aircraft C-130H. *SOLA*, *10*, 23–28. https://doi.org/10.2151/sola.2014-006

Ito, T., Bracco, A., Deutsch, C., Frenzel, H., Long, M., & Takano, Y. (2015). Sustained growth of the Southern Ocean carbon storage in a warming climate. *Geophysical Research Letters*, *42*, 4516–4522. https://doi.org/10.1002/2015GL064320

Kanamitsu, M. (1989). Description of the NMC Global Data Assimilation and Forecast System. *Weather and Forecasting*, *4*, 335–342.

Karion, A., Sweeney, C., Wolter, S., Newberger, T., Chen, H., Andrews, A., et al. (2013). Long-term greenhouse gas measurements from aircraft. *Atmospheric Measurement Techniques*, *6*(3), 511–526. https://doi.org/10.5194/amt-6-511-2013

Keeling, R., Blaine, T., Paplawsky, B., Katz, L., Atwood, C., & Brockwell, T. (2004). Measurement of changes in atmospheric Ar/N_2 ratio using a rapid-switching, single-capillary mass spectrometer system. *Tellus B*, *56B*(4), 322–338.

Keeling, R., & Manning, A. (2014). Studies of recent changes in atmospheric O_2 content. In *Treatise on Geochemistry*, (2nd ed., Vol. 5, pp. 385–404). Oxford, UK: Elsevier Ltd. https://doi.org/10.1016/B978-0-08-095975-7.00420-4

Keeling, R., Manning, A., McEvoy, E., & Shertz, S. (1998). Methods for measuring changes in atmospheric O_2 concentration and their application in Southern Hemisphere air. *Journal of Geophysical Research*, *103*(D3), 3381–3397.

Keeling, R., Najjar, R., Bender, M., & Tans, P. (1993). What atmospheric oxygen measurements can tell us about the global carbon cycle. *Global Biogeochemical Cycles*, *7*(1), 37–67.

Keeling, R., Piper, S., & Heimann, M. (1996). Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration. *Nature*, *381*, 218–221.

Keeling, R., & Shertz, S. (1992). Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature*, *358*, 723–727.

Khatiwala, S., Primeau, F., & Hall, T. (2009). Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean. *Nature*, *462*(7271), 346–349.

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D., Van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*(6253), 1221–1224.

Langenfelds, R. L. (2002), Studies of the global carbon cycle using atmospheric oxygen and associated tracers, PhD thesis, University of Tasmania.

Lauderdale, J. M., Dutkiewicz, S., Williams, R. G., & Follows, M. J. (2016). Quantifying the drivers of ocean-atmosphere CO_2 fluxes. *Global Biogeochemical Cycles*, *30*, 983–999. https://doi.org/10.1002/2016GB005400

Lenton, A., Tilbrook, B., Law, R. M., Bakker, D., Doney, S. C., Gruber, N., et al. (2013). Sea–air CO_2 fluxes in the Southern Ocean for the period 1990–2009. *Biogeosciences*, *10*(6), 4037–4054. https://doi.org/10.5194/bg-10-4037-2013

Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., & Grainger, C. (2003). A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model. *Journal of Geophysical Research*, *108*(D16), 4493. https://doi.org/10.1029/2002JD003161

Mongwe, N., Vichi, M., & Monteiro, P. (2018). The seasonal cycle of pCO_2 and CO_2 fluxes in the Southern Ocean: Diagnosing anomalies in CMIP5 Earth system models. *Biogeosciences*, *15*, 2851–2872. https://doi.org/10.5194/bg-15-2851-2018

Monteiro, P., Gregor, L., Lévy, M., Maenner, S., Sabine, C., & Swart, S. (2015). Intraseasonal variability linked to sampling alias in air-sea CO_2 fluxes in the Southern Ocean. *Geophysical Research Letters*, *42*, 8507–8514. https://doi.org/10.1002/2015GL066009

Munro, D., Lovenduski, N., Takahashi, T., Stephens, 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 (2002–2015). *Geophysical Research Letters*, *42*, 7623–7630. https://doi.org/10.1002/2015GL065194

Petron, G., A. Crotwell, P. Lang, and E. Dlugokencky (2018), Atmospheric carbon monoxide dry air mole fractions from the NOAA ESRL Carbon Cycle Cooperative Global Air Sampling Network, 1988-2017, Version: 2018-10-17, ftp://aftp.cmdl.noaa.gov/data/trace_gases/co/flask/surface/, accessed: 2019-07-11.

Santoni, G., Daube, B., Kort, E., Jiménez, R., Park, S., Pittman, J., et al. (2014). Evaluation of the airborne quantum cascade laser spectrometer (QCLS) measurements of the carbon and greenhouse gas suite—CO_2, CH_4, N_2O, and CO—during the CalNex and HIPPO campaigns. *Atmospheric Measurement Techniques*, *7*(6), 1509–1526. https://doi.org/10.5194/amt-7-1509-2014

Steinbach, J. (2010), Enhancing the usability of atmospheric oxygen measurements through emission source characterization and airborne measurements, PhD thesis, Friedrich-Schiller-Universität.

Stephens, B. (2017), ORCAS merge products. Version 1.0. UCAR/NCAR Earth Observing Laboratory.

Stephens, B., Brailsford, G., Gomez, A., Riedel, K., Mikaloff-Fletcher, S., Nichol, S., & Manning, M. (2013). Analysis of a 39-year continuous CO_2 record from Baring Head, New Zealand. *Biogeosciences*, *10*, 2683–2697. https://doi.org/10.5194/bg-10-2683-2013

Stephens, B., Keeling, R., Heimann, M., Six, K., Murnane, R., & Caldeira, K. (1998). Testing global ocean carbon cycle models using measurements of atmospheric O_2 and CO_2 concentration. *Global Biogeochemical Cycles*, *12*(2), 213–230.

Stephens, B., Keeling, R., & Paplawsky, W. (2003). Shipboard measurements of atmospheric oxygen using a vacuum-ultraviolet absorption technique. *Tellus*, *55B*, 857–878.

Stephens, B., Long, M., Keeling, R., Kort, E., Sweeney, C., Apel, E., et al. (2018). The O_2/N_2 Ratio and CO_2 Airborne Southern Ocean (ORCAS) study. *Bulletin of the American Meteorological Society*, *99*(2), 381–402.

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, *56*(8), 554–577. https://doi.org/10.1016/j.dsr2.2008.12.009

UCAR/NCAR - Earth Observing Laboratory (2005). NSF/NCAR GV HIAPER Aircraft. UCAR/NCAR - Earth Observing Laboratory. https://doi.org/10.5065/D6DR2SJP Retrieved December 14, 2016