
Judith Hauck1, Luke Gregor2, Cara Nissen3, Lavinia Patara4, Mark Hague1, Precious Mongwe5, Seth Bushinsky6, Scott C. Doney7, Nicolas Gruber2, Corinne Le Quéré8, Manfredi Manizza9, Matthew Mazloff9, Pedro M. S. Monteiro9,10, and Jens Terhaar11,12,13

1 Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany, 2 Environmental Physics, Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Zürich, Switzerland, 3 Department of Atmospheric and Oceanic Sciences and Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, USA, 4 GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, 5 Southern Ocean Carbon-Climate Observatory, CSIR, Pretoria, South Africa, 6 University of Hawai‘i Mānoa, Honolulu, HI, USA, 7 Department of Environmental Sciences, University of Virginia, Charlottesville, VA, USA, 8 School of Environmental Sciences, University of East Anglia, Norwich, England, 9 Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA, 10 School for Climate Studies, Stellenbosch University, Stellenbosch, South Africa, 11 Climate and Environmental Physics, Physics Institute, University of Bern, Bern, Switzerland, 12 Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland, 13 Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Falmouth, MA, USA

Abstract We assess the Southern Ocean CO₂ uptake (1985–2018) using data sets gathered in the REgional Carbon Cycle Assessment and Processes Project Phase 2. The Southern Ocean acted as a sink for CO₂, with close agreement between simulation results from global ocean biogeochemistry models (GOBMs, 0.75 ± 0.28 Pg C yr⁻¹) and pCO₂-observation-based products (0.73 ± 0.07 Pg C yr⁻¹). This sink is only half that reported by RECCAP1 for the same region and timeframe. The present-day net uptake is to first order a response to rising atmospheric CO₂, driving large amounts of anthropogenic CO₂ (Cₐ) into the ocean, thereby overcompensating the loss of natural CO₂ to the atmosphere. An apparent knowledge gap is the increase of the sink since 2000, with pCO₂-products suggesting a growth that is more than twice as strong and uncertain as that of GOBMs (0.26 ± 0.06 and 0.11 ± 0.03 Pg C yr⁻¹ decade⁻¹, respectively). This is despite nearly identical pCO₂ trends in GOBMs and pCO₂-products when both products are compared only at the locations where pCO₂ was measured. Seasonal analyses revealed agreement in driving processes in winter with uncertainty in the magnitude of outgassing, whereas discrepancies are more fundamental in summer, when GOBMs exhibit difficulties in simulating the effects of the non-thermal processes of biology and mixing/circulation. Ocean interior accumulation of Cₐ points to an underestimate of Cₒ uptake and storage in GOBMs. Future work needs to link surface fluxes and interior ocean transport, build long overdue systematic observation networks and push toward better process understanding of drivers of the carbon cycle.

Plain Language Summary The ocean takes up CO₂ from the atmosphere and thus slows climate change. The Southern Ocean has long known to be an important region for ocean CO₂ uptake. Here, we bring together all available data sets that estimate the Southern Ocean CO₂ uptake, from models that simulate ocean circulation and physical and biological processes that affect the ocean carbon cycle, from surface ocean observation-based estimates, from atmospheric transport models that ingest atmospheric CO₂ observations, and from interior ocean biogeochemical observations. With these data sets, we find good agreement on the mean Southern Ocean CO₂ uptake 1985–2018, which is 50% smaller than previous estimates when recalculated for the time period and spatial extent used in the previous estimate. However, the estimates of the temporal change of the Southern Ocean CO₂ uptake differ by a factor of two and thus are not in agreement. We further highlight that knowledge gaps exist not only in winter when observations are typically rare, but equally in summer when biology plays a larger role, which is typically represented too simplistically in the dynamic models.

1. Introduction

The Southern Ocean (Figure 1) is the primary conduit between the surface and the deep ocean (Morrison et al., 2022; Talley, 2013) making it a key region for the global carbon cycle and the climate system across time-scales from paleo to present day and into the future (Canadell et al., 2021). First, water mass formation of
Antarctic surface water occurs during large-scale upwelling of deep, old and carbon-rich water masses due to strong westerly winds (Marshall & Speer, 2012; Russell et al., 2006). Part of this water moves northwards by Ekman transport and contributes to the formation of Southern Water and intermediate waters (Ito et al., 2010; Morrison et al., 2022; Sallée et al., 2012) together with subtropical water masses (Jindoue et al., 2016). Another part moves southward and circulates in the large gyres of the Weddell and Ross Seas (Klett et al., 2005). A fraction of these Antarctic surface waters densify on the Antarctic shelves through cooling and brine rejection during sea-ice formation on the Antarctic shelves to then flow down the Antarctic slope and form Antarctic Bottom Water (Jacobs, 2004; Orsi et al., 1999).

In pre-industrial times, the Southern Ocean was a net source of CO$_2$ to the atmosphere due to upwelling of carbon-rich deep waters (Mikaloff Fletcher et al., 2007). Importantly, the large-scale upwelling that drove the natural outgassing fluxes in the polar and subpolar Southern Ocean still occurs today. However, since industrialization, increasing atmospheric levels of CO$_2$ have shifted the equilibrium of CO$_2$ partial pressure between the ocean and the atmosphere in the favor of the latter, thus overcompensating the natural outgassing (e.g., Hoppema, 2004). The contemporary net flux in the Southern Ocean can thus be understood as the sum of the outgassing of natural CO$_2$ and uptake of anthropogenic CO$_2$ (Gruber et al., 2009; Gruber, Landschützer, & Lovenduski, 2019). Importantly, the Southern Ocean has acted as the primary region of uptake for anthropogenic CO$_2$ (Caldeira & DuFry, 2000; Frölicher et al., 2015; Khatiwala et al., 2009; Mikaloff Fletcher et al., 2006; Orr et al., 2001; Sarmiento et al., 1992), which is attributed to upwelling of old water masses (with low anthropogenic carbon) in a region of high wind speeds, as well as subsequent transport of excess carbon from the surface into the ocean interior through the formation of Subantarctic Mode and Antarctic Intermediate Waters (Bopp et al., 2015; Langlais et al., 2017; Mikaloff Fletcher et al., 2006; Sallée et al., 2012; Waugh et al., 2006). In the absence of evidence of substantial changes in the biological carbon pump over the past decades, the role of biology for anthropogenic carbon uptake is thought to be small (Holzer & DeVries, 2002; Murnane et al., 1999). However, the biological carbon pump can have a strong imprint on the net fluxes during the summer when primary production draws down natural CO$_2$ (De Baar et al., 1995; Ryan-Keogh et al., 2023; Smetsacek et al., 2012) and possibly other trace metals (Browning et al., 2021).

While the general importance of the Southern Ocean for the ocean carbon sink is recognized, it is also the region with the largest uncertainty in the mean and trend of the sink (Friedlingstein et al., 2022; Hauck et al., 2020). This is partly because the observation-based estimates and model-based estimates measure different components of the ocean carbon sink, and assumptions on fluxes associated with river discharge need to made, which carry high uncertainty themselves (Aumont et al., 2001; Lacroix et al., 2020). Further, the decadal variability of the Southern Ocean and the underlying mechanisms thereof are a key contributor to the uncertainty and are a topic of continued discussion (Canadell et al., 2021; Gruber, Landschützer, & Lovenduski, 2019; Hauck et al., 2020; Landschützer et al., 2015; Le Quéré et al., 2007; McKinley et al., 2020). A stagnation in the growth of the Southern Ocean carbon sink in the 1990s is commonly attributed to a strengthening of the westerly winds and associated intensified upwelling of carbon- and nutrient-rich deep water (Hauck et al., 2013; Le Quéré et al., 2007; Lovenduski et al., 2007). Indeed, evidence for this stronger upwelling is indirectly observed by enhanced surface nutrient concentrations in all Southern Ocean basins (Ayers & Strutton, 2013; Hoppema et al., 1995, 1999; E. Jones et al., 2012, 2015) despite it being generally considered inefficient due to widespread limitation by iron (De Baar et al., 1995; Ryan-Keogh et al., 2023; Smetsacek et al., 2012) and possibly other trace metals (Browning et al., 2021).

The Southern Ocean carbon sink is projected to continue to play an important role in the future carbon cycle as shown by Earth System Model simulations (Canadell et al., 2021; Hauck et al., 2015; Kessler & Tijputra, 2016; Terhaar et al., 2021). However, there are indications that system changes may occur, such as a shift to a larger proportion of the CO$_2$ uptake occurring in the polar Southern Ocean (Hauck et al., 2015), and a strong sensitivity
Figure 1. Study region. The Southern Ocean covers three biomes: The subtropical seasonally stratified (STSS), the subpolar seasonally stratified (SPSS), and the ice (ICE) biome. The biomes are defined following Fay and McKinley (2014). We further consider the Atlantic, Pacific, and Indian Ocean sectors separately in parts of the analysis. The dashed lines show the REgional Carbon Cycle Assessment and Processes Project Phase 2 Southern Ocean northernmost extent (35°S), the RECCAP1 Southern Ocean northernmost extent (44°S), and RECCAP1’s boundary for the circumpolar region (58°S).

**Global Biogeochemical Cycles**

10.1029/2023GB007848

In this study, we aim to synthesize and assess information on the Southern Ocean carbon sink over the period 1985 to 2018 in the framework of the REgional Carbon Cycle Assessment and Processes project, Phase 2 (RECCAP2). This work builds on a previous assessment, RECCAP Phase 1 (referred to as RECCAP1 for clarity), for the period 1990 to 2009 (Lenton et al., 2013). In RECCAP1, the Southern Ocean was defined as the ocean south of 44°S (building on earlier classification in the atmospheric inversion community), which, however, cut through the major anthropogenic CO2 uptake region at the northern edge of the Southern Ocean. The assessment was based on five global ocean biogeochemical models, 11 atmospheric inversions, 10 ocean inversions and a single pCO2 observation-based data set, the climatology of Takahashi et al. (2009). RECCAP1 resulted in a best estimate of the net Southern Ocean CO2 uptake (1990–2009) of 0.42 ± 0.07 PgC yr⁻¹ based on all models (including inversions), with a surface pCO2-based climatology (Takahashi et al., 2009) suggesting a lower number of 0.27 ± 0.13 PgC yr⁻¹ Lenton et al. (2013). The interannual variability was estimated to be ±25% around this mean value. The largest proportion of the mean flux occurred in the region 44–58°S which spans large parts of the Subantarctic Zone and of the Polar Frontal Zone with similar contributions from the Atlantic, Pacific and Indian Ocean sectors. In the Antarctic Zone (south of 58°S), individual estimates did not agree on the sign of the net CO2 flux.

A major advance since RECCAP1 is the release and ongoing updates of the Surface Ocean CO2 Atlas (SOCAT, Bakker et al., 2016), which currently provides 33.7 million quality-controlled and curated surface ocean pCO2 measurements with an accuracy of <5 μatm in the 2022 release (Bakker et al., 2022). The release of SOCAT allowed for the development of the surface ocean pCO2 observation-based products (pCO2-products) that interpolate and extrapolate sparse ship-based observations from SOCAT to global coverage. Based on these maps of surface pCO2, the air-sea CO2 flux is then calculated using gas-exchange parameterizations and input data fields such as sea surface temperature (SST) and wind fields (Wanninkhof, 2014). Since RECCAP1, a diverse set of statistical and machine-learning approaches have been developed (e.g., Chau et al., 2022; Gregor et al., 2019; Landschützer et al., 2014; Rödenbeck et al., 2014). The pCO2-products allowed for observation-based investigation of interannual and decadal variability. They confirmed the reported stagnation of the Southern Ocean carbon sink in the 1990s (Le Quéré et al., 2007), and identified the aforementioned reinvigoration in the 2000s (Landschützer et al., 2015; Ritter et al., 2017). However, these pCO2-products have made the Southern Ocean's long-standing issue of sparse observations even more evident. Observing system simulation experiments (OSSEs) have shown that these methods are prone to regional and temporal biases (Denvil-Sommer et al., 2021) and some pCO2-products may overestimate the decadal variability by 30% (Gloegé et al., 2021). In fact, a recent study showed that the SOM-FFN pCO2-product used in the reinvigoration study of Landschützer et al. (2015) overestimates the model-based decadal trend 2000–2018 by 130% in an ocean model subsampling experiment (Hauck et al., 2023). However, these OSSEs have also shown that augmenting ship-based observations with well-placed, high accuracy pCO2 observations from autonomous platforms can reduce these biases (Denvil-Sommer et al., 2021; Djuchouang et al., 2022; Gloegé et al., 2021; Hauck et al., 2023).

The gap in ship-based pCO2 observations is slowly being addressed by a second major advance, that is autonomous measurement devices. Among these are pH-equipped biogeochemical Argo floats (BGC-floats) (Johnson et al., 2017; Williams et al., 2016). With this approach, float pH measurements are combined with multi-linear regression-derived alkalinity (Carter et al., 2016, 2018, 2021; Williams et al., 2016), to estimate pCO2. Although uncertainties of the BGC-float based estimates of pCO2 are, to date, higher (theoretical uncertainty of 11 μatm, Williams et al., 2017) than for direct pCO2 measurements (2 μatm, Bakker et al., 2016), some of these indirect pCO2 estimates fill critical gaps in the sparsely sampled winter months. These novel data, either on their own (Gray et al., 2018) or as additional input for pCO2-products (Bushinsky et al., 2019), reported a strong winter outgassing of CO2 in the subpolar Southern Ocean for the years 2015 through 2017 that also led to a substantially smaller estimate of the annual Southern Ocean CO2 uptake for these years. However, these larger-than-expected
winter outgassing estimates were challenged by airborne flux estimates and direct pCO₂ measurements from a circumpolar navigation by an uncrewed sailing drone (Long et al., 2021; Sutton et al., 2021). The sailing drone observations were in agreement with ship-based pCO₂-product estimates throughout all seasons (Sutton et al., 2021). The authors attributed the discrepancy between BGC-floats and other estimates to either a bias of the float measurement devices or interannual variability. In support of the latter argument, the BGC-Argo-based air-sea CO₂ flux in the years 2017–2019 also did not reveal the strong winter outgassing signal of the years 2015 and 2016 (Sutton et al., 2021).

Another advance since RECCAP1 is that more global ocean biogeochemical models (GOBMs) have become available with improvements in resolution and physical and biogeochemical process representation (Friedlingstein et al., 2022; Wanninkhof et al., 2013). While the ability of the GOBMs to capture interannual variability of air-sea CO₂ fluxes (FCO₂) was questioned by the larger variability of pCO₂-product estimates (Le Quéré et al., 2018), the lower interannual variability of GOBMs now falls within the range of the larger ensemble of pCO₂-products (Hauck et al., 2020; McKinley et al., 2020). For the decadal variability of FCO₂ there is a moderate agreement between GOBMs and pCO₂-products on a stagnation of the sink in the 1990s and an increase of the sink in 2002–2011 but with a larger amplitude of the multi-year/decadal variability in the pCO₂-products (Gruber et al., 2023; Hauck et al., 2020; McKinley et al., 2020). Although the GOBMs compare reasonably well to global and Southern Ocean observations of surface ocean pCO₂ (Hauck et al., 2020), their estimates of the global ocean carbon sink remain below those of interior ocean anthropogenic carbon accumulation estimates from 1994 to 2007 (Gruber, Clement, et al., 2019), atmospheric inversions, observed O₂/N₂ ratios (Friedlingstein et al., 2022; Tohjima et al., 2019), and a similar underestimation was found in Earth System Models (Terhaar et al., 2022).

The final major advance in the last decade are regional and global data-assimilating global ocean biogeochemistry models (Carroll et al., 2020; Verdy & Mazloff, 2017). These models bring together the process-based knowledge from GOBMs, but use data assimilation schemes to minimize mismatches between simulated fields, and physical and biogeochemical observations.

Despite these recent advances in observations and models, the Southern Ocean is still the region with the largest discrepancy in mean CO₂ flux (although within the uncertainty of the fluxes associated with river discharge, which are implicitly included in the observation-based estimates, but not in the models, see Sections 2.2.1 and 2.3.1) and variability, as well as largest model spread (Canadell et al., 2021; Fay & McKinley, 2021; Friedlingstein et al., 2022). In this study, we aim to quantify the Southern Ocean (following the RECCAP2 biome shown in Figure 1) surface CO₂ fluxes and interior storage of anthropogenic carbon over the period 1985–2018 from different classes of models and observations, and to identify knowledge gaps and ways forward.

This study is organized in the following way. In our methods, we describe the region (Section 2.1), the data sets that we use throughout this synthesis (Section 2.2), and how the data were processed (Section 2.3). Our results contain first the estimates of the mean fluxes 1985–2018 and their decomposition into anthropogenic and natural fluxes, and atmospheric CO₂ versus climate effects (Section 3.1). This is followed by an analysis of summer and winter fluxes and the full seasonal cycle, where we also decompose pCO₂ into seasonal thermal and non-thermal contributions (Section 3.2). In an annually resolved perspective, we revisit mean fluxes accounting for data sets with short temporal coverage (Section 3.3) before we then analyze the regionally averaged temporal trends of CO₂ flux (Section 3.4). The latter also encompasses a pCO₂ comparison with in situ pCO₂ observations, as well as atmospheric CO₂ and climate effects as drivers of the trends. In the final part of the results, the study then evaluates the GOBM simulation results with observation-based estimates of ocean interior storage of anthropogenic carbon in the Southern Ocean (Section 3.5). The discussion first summarizes the results with a comparison of the RECCAP1 and RECCAP2 results (Section 4.1). We also discuss the drivers of the seasonal cycle (Section 4.2), the interannual and decadal variability (Section 4.3), and the zonal asymmetry of the fluxes in the Southern Ocean (Section 4.4). Lastly, we discuss how our study links with and can inform observational programs (Section 4.5), before presenting a conceptual characterization of the Southern Ocean carbon cycle in the conclusions (Figure 14 in Section 5).

2. Methods

2.1. Regions

We use the RECCAP2 regions (De Vries et al., 2023) to define the Southern Ocean and its northern boundary (Figure 1). This definition of the Southern Ocean covers the subtropical seasonally stratified biome (STSS), the
Global Biogeochemical Cycles

2.2. Data Sets

Here, we introduce data sets across four different data classes that are used for the assessment of the Southern Ocean CO$_2$ fluxes and storage, namely: ocean biogeochemistry models (14), surface pCO$_2$-based data-products (11), data assimilated and ocean inverse models (3), and atmospheric inversion models (6).

2.2.1. Ocean Biogeochemistry Models

We used 13 global ocean biogeochemistry models (GOBMs) and 1 regional ocean biogeochemistry model (Table 1). These models simulate ocean circulation and biogeochemical fluxes caused by physics (advection, mixing, gas-exchange) and by biological processes. They are forced with atmospheric fields from reanalysis products, for example, by either heat and freshwater fluxes directly or by air temperature, wind speed, precipitation and humidity, which are converted to heat and freshwater fluxes using bulk formulae (see references in Table 1; Large et al., 1994). From these 14 models, 11 models are global ocean models with roughly 1° × 1°
Table 2
Overview of Data Sets Used in This Paper (Continued)

<table>
<thead>
<tr>
<th>Data set</th>
<th>Time period</th>
<th>Specific information</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>pCO₂-products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AOML-EXTRAT</td>
<td>1998–2018</td>
<td></td>
<td>Wanninkhof et al. (2023)</td>
</tr>
<tr>
<td>CMEMS-LSCE-FFNN</td>
<td>1985–2018</td>
<td></td>
<td>Chau et al. (2022)</td>
</tr>
<tr>
<td>Jena-CarboScope (Mixed layer scheme)</td>
<td>1985–2018</td>
<td></td>
<td>Rödenbeck et al. (2013, 2022)</td>
</tr>
<tr>
<td>LDEO-HPD</td>
<td>1985–2018</td>
<td></td>
<td>Glocge et al. (2022)</td>
</tr>
<tr>
<td>NIES-ML3</td>
<td>1985–2018</td>
<td></td>
<td>Zeng et al. (2022)</td>
</tr>
<tr>
<td>MPI-SOM-FFN (SOCCOM)</td>
<td>2015–2018</td>
<td></td>
<td>Bushinsky et al. (2019)</td>
</tr>
<tr>
<td>LDEO-climatology (Takahashi legacy)</td>
<td>Climatology</td>
<td></td>
<td>Takahashi et al. (2009)</td>
</tr>
<tr>
<td>Atmospheric inversions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UoE</td>
<td>2001–2020</td>
<td>CarboScope pCO₂-product</td>
<td>Feng et al. (2016)</td>
</tr>
<tr>
<td>CMS-Flux</td>
<td>2010–2020</td>
<td>Takahashi climatology</td>
<td>Liu et al. (2021)</td>
</tr>
</tbody>
</table>

**Note.** Sorted by data class, here: pCO₂-products and atmospheric inversions. The atmospheric inversions were provided only since 1990.

resolution, and two global models (FESOM_REcoM_HR and ORCA025-GEOMAR) and the regional model (ROMS-SouthernOcean-ETHZ) are available in ca. 0.25° × 0.25° resolution. Details of global model set-ups are given in DeVries et al. (2023). The ROMS-based regional Southern Ocean model has a northern boundary at 24°S.

For the ocean-models listed above, up to four different simulations were provided (see also Table S1 in Supporting Information S1 and DeVries et al., 2023). These differ in whether atmospheric CO₂ and all other atmospheric forcing variables vary on interannual time scales, are repeated for a single year, or follow a multi-year climatology. In simulation A, the historical run, both atmospheric CO₂ and all other physical forcing variables vary on interannual time scales. In simulation B, the preindustrial control run, a repeated year or climatological physical atmospheric forcing is used, and the atmospheric CO₂ levels are held constant at pre-industrial levels. In simulation C, the atmospheric CO₂ varies interannually and only the physical atmospheric forcing is climatological. In simulation D, the atmospheric CO₂ levels are held constant at pre-industrial levels, whereas the physical atmospheric forcing varies interannually. These simulations allow for the separation of the effects of the increase in atmospheric CO₂ and climate change and variability on air-sea CO₂ fluxes: the steady-state and non-steady state components of both natural and anthropogenic carbon. Here anthropogenic refers to the direct effect of increasing atmospheric CO₂ and non-steady state encompasses the effects of climate change and variability. For a detailed explanation, please see DeVries et al. (2023) and further explanation in Le Quéré et al. (2010), McNeil and Matear (2013), Hauck et al. (2020), Crisp et al. (2022), and Gruber et al. (2023). Simulation A includes all components of the carbon fluxes. In the control simulation B, only the steady-state component of natural carbon is considered. In simulation C, only the steady-state components of both natural and anthropogenic carbon are accounted for. Finally, in simulation D, only the steady state and non-steady state components of natural carbon are represented.

The majority of models do not account for the river-induced outgassing of carbon (DeVries et al., 2023; Terhaar et al., 2023), hence the air-sea CO₂ flux in simulation A corresponds to the $S_{OCEA}$ definition used in the Global
Carbon Budget (Friedlingstein et al., 2022), which differs from \( pCO_2 \)-product estimates by the river-induced term. Note that the river-induced term will be discussed in greater detail in Section 4.1. In addition, simulation A may include a model bias (mean offset) and drift (temporally changing offset). We assess the model drift of the air-sea \( CO_2 \) flux by calculating the linear trend of the integrated \( CO_2 \) flux time series for the period 1985–2018 in simulation B for each model and each biome. The time series plots and the linear trends reported in Figure 8 are drift corrected by subtracting the trend from simulation B. We note that this drift-correction only marginally impacts the reported trends in the result section, as the trends in simulation B are small compared to the mean fluxes for all models (see Text S1 and Figure S1 in Supporting Information S1). In contrast to a global bias (any deviation of the global mean \( CO_2 \) flux from 0 in simulation B, see Hauck et al. (2020)), the regional bias in the simulated flux cannot be assessed by the set of simulations as it cannot be separated from the natural steady-state air-sea \( CO_2 \) flux (Terhaar et al., 2023), which is non-zero on a regional level.

We use the full suite of models in all analyses, with two exceptions. First, we excluded the MPIOM-HAMOCC model in all seasonal analyses (Figures 4–7) because its amplitude of the seasonal cycle is a factor 3–6 larger than in the other models in the three main Southern Ocean biomes (Figure S2 in Supporting Information S1), and including this outlier would skew the ensemble mean disproportionately. The exaggerated seasonal cycle in the MPIOM-HAMOCC model was found in earlier studies and is attributed to excessive net primary production in the Southern Ocean (Mongwe et al., 2018). Second, the decomposition into natural and anthropogenic \( CO_2 \) fluxes was not possible with GOBMs that only provided simulations A and B (MOM6-Princeton and FESOM-REcoM-HR). See Section 2.3.5 for further restrictions on GOBM use and interpretation for the interior ocean anthropogenic carbon accumulation.

### 2.2.2. Surface \( pCO_2 \)-Based Data-Products

As a second data class, we use surface ocean \( pCO_2 \) observation-based data products (\( pCO_2 \)-products) (Table 2, for more details see DeVries et al. (2023)). These \( pCO_2 \)-products extrapolate or interpolate sparse ship-based measurements of \( pCO_2 \) using statistical modeling approaches. All \( pCO_2 \)-based data-products use SOCAT as the target data set. The majority of \( pCO_2 \)-products use similar gridded prediction data sets to fill the gaps, including SST, sea surface salinity (SSS), mixed-layer depth, and chlorophyll-a estimates for the open ocean. We use eight such \( pCO_2 \)-products that all cover the full time-series 1985–2018 for the ensemble mean of \( pCO_2 \)-products. AOML_EXTRAT covers a shorter period, and is thus not included in the ensemble mean 1985–2018, but is included in the ensemble mean 2015–2018. The largest methodological difference between the \( pCO_2 \)-products stems from the algorithm choice. The majority of the methods use regression approaches (a.k.a. machine learning) such as artificial neural networks (e.g., MPI-SOM-FFN) and gradient boosted decision trees (e.g., CSIR-ML6) to capture the relationship between the ship-based measurements and the predictor variables. The Jena-CarboScope product includes a mechanistic understanding of mixing, entrainment, and fluxes of \( CO_2 \) into and out of the mixed layer (Rödenbeck et al., 2014). The HPD-LDEO method adjusts global ocean biogeochemistry model estimates of \( pCO_2 \) to be closer to observed ship-based measurements and is thus an observation-based posterior correction to the GOBM estimates (Gloege et al., 2022).

Further, two additional variants of MPI-SOM-FFN and Jena-CarboScope by Bushinsky et al. (2019, ship + float estimates are used here) include additional BGC-float-derived \( pCO_2 \) for the Southern Ocean (referred to as BGC-float \( pCO_2 \)-products, 2015–2018). We also use the Watson2020 product, which is a neural network approach (based on MPI-SOM-FFN) but applies an adjustment to SOCAT \( pCO_2 \), that accounts for the difference between ship intake temperature and satellite SST (Watson et al., 2020). The BGC-float \( pCO_2 \)-products (2015–2018) and Watson2020 (1988–2018) are not included in the \( pCO_2 \)-product ensemble averages, as they are based on fundamentally different \( pCO_2 \) values. We also use a monthly climatology product (LDEO-clim) that is centered on the year 2010 (Takahashi et al., 2009). The LDEO-clim product fills the gaps using a combination of inverse distance weighted interpolation and a diffusive-advective interpolation scheme (Takahashi et al., 2009). Note that this product is only used in representations of the seasonal cycle, and not for trend analyses. All these \( pCO_2 \)-products estimate the bulk air-sea \( CO_2 \) flux with:

\[
FCO_2 = K_0 \cdot k_w \cdot (pCO_2^{atm} - pCO_2^{atm}) \cdot (1 - \text{ice})
\]

where \( K_0 \) is the solubility of \( CO_2 \) in seawater, \( k_w \) is the gas transfer velocity, \( pCO_2^{atm} \) is the oceanic estimate of \( pCO_2 \) from the \( pCO_2 \)-product, \( pCO_2^{atm} \) is the atmospheric \( pCO_2 \), and ice is the sea-ice fraction, with the majority of the open ocean having a fraction of 0. Other than \( pCO_2^{atm} \), \( k_w \) is the largest source of uncertainty in the
calculation of bulk air-sea CO₂ fluxes (Wanninkhof, 2014; Fay et al., 2021). However, most of the pCO₂-products use a quadratic formulation of kₗ as described by Wanninkhof et al. (1993) meaning that the product spread is reduced due to similar choices—details are shown in Global chapter's Table S2 in Supporting Information S1 (DeVries et al., 2023). An exception is the Watson2020 product (Watson et al., 2020) that calculates sea CO₂ fluxes using the formulation described in Woolf et al. (2016) where a cool and salty skin adjustment is applied.

2.2.3. Data-Assimilated Models

We use three data-assimilating models (Table 1). The Biogeochemical Southern Ocean State Estimate (B-SOSE, Verdy & Mazloff, 2017) is an eddy-permitting 1/6-degree data-assimilating model, which assimilates the data from Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) BGC-Argo floats as well as shipborne and other autonomous observations (i.e., GLODAP and SOCAT) over the period 2013–2018. In situ and satellite observations of the physical state are also assimilated. B-SOSE is based on the MIT general circulation model (MITgcm, Campin et al., 2011) and uses software developed by the consortium for Estimating the Circulation and Climate of the Ocean (ECCO, Stammer et al., 2002; Wunsch & Heimbach, 2013) to build on the SOSE physical model framework by adding the Nitrogen version of the Biogeochemistry with Light, Iron, Nutrients, and Gases (N-BLING; evolved from Galbraith et al. (2010)) biogeochemical model. Consistency with the data is achieved by systematically adjusting the model initial conditions and the atmospheric state through the 4D-Var assimilation methodology. This B-SOSE assimilation methodology does not break the model biogeochemical or physical budgets. The budgets are closed, which allows one to understand signal attribution, though limits the control over the solution. For this reason B-SOSE is only consistent with the data on the timescales longer than approximately 90 days; the mesoscale eddies are reproduced statistically and not deterministically. Even with this assimilation methodology some seasonal biases still exist, and B-SOSE is still a work in progress.

The ECCO-Darwin data-assimilation model (Carroll et al., 2020) is based on a global ocean and sea ice configuration (about 1/3°) of the MIT general circulation model and is available from January 1992 to December 2017. Besides being global and covering a longer duration than B-SOSE, this product also uses a different biogeochemical model and assimilation technique. The ECCO circulation estimates used in this version are coupled online with the Darwin ecosystem model (Dutkiewicz et al., 2009), which represents the planktonic ecosystem dynamics coupled with biogeochemical cycles in the ocean. The Wanninkhof (1992) parameterization of gas transfer velocity is used and pCO₂ atm is the National Oceanic and Atmospheric Administration Marine Boundary Layer Reference product (Drugokenkey et al., 2021). The biogeochemical observations used to evaluate and adjust ECCO-Darwin include (a) surface ocean fugacity (ICO₂) from the monthly gridded Surface Ocean CO₂ Atlas (SOCATv5, Bakker et al., 2016), (b) GLODAPv2 ship-based profiles of NO₃, PO₄, SiO₂, O₂, dissolved inorganic carbon (DIC), and alkalinity (Olsen et al., 2016), and (c) BGC-Argo float profiles of NO₃ and O₂ (Drucker & Riser, 2016; Riser et al., 2018). To adjust the model’s fit to the global biogeochemical observations, the Green’s function approach is used to adjust biogeochemical initial conditions and model parameters.

OCIMv2021 is an inverse model that assimilates observations of temperature, salinity, CFCs, and radiocarbon to achieve an estimate of the climatological mean ocean circulation (DeVries, 2022). This steady-state circulation model is used together with an abiotic carbon cycle model and atmospheric CO₂ forcing to simulate anthropogenic carbon uptake and its redistribution within the ocean. It uses a monthly time-step and simulates the period 1780 to 2018. No assimilation takes place during this period.

2.2.4. Atmospheric Inversions

Six atmospheric inversions are available for our analysis (Table 2). Atmospheric inversions make use of the worldwide network of atmospheric CO₂ observations. They ingest a data set of fossil fuel emissions, which are assumed to be well known, into an atmospheric transport model and then solve for the spatio-temporal distribution of land and ocean CO₂ fluxes while minimizing the mismatch with atmospheric CO₂ observations (Friedlingstein et al., 2022). Thus, the resulting land and ocean carbon fluxes are bound to the atmospheric CO₂ growth rate, but the estimated regional fluxes depend on the number of stations in the observational network. The inversions also start from prior estimates of land and ocean fluxes. For four inversion data sets that we use here, the ocean prior is taken from pCO₂-products that are used in this analysis as well (Table 2). One inversion (UoE) uses the Takahashi climatology as a prior and one (CMS-Flux) an ocean biogeochemical model. The atmospheric inversions are thus not independent from the other data classes (Friedlingstein et al., 2022, their Table A4). The atmospheric inversion data were submitted for RECCAP in the same version as in the Global Carbon Budget 2021 (Friedlingstein et al., 2022).
et al., 2022), but only since 1990. The three inversions starting later (2001 or 2010) are only included in averages reported for 2015–2018 (Figures 4 and 5), and as individual lines in the time-series figure (Figure 8).

2.3. Processing

Throughout this study, we report the air-sea CO₂ exchange as the net flux (FCO₂), which is the sum of natural, anthropogenic, and river-induced air-sea CO₂ flux (see e.g., Crisp et al., 2022; DeVries et al., 2023; Hauck et al., 2020). As the GOBMs vary widely in their choices on river carbon and nutrient input into the ocean and burial at the seafloor (see DeVries et al., 2023; Terhaar et al., 2023), an adjustment is applied to make all data classes comparable.

2.3.1. River Flux Adjustment

Globally, the majority of GOBMs produce a small imbalance of riverine carbon inflow and burial globally (<0.14 PgC yr⁻¹), which is smaller than the current best estimate of river-induced CO₂ ocean outgassing of 0.65 PgC yr⁻¹ (Regnier et al., 2022). The imbalances are due to manifold choices and illustrate the lack of a closed land-ocean carbon loop in the GOBMs. As the GOBMs do not adequately account for the river discharge and its fate within the ocean, and thus for river-derived ocean CO₂ outgassing (Terhaar et al., 2023), we account for this outgassing by using the spatial patterns of river-induced air-sea CO₂ fluxes from Lacroix et al. (2020) that are scaled to the global value of 0.65 PgC yr⁻¹ (Regnier et al., 2022). Southern Ocean outgassing from rivers amounts to 0.04 PgC yr⁻¹, that is, around 6% of the global river flux. It is distributed over the Southern Ocean biomes as follows (positive outgassing): 0.00036 PgC yr⁻¹ in the ICE biome, 0.053 PgC yr⁻¹ (SPSS biome), −0.014 (STSS biome). The estimated riverine CO₂ fluxes were added to biome-integrated fluxes in simulation A for all GOBMs, so that these are comparable to the pCO₂-products. They are not added to spatial maps of CO₂ fluxes due to large uncertainties in the regional attribution by Lacroix et al. (2020). The riverine fluxes are one (ICE) to multiple (SPSS, STSS) orders of magnitude smaller than the mean fluxes quantified in this study. The uncertainty associated with the river flux adjustment is discussed in Section 4.1.

2.3.2. Treatment of Different Area Coverage

Air-sea CO₂ fluxes in all data classes were integrated over the area available for each GOBM, pCO₂-product etc., that is, fluxes were not scaled to the same ocean area here. Relative to the ocean area in the RECCAP mask, the covered ocean areas in the GOBMs and data-assimilating models corresponds to 96.2%–100% (minimum for CCSM-WHOI) and to 95.6%–100% in the GOBMs and data-assimilating models corresponds to 96.2%–100% (minimum for JMA-MLR). These differences mainly stem from the ICE biome. We assume that the discrepancy arising from differences in covered area are smaller than the uncertainty arising from any extrapolation to the same area.

2.3.3. Trend Calculation

The reported linear trends between 1985 and 2018, 1985 and 2000 or 2001 and 2018 were obtained from a linear regression of the annual mean time series (using Python’s scikit-learn). The slope of the respective fit was multiplied by a factor of 10 to obtain decadal CO₂ flux trends. The start and end year of the trend calculation were chosen based on the RECCAP2 protocol and are supported by the reported stagnation of the Southern Ocean carbon sink in the 1990s and the reinvigoration since the early 2000s (Hauck et al., 2020; Landschützer et al., 2015; Le Quéré et al., 2007).

2.3.4. pCO₂ Decomposition

To separate temperature driven changes in pCO₂ from biological processes and mixing-driven entrainment, pCO₂ is decomposed into thermal and non-thermal components (Takahashi et al., 1993). The thermal component (pCO₂⁺) is calculated as

\[ pCO₂⁺ = pCO₂ \cdot e^{0.0423 \cdot ΔT} \]  

(2)

where pCO₂ is the annual mean of pCO₂ and ΔT difference of the monthly mean temperature from the annual mean temperature. The non-thermal contribution (pCO₂⁻) is estimated as the difference of the thermal contribution (pCO₂⁺) from the monthly averaged pCO₂. The first derivatives of these two components are subtracted from each other to create the pCO₂ seasonal driver metric, denoted as \( λpCO₂ \):

\[ λpCO₂ = \left| \frac{δpCO₂⁺}{δt} \right| - \left| \frac{δpCO₂⁻}{δt} \right| \]  

(3)
Global Biogeochemical Cycles

2.3.5. Anthropogenic Carbon Inventories

Anthropogenic CO₂ (C_{anth}) is defined as the change in ocean DIC since preindustrial times due to the direct effect of increasing CO₂ concentration in the atmosphere. It is computed as the DIC difference between experiments A and D. The accumulation of C_{anth} can be separated into a steady-state component (C_{anth}^{ss}, DIC difference between experiments C and B), that is influenced only by the increased atmospheric CO₂, and a nonsteady state component (C_{anth}^{ns}), which considers the effect of climate variability and change on C_{anth} (and which is maximally 10%–20% of C_{anth}). Text S2 and Figures S3 and S4 in Supporting Information S1). Here we focus mainly on the change in C_{anth} that has occurred over the period 1994–2007 (hereafter ΔC_{anth}), to correspond to the years covered by the eMLR(C*) observation-based estimate (Gruber, Clement, et al., 2019). The eMLR(C*) method differs fundamentally from past indirect or model-based methods used to estimate C_{anth} accumulated since pre-industrial times (DeVries, 2014; Gruber et al., 1996; Sabine et al., 2004; Waugh et al., 2006). Of these, we used the 1800–1994 cumulative C_{anth} estimate based on Sabine et al. (2004), which is characterized by an uncertainty of about 20% globally (Matsumoto & Gruber, 2005; Sabine et al., 2004). In terms of GOBMs, we used all those listed in Table 1, with the exception of FESOM-REcoM-HR and MOM6-Princeton who provided only experiments A and B. For most GOBMs, we analyze C_{anth}^{ss} to allow for a more accurate comparison with the observation-based data set (eMLR(C*)). However, for MPIOM-HAMOCC and CNRM-ESM2-1 it was only possible to compute C_{anth}^{ss} because of physical forcing inconsistencies between experiments A and D. We believe that the advantage of including all GOBMs in the analysis outweighs the disadvantages of having an incomplete definition of C_{anth} among GOBMs. It should be noted that the spin-up procedure of ROMS-SouthernOcean-ETHZ, which uses atmospheric CO₂ from 1969 to 1978 (for a 10 year spin-up of the biogeochemical component), makes it suitable only for the analysis of ΔC_{anth} between 1994 and 2007, and not of cumulative C_{anth} until 1994 nor of air-sea C_{anth} fluxes in specific years. As explained in the RECCAP2 model evaluation chapter (Terhaar et al., 2023), all GOBMs are forced with a very similar atmospheric CO₂ mixing ratio (xCO₂) over the historical period. However, the atmospheric xCO₂ in the pre-industrial control simulations across the GOBM ensemble varies between 278 and 287.4 ppm, leading to an underestimate of the C_{anth} storage for those models with a late starting date (Terhaar et al., 2023).

3. Results

3.1. Mean Air-Sea CO₂ Fluxes 1985–2018

We start with a comparison of the average air-sea CO₂ flux in the two data classes (GOBMs, pCO₂-products) that cover the full period 1985–2018. We exclude data classes with fewer products for the sake of robustness, and show the comparison between all data classes in Sections 3.2 and 3.4. The mean net Southern Ocean air-sea CO₂ flux 1985–2018 by the GOBM ensemble is −0.75 ± 0.28 and −0.73 ± 0.07 PgC yr⁻¹ (flux into the ocean) for the pCO₂-product ensemble mean (Figure 2a). While both ensemble means result in an almost identical ocean uptake of CO₂, the GOBM ensemble spread is four times larger.

All Southern Ocean regions are sinks of CO₂ based on the ensemble averages of the GOBMs and pCO₂-products (Figure 2). The subtropical seasonally stratified biome (STSS), which is a subduction area with deep winter mixed layer depth and intermediate chlorophyll concentration (Fay & McKinley, 2014), is the largest sink according to all data sets (GOBMs: −0.53 ± −0.17 PgC yr⁻¹, pCO₂-based products: −0.62 ± 0.05 PgC yr⁻¹, Figure 2a). Second is the subpolar seasonally stratified biome (SPSS) (GOBMs: −0.13 ± 0.14 PgC yr⁻¹, pCO₂-products: −0.06 ± 0.02 PgC yr⁻¹), which is characterized by upwelling of old water, rich in natural carbon but with low anthropogenic carbon content. The upwelled water is also rich in nutrients, and thus a region with important biological activity. Note that three GOBMs simulate the SPSS to be a source of CO₂ to the atmosphere. The marginal sea ice (ICE) biome is the weakest CO₂ sink (GOBMs: −0.09 ± 0.13 PgC yr⁻¹, pCO₂-products: −0.06 ± 0.02 PgC yr⁻¹).
−0.04 ± 0.02 PgC yr⁻¹ due to sea ice acting as a lid that prevents carbon outgassing in winter, and is the smallest of all three biomes covering an area of about 60% the size of STSS or SPSS (Fay & McKinley, 2014). Four individual models suggest that the ICE biome is a weak outgassing region, but no other data set supports this. Though

Figure 2. Temporal average of the Southern Ocean CO₂ net flux ($F_{CO₂}$). A positive flux denotes outgassing from ocean to atmosphere. The temporal average is calculated over the period 1985 to 2018 for the global ocean biogeochemistry models (GOBMs) and $pCO₂$-based data-products (Tables 1 and 2). (a) The green and blue bar plots show the ensemble mean of the GOBMs and $pCO₂$-based data-products, and open circles indicate the individual GOBMs and $pCO₂$-products. The ensemble standard deviation (1σ) is shown by the error bars. The river flux adjustment added to the GOBMs is small (0.04 PgC yr⁻¹), its distribution over the biomes is described in Section 2.3.1. (b) Zonal mean flux density of the different data sets. Thick green and blue lines show the ensemble means, and thin green and blue lines show the individual GOBMs and $pCO₂$-products. Approximate boundaries for biomes are marked with black points on the x-axis. (c, d) Maps of spatial distribution of net CO₂ flux for ensemble means of GOBMs, and $pCO₂$-products.
small, the differences between GOBM and $pCO_2$-product $FCO_2$ at the biome-level compensate each other leading to the near-identical estimates for the whole Southern Ocean (Figures 2a and 2b). In a zonal mean view (Figure 2b), the smallest uptake occurs between 62 and 55ºS and the largest uptake around 40ºS. However, the amplitude differs between data classes, with the $pCO_2$-products having a larger difference between minima and maxima (1.96 mol C m$^{-2}$ yr$^{-1}$), than the GOBM ensemble mean (1.19 mol C m$^{-2}$ yr$^{-1}$). Some of the individual GOBMs deviate from this pattern (see Figure S5a in Supporting Information S1 for zonal means of individual models).

Regionally, significant differences emerge between the Atlantic, Indian and Pacific sectors of the Southern Ocean (Figures 2c and 2d). Within the STSS, large CO$_2$ fluxes into the ocean occur in the Atlantic and Indian sector across all data classes (Figures 2b and 2c), mean flux density: $-1.93$ and $-2.05$ mol C m$^{-2}$ yr$^{-1}$ for GOBMs and $pCO_2$-products, respectively, in the Atlantic sector, $-1.44$ and $-1.89$ mol C m$^{-2}$ yr$^{-1}$ in the Indian sector, and $-1.22$ and $-1.54$ mol C m$^{-2}$ yr$^{-1}$ in the Pacific sector). CO$_2$ outgassing locations differ across the data classes. In the GOBM ensemble mean, the outgassing is mainly confined to the Indian sector of the SPSS, whereas it is more widely spread in the $pCO_2$-product ensemble mean covering the Pacific and Indian sectors of the SPSS and the Indian sector in the ICE biome. The smooth appearance of the outgassing signal in the GOBM and $pCO_2$-product ensemble means may be partly attributable to averaging over multiple data sets and months and years.

### 3.1.1. Decomposition Into Anthropogenic and Natural Carbon Fluxes and Climate Versus Atmospheric CO$_2$ Effects on the Mean CO$_2$ Flux

With the aid of the additional model simulations, we can decompose the net Southern Ocean air-sea CO$_2$ flux into natural and anthropogenic components, and separate the indirect effects of physical climate change and the direct geochemical effect of increasing atmospheric CO$_2$ mixing ratios. The GOBM ensemble mean indicates that the natural Southern Ocean carbon cycle without anthropogenic perturbation would be a small CO$_2$ source to the atmosphere of 0.05 PgC yr$^{-1}$, although with a large model spread as indicated by the standard deviation of 0.25 PgC yr$^{-1}$ (Figure 3 green bars). In fact, six GOBMs simulate negative natural CO$_2$ fluxes, that is, into the ocean, and six GOBMs simulate positive natural fluxes, that is, out of the ocean. This also illustrates that the GOBM spread of net fluxes (gray bars in Figure 3, standard deviation: 0.28 PgC yr$^{-1}$) is, to the first order, dominated by the model differences of natural fluxes (standard deviation: 0.25 PgC yr$^{-1}$), which may contain artifacts from model biases and drift (Terhaar et al., 2023). The spread of anthropogenic fluxes is smaller (0.13 PgC yr$^{-1}$, dark red in Figure 3). The small natural outgassing signal in the ensemble mean is a balance of natural CO$_2$ uptake in the STSS ($-0.26 \pm 0.14$ PgC yr$^{-1}$) and outgassing in the SPSS ($0.21 \pm 0.11$ PgC yr$^{-1}$) and ICE ($0.10 \pm 0.12$ PgC yr$^{-1}$) biomes. This is in qualitative agreement with the patterns of natural CO$_2$ fluxes by Mikaloff Fletcher et al. (2007) who report outgassing of natural CO$_2$ in the Southern Ocean with the largest contribution from the subpolar region.
Figure 4. Average winter (June–August) air-sea CO$_2$ fluxes (FCO$_2$) in the period 2015–2018, (a) averaged over biomes, (b) zonal mean flux density, and (c–f) maps of flux density. Same as Figure 2, but including also data sets with shorter coverage, and a map of the CO$_2$ flux from the BGC-float pCO$_2$-products (panel (e)), and Biogeochemical Southern Ocean State Estimate (f), and hence focusing on the period 2015–2018 for all data sets for comparability. Note that the MPI model is excluded here. The zonal mean of individual models are presented in Figure S5c in Supporting Information S1.
Figure 5. Average summer (December–February) air-sea CO\textsubscript{2} fluxes (FCO\textsubscript{2}) in the period 2015–2018. Same as Figure 4, but for summer. The zonal mean of individual models are presented in Figure S5b in Supporting Information S1.
The anthropogenic perturbation (−0.79 ± 0.13 Pg C yr\(^{-1}\), dark red Figure 3) has turned the SPSS and ICE biomes, and possibly the entire Southern Ocean, from source to sink. The large anthropogenic flux contribution in the SPSS (−0.38 ± 0.08 Pg C yr\(^{-1}\)) suppresses the natural CO\(_2\) outgassing flux. The STSS is a sink for both natural and anthropogenic flux components. The direct effect of increasing atmospheric CO\(_2\) enhances the Southern Ocean sink by −0.74 ± 0.11 Pg C yr\(^{-1}\) (light red Figure 3) and is the largest signal in the anthropogenic perturbation. A smaller component stems from the climate change effect on this steady state CO\(_2\)-induced flux (Figure S6 in Supporting Information S1). The direct CO\(_2\) effect is largest in the SPSS (−0.34 ± 0.06 Pg C yr\(^{-1}\)) where old water masses reach the surface that are undersaturated in anthropogenic carbon, followed by the STSS and ICE biomes (−0.23 ± 0.03 and −0.17 ± 0.03 Pg C yr\(^{-1}\)). In the upwelling regions, the primary effect of rising atmospheric CO\(_2\) is thus to suppress the outgassing of natural carbon.

The effect of physical climate change and variability, that is, warming and changes in wind speed patterns and strength that provoke changes in circulation (Hauck et al., 2013; Le Quéré et al., 2007; Lovenduski et al., 2007), reduces the CO\(_2\) flux into the ocean (+0.04 ± 0.07 Pg C yr\(^{-1}\), yellow in Figure 3), but is overall small in comparison to the direct CO\(_2\) effect. This climate change induced outgassing stems nearly entirely from the SPSS (+0.04 ± 0.04 Pg C yr\(^{-1}\)), with the largest contribution from the Indian sector followed by the Pacific (Figure S7 in Supporting Information S1). Thus, the climate change effect amplifies the natural CO\(_2\) outgassing, which is also the largest in the Indian and Pacific sectors of the SPSS. The climate effect is a combination of climate effects on natural and anthropogenic CO\(_2\) fluxes, which partly oppose each other (Figure S6 in Supporting Information S1).

3.2. The Seasonal Cycle of Air-Sea CO\(_2\) Fluxes in the Southern Ocean 2015–2018

We now shift our focus to seasonal fluxes by separating fluxes into separate winter (Figure 4) and summer (Figure 5) mean CO\(_2\) fluxes. For this, we examine the period 2015–2018, for which all data sets are available (see Figure S8 in Supporting Information S1 for an annual mean figure for 2015–2018 and Table S4 in Supporting Information S1 for an overview of numbers and comparison to annual means 2015–2018 and 1985–2018). The characteristics of the seasonal cycle and differences between data classes described below are robust to choice of period (2015–2018, 1985–2018).

3.2.1. Winter

In winter, all but two data sets (one GOBM and BGC-float pCO\(_2\)-products) agree that the Southern Ocean is a sink of CO\(_2\) (GOBMs: −0.83 ± 0.40 Pg C yr\(^{-1}\), pCO\(_2\) products: −0.51 ± 0.17 Pg C yr\(^{-1}\); Figure 4a, Table S4 in Supporting Information S1). The general pattern of strong uptake toward the north and a reduction toward the south is common to all data classes, though exceptions for individual GOBMs do exist (Figures 4a and 4b). Expounding on this, the strong uptake in the STSS is shown by all data sets, but further south the coherence disintegrates. Within the SPSS, there is considerable variation in position and magnitude of maximum outgassing with some GOBMs being a sink along the entire zonal mean (Figures 4a and 4b). Toward the southern reaches of the ICE biome, fluxes are more coherent as they are constrained by sea-ice cover in winter (Figure 4b). For the zonal means of individual GOBMs, see Figure S5 in Supporting Information S1.

The divergence between data class average flux estimates for the Southern Ocean are explained nearly entirely by differences in the SPSS (GOBMs: −0.15 ± 0.32 Pg C yr\(^{-1}\) and pCO\(_2\) products: 0.13 ± 0.13 Pg C yr\(^{-1}\), in Figure 4a). Note also that the spread of the individual GOBMs is the largest in the SPSS (0.32 Pg C yr\(^{-1}\)), although it is also substantial in the other biomes (STSS: 0.29 Pg C yr\(^{-1}\), ICE: 0.13 Pg C yr\(^{-1}\)) (Figure 5a). The SPSS is also where we see the largest impact of the inclusion of floats in the BGC-float pCO\(_2\)-products (Figures 4d and 4e), with the mean outgassing flux more than doubling that of the regular pCO\(_2\)-product ensemble. Interestingly, differences between B-SOSE and BGC-float pCO\(_2\)-products are large (∼1 Pg C yr\(^{-1}\), whole Southern Ocean) and FCO\(_2\) from these data sets even have a different sign for whole Southern Ocean and SPSS biome (Figure 4) although they both include BGC float data.

The zonal differences and features of fluxes between data classes are also most distinct in the SPSS (Figures 4c–4f). In short, the Atlantic sector of the SPSS has the lowest flux (weak source or even sink), while the Indian and Pacific sectors dominate the outgassing. The data-assimilated model B-SOSE has stronger localized outgassing compared with the other data classes but bear in mind that B-SOSE is only one data sets (Figure 4f), while the other data classes (Figures 4c–4e) represent up to 13, thus potentially averaging out local signals. The outgassing hotspot at the boundary between the Atlantic and Indian sectors of the SPSS can also be recognized in the
The second hotspot in the western Pacific SPSS is not distinguishable in the other data sets.

### 3.2.2. Summer

In summer, GOBMs, pCO$_2$-products and inversions largely show CO$_2$ uptake within the three Southern Ocean biomes, and outgassing north of the STSS (Figures 5a and 5b, Table S4 in Supporting Information S1). In contrast to winter, the GOBM ensemble mean for summer 2015–2018 ($-1.04 \pm 0.77$ PgC yr$^{-1}$) underestimates the CO$_2$ uptake relative to the pCO$_2$-product ensemble mean ($-1.45 \pm 0.17$ PgC yr$^{-1}$, Figure 5a). This also holds true for the data-assimilated models, where B-SOSE even simulates outgassing in the SPSS (Figures 5a, 5b, and 5f). Otherwise, the data-assimilated models, B-SOSE and ECCO-Darwin, deviate substantially from the other data classes. The differences between pCO$_2$-products with and without BGC-float data are hardly apparent in summer (Figure 5a, compared to Figure 4a). This could be due to a smaller discrepancy between ship and data-ships in summer, and/or a dominance of SOCAT data in summer for the ship + float estimate. For context, for the period 2015 through 2018, BGC-float data account for up to 70% of winter pCO$_2$ monthly by 1° × 1° measurements in the Southern Ocean (SOCAT + floats), while in summer the floats represent only 20% (Bakker et al., 2016; Bushinsky et al., 2019).

While the STSS was a region of coherence between data classes in winter (Figure 4), it is the main source of the discrepancy between the GOBM and pCO$_2$-product ensemble means in summer (GOBMs: $-0.40 \pm 0.28$ PgC yr$^{-1}$, pCO$_2$-products: $-0.73 \pm 0.07$ PgC yr$^{-1}$). The discrepancy is comparatively smaller in the SPSS (GOBMs: $-0.33 \pm 0.34$ PgC yr$^{-1}$, pCO$_2$-products: $-0.41 \pm 0.06$ PgC yr$^{-1}$). We note that CO$_2$ fluxes for both GOBMs and pCO$_2$-products show less variation from ICE to STSS in summer compared to winter (Figure 4b vs. Figure 5b, respectively). There is, nevertheless, an offset with lower GOBM CO$_2$ uptake than in pCO$_2$-products north of 55°S, and vice versa to the south. Also, the GOBM spread in the represented magnitude of the fluxes is large. In absolute terms, the GOBM ensemble spread of fluxes in summer (from $-2.03$ to $+0.28$ PgC yr$^{-1}$) is larger than in winter (from $-1.36$ to $0.12$ PgC yr$^{-1}$) or than the spread in the annual mean (from $-1.30$ to $-0.38$ PgC yr$^{-1}$; see Figure S5b in Supporting Information S1 for zonal means of individual GOBMs). This mirrors the difficulty in representing the balance between physical and biological processes in summer, which is further assessed in the next two Sections 3.2.3 and 3.2.4.

### 3.2.3. The Full Seasonal Cycle

We diagnose distinctly different seasonal cycles in the three biomes. The ICE biome has a rather clear maximum uptake in summer in the GOBM and pCO$_2$-product ensemble means, as well as most individual data sets (Figure 6a). In the STSS, the pCO$_2$-products suggest a weak seasonal cycle with a maximum uptake in autumn (Figure 6c), while the majority of GOBMs simulate a maximum CO$_2$ uptake in winter and a substantially smaller flux in summer.

The largest disagreement occurs in the SPSS, where the seasonal cycle transitions from winter outgassing in the ICE biome to summer outgassing in the STSS biomes. Here, in the SPSS, atmospheric inversions and pCO$_2$-products (including the BGC-float pCO$_2$-products), suggest the maximum CO$_2$ uptake to be in summer. In winter, the BGC-float pCO$_2$-products more than double the estimates of outgassing relative to the other pCO$_2$-products (Figure 6b). The GOBM ensemble average roughly agrees with this seasonal pattern, but simulates a reduced seasonal cycle amplitude (Figure 6b). The reduced amplitude in the SPSS is also due to different seasonal cycle phasing within the GOBM ensemble, with 8 out of 13 GOBMs simulate the maximum uptake between November and January (Figures 6b and 6d–6r), illustrating the complexity of representing the transition between the different seasonal cycle regimes. As already seen in Figures 4 and 5, the seasonal cycle differs substantially between B-SOSE and BGC-float pCO$_2$-products, particularly in the SPSS despite the new information provided by the BGC-floats. Further, the two data-assimilating models differ by a similar amount as other GOBMs differ from each other, suggesting that the model dynamics dominate over assimilated data.

In summary, most GOBMs and pCO$_2$-products agree on a summer peak in the ICE biome (but exceptions exist, Figures 6d–6r), and a winter peak at the northern boundary of the Southern Ocean. The largest discrepancy between data sets is where and how swift this transition occurs. While the use of static biomes adds to the discrepancies seen in the averaged seasonal cycles (Figures 6a–6c), the disagreement between the phasing of individual GOBMs is likely a much larger contributor to these discrepancies (Figures 6d–6p). We now turn to an investigation of the thermal and non-thermal effects on the seasonal cycle, which may help explain these discrepancies.
3.2.4. Thermal Versus Non-Thermal Effects on the Seasonal Cycle

The seasonal cycle of CO$_2$ fluxes in the Southern Ocean is a balancing act between competing thermal and non-thermal drivers (Mongwe et al., 2016, 2018; Prend et al., 2022). DIC draw-down by biological production leads to a summer minimum in pCO$_2$ and maximum in CO$_2$ uptake (Figures 7a–7c, Metzl et al., 2006; Mongwe et al., 2018). Whereas in autumn and winter, upwelling and entrainment of DIC-rich water into the mixed layer leads to a maximum in pCO$_2$ and a minimum in CO$_2$ uptake or even outgassing. Seasonal variations in mixed layer temperature further affect the solubility of CO$_2$, with lower (higher) temperatures increasing (decreasing) solubility and thus promoting CO$_2$ uptake (outgassing) (Takahashi et al., 2002).

The thermal and non-thermal components of pCO$_2$ can be decomposed to determine the dominant driver on monthly timescales (Figure 7; Mongwe et al., 2018). Here, we do this by estimating the absolute difference of the rate of change of the thermal ($p$CO$_{2\text{th}}$) and non-thermal ($p$CO$_{2\text{nt}}$) components (Figure 7; Equation 3). The contribution of salinity and total alkalinity to seasonal pCO$_2$ changes are small in the Southern Ocean and compensate for each other on a seasonal scale (e.g., Lauderdale et al., 2016; Sarmiento & Gruber, 2006), thus we here consider pCO$_{2\text{nt}}$ to be predominantlyDIC-driven.

In general, the seasonal cycle phasing of the thermal component of the GOBMs agrees well with those of the pCO$_2$-products (Figures 7d–7f). This should not come as a surprise, as GOBMs are forced by atmospheric reanalyses which assimilate observed SST (Doney et al., 2007). As a result, pCO$_{2\text{th}}$ of the pCO$_2$ seasonal cycle in the GOBMs (forced by reanalyses) compare much better to pCO$_2$ derived from the pCO$_2$-products than fully coupled Earth System Models (Mongwe et al., 2016, 2018). The non-thermal contribution is thus the primary reason for the spread between GOBMs, and for the differences between GOBMs and pCO$_2$-products (Figures 7d–7f). Thus, we group GOBMs based on whether they are predominantly DIC or thermally driven across all three biomes (Figures 7g–7i, Table S2 in Supporting Information S1), which we term DIC-dominant or DIC-weak respectively.

In DIC-weak GOBMs, the strong underestimation of pCO$_{2\text{nt}}$ causes these models to be too strongly temperature driven across the year (Figure 7). This reverses the seasonal cycle of pCO$_2$ (Figure 7c) and shifts the timing of uptake toward the colder months (when CO$_2$ solubility is largest). While in spring and summer, the role of biologically driven uptake is suppressed in favor of warming driven outgassing. This effect is largely confined to the STSS and to a lesser extent also the SPSS, and can account for the mismatch in the seasonal cycle of pCO$_2$ and CO$_2$ flux seen in some GOBMs (Figures 7b and 7c and Figures 6b and 6c). For example, in the STSS, nearly all GOBMs and specifically all DIC-weak GOBMs have a shifted season of maximum uptake from summer to winter/spring, that is, toward the colder months (Figures 7c and 7f and Table S2 in Supporting Information S1).

In terms of the underlying mechanisms driving the underestimated pCO$_{2\text{nt}}$, we hypothesize that a lack of deep vertical mixing in winter leads to too little entrainment of DIC-rich deep waters, while simultaneously allowing for too early primary production (which may then shift the growing season earlier and reduce biologically driven summer uptake). Notably, the bias in pCO$_2$ is largest in summer (DJF), followed by autumn (MAM), and is about twice as large in the DIC-weak GOBMs than in the DIC-dominant GOBMs (Figure S13 in Supporting Information S1). This further supports the lesser importance of thermal processes in the STSS and SPSS regions evident in the pCO$_2$-products.

In the ICE biome, fluxes of the GOBMs and pCO$_2$-products tend to agree much more closely in terms of their representation of the seasonal cycle (Figure 6a). This is likely related to the strong role the seasonal advance and retreat of sea ice plays in air-sea CO$_2$ fluxes, primarily through its effect as a physical barrier. The large difference in the seasonal cycle of pCO$_2$ (Figure 7a) between the GOBMs and pCO$_2$-products thus has little impact on the fluxes (Figure 6a). Further, sea ice can also have an effect on vertical mixing and light availability, thus impacting both physical and biological pathways of DIC into and out of the mixed layer (Bakker et al., 2008; Shadwicke et al., 2021; M. Yang et al., 2021).

3.3. Mean Air-Sea CO$_2$ Flux of All Data Classes

We next inspect the mean fluxes for data sets that are not available for the full time-period in an annually resolved perspective (Figure 8) in this section and the temporal evolution of the air-sea CO$_2$ fluxes from 1985 to 2018 in the next Section 3.4.

Acknowledging some agreement between GOBMs and pCO$_2$-based product ensemble means despite large spread across GOBMs (Figure 8 bars), substantial deviations among individual data sets appear. B-SOSE (2015–2018)
suggests a 0.25 PgC yr\(^{-1}\) smaller uptake than the GOBM and \(pCO_2\)-product ensemble means for the entire Southern Ocean (Figure 8a). ECCO-Darwin has the largest flux estimate into the ocean in the SPSS and the entire Southern Ocean (1.30 PgC yr\(^{-1}\), 1985–2018). Notably, the two data-assimilated models B-SOSE and ECCO-Darwin differ by a factor of 2 for the Southern Ocean wide estimate. In agreement with previous reports (Bushinsky et al., 2019), BGC-float \(pCO_2\)-products suggest Southern Ocean uptake to be 40% (0.4 PgC yr\(^{-1}\)) smaller than the \(pCO_2\)-products without BGC-float data (2015–2018). This discrepancy originates largely in the SPSS, where the BGC-float \(pCO_2\)-products estimate outgassing of 0.14 PgC yr\(^{-1}\), and the ensemble mean of the SOCAT-only-based \(pCO_2\)-products estimate a CO\(_2\) uptake of −0.13 PgC yr\(^{-1}\).

Smaller contributions to the deviation stem from the STSS and ICE biomes where BGC-float \(pCO_2\)-products report a smaller uptake by 0.14 PgC yr\(^{-1}\) when compared with the regular \(pCO_2\)-products. The Watson2020-product is generally close to the other \(pCO_2\)-products, with the exception of the SPSS where it suggests a flux of −0.18 PgC yr\(^{-1}\) (1985–2018), which is larger than any other \(pCO_2\)-product. The origin of the large SPSS difference in Watson2020 could, in part, be attributed to subtle differences in method choices in addition to different flux parameterizations (Watson et al., 2020). The atmospheric inversions produce a somewhat lower sink (−0.64 PgC yr\(^{-1}\), average over three inversions 1985–2018), but are generally close to the \(pCO_2\)-products, as they mostly use surface \(pCO_2\)-products as a prior (Table 2 and Friedlingstein et al., 2022).

### 3.4. Temporal Variability and Trends in Air-Sea CO\(_2\) Flux 1985–2018

While the STSS was a net-sink region throughout the period (Figure 8), the SPSS and ICE have turned from neutral (around 0 PgC yr\(^{-1}\)) to net sink regions since 1985, based on GOBM and \(pCO_2\)-product ensemble mean estimates. In the case of the \(pCO_2\)-products, the trends should be interpreted in the context of the sparse observations, particularly prior to the 2000s (Figures 9a and 10c). However, most individual GOBMs support the shift from neutral to uptake, with only two GOBMs (CCSM-WHOI and EC-Earth3) simulating either the ICE or SPSS biomes as regions of outgassing after 2015.

The temporal variability is quantified as the amount of “interannual variability” (IAV). This is calculated as the standard deviation of the detrended time-series, as defined in Rödenbeck et al. (2015) and Friedlingstein et al. (2022), which, in reality, captures both interannual and decadal variability components. Following this definition, the \(pCO_2\)-products have a larger interannual variability for the Southern Ocean wide integrated flux (0.09 PgC yr\(^{-1}\), range 0.04–0.16 PgC yr\(^{-1}\)) compared to the GOBMs (0.06 PgC yr\(^{-1}\), range 0.03–0.10 PgC yr\(^{-1}\), Table S3 in Supporting Information S1). Notably, the MPI-SOM-FFN \(pCO_2\)-product, which formed the basis of previous reports on Southern Ocean decadal variability (Landschützer et al., 2015), has the largest IAV of 0.16 PgC yr\(^{-1}\), about 60% larger than the next largest \(pCO_2\)-product IAV. This is in line with previous studies that found that the MPI-SOM-FFN approach may overestimate Southern Ocean variability by 30% (Gloege et al., 2021) and the decadal trend 2000–2018 by 130% (Hauck et al., 2023). The amplitude of interannual variability in the atmospheric inversion ensemble mean can only be calculated for the period since 1990 based on three ensemble members. The resulting IAV is higher than in any other data class (0.15 PgC yr\(^{-1}\), range 0.1–0.2 PgC yr\(^{-1}\)), which appears to be variability on multi-year to decadal scale (Figure 8), but is probably less robust due to the small ensemble size and different period.

Within the Southern Ocean, the strongest IAV is found in the SPSS region (0.04 PgC yr\(^{-1}\) GOBMs, 0.05 PgC yr\(^{-1}\) \(pCO_2\)-products, 0.08 PgC yr\(^{-1}\) atm. inversions), followed by the STSS (0.02 PgC yr\(^{-1}\) GOBMs, 0.03 PgC yr\(^{-1}\) \(pCO_2\)-products, 0.06 PgC yr\(^{-1}\) atm. inversions) and ICE biome (0.02 PgC yr\(^{-1}\) for GOBMs and \(pCO_2\)-products, 0.03 PgC yr\(^{-1}\) atm. inversions). Within the subpolar biome, the Indo-Pacific sector has a higher IAV (0.02 PgC yr\(^{-1}\) \(pCO_2\)-products, 0.04–0.05 PgC yr\(^{-1}\) atm. inversions) than the Atlantic sector (0.01 PgC yr\(^{-1}\) \(pCO_2\)-products, 0.02 PgC yr\(^{-1}\) \(pCO_2\)-products and atm. inversions). The large contribution to interannual variability in the SPSS may well be linked to the largest amplitude of the seasonal cycle of CO\(_2\) flux (see Section 3.2.3).

To assess the decadal-scale trends, we fit linear trends to the periods 1985–2000 and 2001–2018 (Section 2.3.3, Figures 8e–8h). This choice is motivated by the fact that the year 2000 marks roughly the mid of the considered time period and the inflection point in global and Southern Ocean CO\(_2\) uptake (Gruber et al., 2023; Landschützer et al., 2015, 2016; Le Quéré et al., 2007). We acknowledge that trend calculations are sensitive to choice of start and end years (Easterling & Wehner, 2009; Fay & McKinley, 2013), but the general statements on the two periods hold also when start and end years are varied (Table S5 in Supporting Information S1). The \(pCO_2\)-products suggest a stagnation of the flux in the STSS, and even a flux decrease in the SPSS prior to 2000. In contrast,
GOBMs suggest a continued increase in the sink in the STSS and SPSS in the same period. In the ICE biome, GOBMs and $p$CO$_2$-products result in an increasing trend (Figure 8h). After 2000, $p$CO$_2$-products and GOBMs agree on a trend toward more CO$_2$ uptake, which is significantly different from zero in all biomes except for $p$CO$_2$-products in the ICE biome (see numbers in Figures 8e–8h). However, they differ substantially in magnitude.
between GOBM and $pCO_2$-product ensemble means, particularly in the STSS (Figure 8f). The discrepancies in the magnitude of the trend act to decrease the gap between GOBM and $pCO_2$-product ensemble means in the SPSS and ICE biomes, but lead to the divergence in the flux estimate in the STSS.

On a sub-biome level (i.e., Atlantic, Indian, and Pacific sectors), all three sectors in the STSS were CO$_2$ sinks throughout the period and had weaker trends (less negative) before 2000 compared to the period after 2000 (Figure S12 in Supporting Information S1 for the DIC-dominant/weak model groups). (g–i): $\dot{\lambda}pCO_2$, the difference of the thermal and non-thermal (dissolved inorganic carbon [DIC]) components of ocean surface $pCO_2$ as in Mongwe et al. (2018). When $\dot{\lambda}pCO_2 > 0$ (red) indicates temperature dominance, and $\dot{\lambda}pCO_2 < 0$ (blue) indicates that the non-thermal component (i.e., DIC) is dominant. The MPI model is excluded in this analysis.

### 3.4.1. Comparison With In Situ $pCO_2$

Here, we evaluate the accuracy of $pCO_2$ across data classes since $pCO_2$ is the dominant driver of air-sea CO$_2$ flux variability at a monthly scale (Landschützer et al., 2016). All data sets are compared with observations (monthly gridded SOCAT v2022 data set Bakker et al., 2016, 2022; Sabine et al., 2013). The RECCAP2 data sets are...
Figure 8. Temporal evolution of the Southern Ocean air-sea CO$_2$ flux for (a) the entire Southern Ocean, the (b) subtropical seasonally stratified, (c) subpolar seasonally stratified, and (d) ice biomes. The ensemble standard deviation (1σ) averaged over the whole time series, is shown by the bars. Panels (e–h) are the same as panels (a–d) for the global ocean biogeochemistry model (GOBM) ensemble average and pCO$_2$-product ensemble average only, with linear trends between 1985–2000 and 2001–2018 as the dashed and dotted lines, respectively. The uncertainty range of the trend is calculated as one standard deviation of the trends across all GOBMs and pCO$_2$-products, respectively. Note the different y-axis scales. The separation into Atlantic, Pacific, and Indian Ocean sectors is shown in Figure S12 in Supporting Information S1.
The assimilation model, ECCO-Darwin, has a negative bias after 2000 (−13.5 ± 3.0 μatm; Figure 9a), which could explain the differences in CO$_2$ uptake between SOCAT observations and estimates for all data classes (Figure 9b). This comparison of simulated to observed pCO$_2$ at observation sites demonstrates that GOBMs are capable of reproducing SOCAT pCO$_2$ and its temporal evolution on large spatial and annual time-scales. Thus, for the period after 2000, the differences in CO$_2$ flux trend for the entire Southern Ocean between GOBMs and pCO$_2$-products (Figure 8) cannot be attributed to differences in pCO$_2$ in the regions where observations were taken. Instead, the differences arise primarily from areas where no pCO$_2$ observations exist, as also concluded in Hauck et al. (2020). The pCO$_2$ time-series calculated from the full coverage results in a lower pCO$_2$ value in the pCO$_2$-products than in the GOBMs (Figure 9a), which could explain the stronger CO$_2$ flux trend in the pCO$_2$-products (Figure 8). This discrepancy between pCO$_2$-products and GOBMs is larger in the last 10 years (2009–2019: 5.8 μatm) than the previous decade (1999–2008: 2.8 μatm, Figure 9a). Nevertheless, the root mean squared difference calculated from monthly mean data is higher in GOBMs than in assimilation products (ECCO-Darwin and Biogeochemical Southern Ocean State Estimate) are kept separate as they have different time series lengths (shown by dashed and solid gray lines, respectively). The light red area plot (right y-axis) shows the number of monthly by 1° × 1° gridded SOCAT observations per year. The comparison of the RECCAP2 GOBMs and pCO$_2$-products with gridded in situ observations (gridded SOCAT v2022 data set Sabine et al., 2013). 

It is important to note that the sampling distribution may play an important role in the magnitude of the biases calculated in any model. The root mean squared difference calculated from monthly mean data is higher in GOBMs than in assimilation products (ECCO-Darwin and Biogeochemical Southern Ocean State Estimate) are kept separate as they have different time series lengths (shown by dashed and solid gray lines, respectively). The light red area plot (right y-axis) shows the number of monthly by 1° × 1° gridded SOCAT observations per year. (a) Time-series of annually averaged pCO$_2$ from GOBMs (green), data-assimilated models (grays), and pCO$_2$-products (blue). The darker shaded lines show the annual mean calculated from the data sets subsampled to match the historic SOCAT sampling. The lighter shades show the annual mean calculated for the full coverage. The dark red line depicts the annual mean pCO$_2$ from SOCAT observations without interpolation. The assimilation products (ECCO-Darwin and Biogeochemical Southern Ocean State Estimate) are kept separate as they have different time series lengths (shown by dashed and solid gray lines, respectively). The light red area plot (right y-axis) shows the number of monthly by 1° × 1° gridded SOCAT observations per year. (b) The bias of pCO$_2$ for all data classes (subsampled to match SOCAT observations, dark lines in panel (a)) relative to SOCAT pCO$_2$ observations (solid dark red line in panel (a)). (c) The root mean squared difference (RMSD) between SOCAT observations and estimates for all data classes. Bias and RMSD were calculated on a monthly by 1° × 1° resolution, and the bias and RMSD were averaged to annual means afterward. A plot of RMSE and bias for subpolar seasonally stratified and subtropical seasonally stratified biomes and different seasons is presented in Figure S13 in Supporting Information S1.
pCO₂ summer sampling bias in the Southern Ocean has long been recognised as a potential source of biases in pCO₂ estimates, particularly for the pCO₂-products that rely heavily on the in situ data (Djeutchouang et al., 2022; Gregor et al., 2017; Metzl et al., 2006; Ritter et al., 2017). The SOCCOM project increased the number of winter samples with pH-enabled profiling floats (from 2014), suggesting stronger outgassing during winter than previously shown (Gray et al., 2018). In RECAP2, the B-SOSE assimilation model and the BGC-float pCO₂-products both make use of this data (Bushinsky et al., 2019; Verdy & Mazloff, 2017). Both of these estimates overestimate pCO₂ relative to SOCAT pCO₂ highlighting the challenge in consolidating ship-based SOCAT and BGC-float data.

3.4.2. Climate Versus CO₂ Effects on Trends in CO₂ Flux

Our analysis so far has indicated that the GOBMs reproduce seasonal temperature effects on CO₂ flux reasonably well (Figures 7d–7f), and a larger uncertainty is associated with imprints of circulation and biological activity. Next, we inspect the zonal mean and spatial patterns of the CO₂ flux trend between 1985 and 2018 (Figure 10).
Perhaps the most striking feature of the latitudinally averaged trends (Figure 10c) is that the uncertainty of the $pCO_2$-products (blue) is much larger than that of the GOBMs (green). This is contrary to the seasonal cycle (Figures 7d–7f), where the $pCO_2$-products are in much closer agreement than the GOBMs. This supports past studies which found that the sparse data in the Southern Ocean has a much larger influence on the accuracy of the decadal variability than on the seasonal variability (Gloege et al., 2021).

The $pCO_2$-products place the largest trend of CO$_2$ uptake in the entire ICE biome; however, note the large variability between $pCO_2$-products (blue shading in Figure 10c). The $pCO_2$-products show a smaller (negative) peak in the trends in the STSS between about 40 and 45°S. The GOBMs, in contrast, exhibit less zonal variability with a large meridional gradient in the ICE biome with the strongest (negative) growth rate of the trend between ~60 and 65°S. The secondary peak in the STSS (observed in the $pCO_2$-products) is not seen in the GOBM ensemble mean (green line).

Decadal variability, particularly for the $pCO_2$-products, leads to weaker trends over the full period due to opposing trends over shorter time scales (see Figure 8, Figures S14 and S15 in Supporting Information S1). For example, the $pCO_2$-products exhibit weakening CO$_2$ uptake in the Pacific and eastern Indian sector of the SPSS for the full period (Figure 10b). However, the same location experiences a growth in the sink for the period 2001–2018 (Figure S14b in Supporting Information S1). Differences between GOBMs and $pCO_2$-products in the ICE biome are large, but the smaller area of the higher latitude region decreases the influence of the discrepancy (Figure 8). Although the difference in flux density between GOBMs and $pCO_2$-products is larger in the ICE biome, the discrepancy in the STSS contributes more to the total flux trend discrepancy due to the larger area of the STSS biome (Figure 8).

We can attribute the growth in the strength of the CO$_2$ sink almost entirely to increasing atmospheric CO$_2$ levels, as shown by simulation C in the GOBMs ($F_{\text{net}}$ as black line in Figure 10c). The trends are only slightly modulated by climate change and variability, driving a weakening in the SPSS and strengthening in the northern part of the ICE biome (green $F_{\text{net}}$ vs. black $F_{\text{net}}$ in Figure 10c). Thus, the effect of climate change and variability is substantially smaller than the uncertainty in the $pCO_2$-products. But then what could be the cause of the discrepancies between the CO$_2$ flux trends of the GOBMs and $pCO_2$-products?

Past work has found that fully coupled Earth System Models in CMIP6 do not capture circulation trends accurately, which is also observed as a deviation in SST trends, with no cooling simulated in the polar Southern Ocean (Beadling et al., 2020). However, we demonstrate that the GOBMs simulate SST trends from 1985 to 2018 rather well, including the cooling in the southern parts (Figure 10d), as also shown for the thermally driven component of the $pCO_2$ seasonal cycle (Figure 8). We can therefore not ascribe the discrepancies between GOBMs and $pCO_2$-products in the trend of the CO$_2$ flux to temperature biases in the SST trends of the GOBMs.

The well-represented SST trends in the GOBMs is related to the choice of forcing the GOBMs with reanalysis data that itself depends on SST observations (Doney et al., 2007). In contrast to fully coupled Earth System models in CMIP6 (Beadling et al., 2020), the suite of models used here capture the decadal trend pattern of warming along the northern flank of the Antarctic Circumpolar Current (ACC), and cooling in the south (Figure 10, Armour et al., 2016; F. Haumann et al., 2020). The lack of warming south of 50°S was previously related to the wind-driven upwelling of deep water that had not yet been exposed to anthropogenic warming and by northward heat transport (Armour et al., 2016). More recently, the cooling was suggested to be caused by increased freshwater export from the ice region, which increases stratification and thus reduces the upward heat flux from below by warm water masses (F. Haumann et al., 2020). While the GOBM ensemble mean captures the latitudinal structure of the SST trend well, it underestimates the magnitude of peak cooling at around 60°S as well as peak warming north of 40°S.

We thus ascribe data sparsity as a reason for potential biases in the trend in the $pCO_2$-products, with the large uncertainty in $pCO_2$-products underlining that their estimate is currently not robust. Trends simulated in GOBMs are consistent within the ensemble, suggesting that they are caused by rising CO$_2$ acting on commonly reproduced characteristics of the Southern Ocean biomes. Nevertheless, biases in ocean circulation, sea ice and biology remain as possible reasons for biases in GOBMs.

### 3.5. Interior Ocean Storage of Anthropogenic Carbon 1994–2007

The focus of this section is the anthropogenic perturbation of DIC in a subset of the GOBMs (see Section 2.2.1), and in particular its accumulation rate over the period 1994 to 2007 ($\Delta C_{\text{anl}}$), in comparison with the eMLR(C*)...
observational estimate (Gruber, Clement, et al., 2019) and the ocean inverse model OCIMv2021 (DeVries, 2022).

The eMLR(C*) product uses a multiple linear regression approach to estimate ΔC_{air} and captures both the influence of CO₂-driven and climate-driven change in sea-air CO₂ fluxes and transports, whereas OCIMv2021 captures only the CO₂-driven changes.

All data classes agree in having the largest ΔC_{air} inventories within and to the north of the STSS biome (Figure 11), whose southern boundary approximately corresponds to the northern edge of the ACC. This pattern is related to the mechanisms by which C_{air} is taken up at the surface and exported to depth (Bopp et al., 2015; Mikaloff Fletcher et al., 2006; Morrison et al., 2022). Subpolar upwelling exposes old C_{air}-poor waters to elevated atmospheric CO₂ concentrations and this, combined with strong winds, drives a large influx of C_{air} in the SPSS biome (Figures 12a–12c). A small fraction of the C_{air} moves southward and is exported within Antarctic Bottom Waters, while the largest fraction is transported northward within the upper cell of the meridional overturning circulation. C_{air} air-sea fluxes remain elevated throughout the northward path, and are reinforced by the deep mixed layers in the regions where mode and intermediate waters are formed, which results in a secondary peak at around 40°S in some GOBMs, diluted by the ensemble mean (Figure 12c).

The effective transport of C_{air} into the ocean interior relies on a number of physical processes, the dominant of which is the northward transport by the overturning circulation of the C_{air} ventilated in the ocean interior by deep winter mixing (Frölicher et al., 2015; Morrison et al., 2022). The absorbed C_{air} spreads northward along density surfaces within mode and intermediate waters (Figures 12d–12f) and is circulated within and out of the Southern Ocean by the subtropical gyres (Frölicher et al., 2015; D. C. Jones et al., 2016; Waugh et al., 2019). As a result, the largest C_{air} inventories are displaced to the north with respect to the maximum air-sea C_{air} influx (Figures 12b and 12c). Another pathway by which the C_{air} inventory can build up without a corresponding surface influx is by southward advection and subsequent subduction of high-C_{air} Subtropical Waters (Iudicone et al., 2016; Morrison et al., 2022).

The observation-based product eMLR(C*) and the ocean inverse model OCIM-v2021 have similar ΔC_{air} accumulation rates when integrated over the Southern Ocean for the period 1994 through 2007 (0.52 and 0.47 PgC yr⁻¹, respectively, Figure 13), but differ in their horizontal (Figure 11) and vertical (Figure 12) patterns. The eMLR(C*) exhibits particularly low ΔC_{air} values at subpolar and high latitudes (Figure 12g), especially in the Pacific sector (Figure 11). The GOBMs multi-model-mean of ΔC_{air} accumulation rates over the same 1994 through 2007 period and integrated within the Southern Ocean (Figure 13) is 0.46 ± 0.11 PgC yr⁻¹, that is, 7% lower than the mean of the two observational estimates considered here. 6 out of the 12 GOBMs fall within the 19% range of the observational eMLR(C*) uncertainty. Two thirds of all GOBMs (hereafter “GOBMs low”) have lower than observed ΔC_{air} accumulation rates (0.39 ± 0.11 PgC yr⁻¹, about 20% lower than the observational estimates). The remaining GOBMs (hereafter “GOBMs high”) have higher than observed ΔC_{air} accumulation rates (0.58 ± 0.07 PgC yr⁻¹, about 17% higher than the observational estimates). “GOBMs high” have a higher ΔC_{air} storage than “GOBMs low” throughout the Southern Ocean (Figures 11c, 11d, and 12c), higher C_{air} air-sea fluxes (Figure 12c), higher SSS in the SPSS and STSS biomes and mixed layer depths especially in the SPSS biome (Texts S3 and S4 and Figure S17 in Supporting Information S1).

Along the zonal mean section, all GOBMs show a southward shift in ΔC_{air} with respect to eMLR(C*) (Figure 25) shown by a north-south dipole in the upper 1 km (Figure 12h), as similarly found in the comparison between OCIM-v2021 and eMLR(C*) (Figure 12g). With respect to OCIM-v2021, GOBMs show higher ΔC_{air} above 1000 m depth and lower ΔC_{air} beneath (Figure 12f). This could point to insufficient ventilation of C_{air} in “GOBMs low” models (Figure S19 in Supporting Information S1), which represent the majority of the GOBMs. The amount of spread and the overall underestimate of ΔC_{air} in the GOBMs is consistent with Earth System Models analyzed by Frölicher et al. (2015) and Terhaar et al. (2021), supporting the argument that biased ocean model dynamics and water mass properties rather than biases in the atmospheric forcing cause the C_{air} underestimate (Bourgeois et al., 2022; Terhaar et al., 2021).

Integrated over the Southern Ocean, we find that the model spread in ΔC_{air} accumulation rates from 1994 to 2007 can be largely explained (81% variance explained) by the spread in accumulated C_{air} until 1994 (Figure 13), suggesting a coherent scaling between long-term and recent C_{air} accumulation rates. The model spread in ΔC_{air} accumulation rates is also related with the spread in C_{air} air-sea fluxes averaged over 1994–2007 (78% variance explained). These results show that past long-term ΔC_{air} accumulation rates are a better predictor for present ΔC_{air} accumulation rate than present C_{air} air-sea fluxes. The reason for this is that C_{air} air-sea fluxes are linked to changes in C_{air} storage through ocean transport, which may differ substantially between models (Bourgeois et al., 2022; Frölicher et al., 2015; Terhaar et al., 2021). This becomes obvious when considering the myriad of processes involved, including the strength of the overturning circulation, the strength of the subtropical gyres,
the isopycnal stirring by mesoscale eddies, and localized subduction dynamics (Morrison et al., 2022; Sallée et al., 2012). The different way in which the GOBMs simulate these transport processes is possibly linked to the large model spread in ΔC_{ant} accumulation rates among GOBMs. Past studies have found that SSS affects the surface ocean density in the formation regions of mode and intermediate waters and could be used as a constraint of the C_{ant} air-sea fluxes, and thus of the C_{ant} storage within the recently ventilated water masses (Terhaar et al., 2021). In this study and in Terhaar et al. (2023), we find that SSS explains a lower variance in the...
ΔC_{ant} accumulation rates ($R^2 = 0.61$; Figure 13) and in the $C_{air}$, air-sea fluxes ($R^2 = 0.57$, Terhaar et al., 2023) with respect to the ESMs ($R^2 = 0.74$) analyzed by Terhaar et al. (2021). The relationship may be weaker due to the different suite of models used in the ESM and GOBM ensembles and remaining biases associated with incomplete spin-up (Terhaar et al., 2023).

### 4. Discussion

#### 4.1. Summary and Progress Since RECCAP1

We provide an updated estimate of the Southern Ocean carbon sink (see Figure 1 for regional extent). The numbers we present (Table 3) are not directly comparable with the RECCAP1 estimate (Lenton et al., 2013) due...
to different region definitions (Figure 1) and periods (1990–2009 vs. 1985–2018). The RECCAP1 regional definition of the Southern Ocean (44–75°S) cut across and missed a large part of the strong CO₂ uptake north of the Subantarctic Front. Recalculating the RECCAP2 numbers for the RECCAP1 region and period each reduces the sink estimate by ∼50% resulting in only ∼25% of its original value. A key result thus is that RECCAP2 estimates a weaker Southern Ocean carbon sink by 50% compared to RECCAP1. In RECCAP2 we find that the SPSS and ICE biomes are a weaker sink than in RECCAP1, and with the absence of the STSS, a region of strong C uptake, the other regions gain importance.
The observational and modeling communities have made substantial progress on quantifying and characterizing the Southern Ocean carbon sink since RECCAP1 (Lenton et al., 2013). The creation of the Surface Ocean CO$_2$ Atlas and its annual updates have marked a step-change by facilitating the development of statistical models (a.k.a. pCO$_2$-products). The large and diverse ensemble of pCO$_2$-products help to identify the robust features of the Southern Ocean carbon sink. The pCO$_2$-products have a relatively small spread compared to the global ocean biogeochemistry models in terms of mean and seasonal cycle, indicating that the uncertainty from differences in mapping methods is small. However, the spread in the trend estimates is in fact larger in the products than in the GOBMs (Figure 10). Further, the narrow spread in mean and seasonal cycle does not include the uncertainties due to sparse pCO$_2$ observations in the Southern Ocean, particularly in winter and before the 2000’s (Ritter et al., 2017). In addition, pCO$_2$-products share the uncertainties with GOBMs associated with the bulk formulation of air-sea CO$_2$ exchange (Fay et al., 2021; Wanninkhof et al., 2009). While they do have their shortcomings, the pCO$_2$ products are an advance for constraining the Southern Ocean carbon sink compared to the atmospheric inversions that were used in RECCAP1 (Lenton et al., 2013). This is because the surface ocean pCO$_2$ observations provide a more direct constraint on the air-sea...
CO₂ flux than the relatively small atmospheric CO₂ signals over the ocean that form the basis of the atmospheric inversions.

4.1.1. Model Deficiencies

The larger GOBM ensemble provides a more representative process-based estimate and the spread in GOBMs has been reduced since RECCAP1 (see Table 3, Lenton et al., 2013). The remaining spread is nevertheless large and points toward critical need for model development, where the largest sources of uncertainty stem from biological process description and circulation, which vary in importance depending on flux component (natural, anthropogenic, etc.), and spatio-temporal scale of interest. In terms of the anthropogenic component, the 12 GOBMs analyzed here have a 24% spread (standard deviation around the mean) in the C acc accumulation rates, which is marginally larger than the ∼20% uncertainty associated with the observational estimates of ΔC acc and C acc (even though caution is warranted when directly comparing the uncertainty estimates, which are computed formally different across data classes; Gruber, Clement, et al., 2019; Sabine et al., 2004). Overall, the GOBM ensemble mean underestimates the observation-based estimates of the C acc accumulation up to 1994 by 19% and the change between 1994 and 2007 by 7%. Admittedly, the GOBM ensemble analyzed here is relatively small, and the underestimation of C acc and ΔC acc is in the range of the uncertainty ranges of the observational estimates.

Despite the caveats of small ensemble size and overlapping uncertainty ranges, we can nonetheless speculate that the detected underestimation is likely related to a combination of physical, chemical, and methodological factors. First, our results point to too little or too shallow ventilation of mode and intermediate waters (Figure 12), the causes of which can be related to insufficient vertical mixing or too sluggish northward export of the subducted waters (Morrison et al., 2022). However, while sea-surface salinity (SSS) was singled out as a strong predictor of C acc−sea fluxes in an ESM ensemble analyzed by Terhaar et al. (2021), in our study and in Terhaar et al. (2023), SSS was not found to be a clear constraint of the anthropogenic CO₂ uptake and its interior storage in the GOBMs. Rather, Terhaar et al. (2023) find that biases in the normalized surface Revelle factor could explain the underestimation of C acc−uptake. Finally, the relatively high pre-industrial CO₂ mixing ratios related to late starting dates in several GOBMs are likely causing an underestimation of the cumulative C acc storage, which is especially large in the Southern Ocean (Terhaar et al., 2023).

For the natural carbon fluxes, the difficulty in capturing the delicate balance between physical and biological processes is clearly manifested by the large model spread (Figure 3). In addition, the different spin-up procedures could play a role. Terhaar et al. (2023) indicate that the natural CO₂ flux component may be biased toward uptake that is too strong, possibly related to GOBMs not being in steady-state (Terhaar et al., 2023), which is particularly relevant in the Southern Ocean where old water masses resurface. While long preindustrial spin-ups would bring the GOBMs closer to steady-state and thus reduce drift, they may come at the cost of less realistic surface conditions and their response to climate change and variability (Séférian et al., 2016). Interestingly, the two data-assimilated GOBMs differ to a large extent, illustrating that dynamical processes in these models may still override information gained from assimilated observations.

4.1.2. GOBM and pCO₂-Product Discrepancies

The averages of the GOBM and pCO₂-product ensembles agree for many key estimates, showing progress over the past 10 years: the mean and spatial distribution of the sink is in good agreement (Figure 2), although discrepancies of the magnitude and, particularly, the trends still persist (Figures 8 and 10; see also Canadell et al., 2021). The fact that these ensemble means agree so well in many respects provides some confidence in the Southern Ocean CO₂ flux estimates because they are nearly independent. However, the agreement of GOBMs and pCO₂-products on the mean CO₂ flux is partly a result of compensation of regional and seasonal discrepancies (Figures 4, 5, and 8).

The concerning discrepancies in the trend cannot be explained by GOBM biases in warming trends as these are well reproduced (Figure 10). Similarly, the GOBM ensemble is not systematically biased toward formation of mode and intermediate waters that is too weak, in contrast to the ESMs, and an effect on the trend of the ocean carbon sink could not be evidenced (Terhaar et al., 2023). Further potential candidates for GOBM biases, which were not explored here, are stratification (Bourgeois et al., 2022), mixing, and mixed layer dynamics, which could also lead to excess carbon accumulation in the surface layer and thus may be the driver for the overestimation of the surface Revelle factor. In the pCO₂-products, the trend is subject to biases by data sparsity (Gloge et al., 2021; Hauck et al., 2023).
Global Biogeochemical Cycles

The agreement in the mean air-sea CO₂ flux between GOBMs and pCO₂-products is also highly susceptible to the choice of river flux adjustment that either locates most outgassing of river-derived carbon in the Southern Ocean (Aumont et al., 2001) or in the tropical Atlantic (Lacroix et al., 2020). Reasons for the discrepancy between Aumont et al. (2001) and Lacroix et al. (2020) may be because of specific choices in nutrient and carbon input, lability of organic matter, resulting ocean model transport (see also the discussion in Terhaar et al., 2023). We here chose to use the river flux adjustment of Lacroix et al. (2020), scaled up to a global value of 0.65 PgC yr⁻¹, resulting in a small adjustment for the Southern Ocean of 0.04 PgC yr⁻¹. In contrast, the Southern Ocean (south of 20°S) adjustment based on Aumont et al. (2001) that is so far used in the Global Carbon Budget is higher by one order of magnitude (0.32 PgC yr⁻¹) and can explain the large mismatch in the mean flux (but not its trend) between GOBMs and pCO₂ products in the Southern Ocean in the Global Carbon Budget (Friedlingstein et al., 2022).

4.2. Seasonal Cycle and Thermal Versus Non-Thermal Drivers

As a community, we have a good understanding of the mechanisms that drive pCO₂ seasonality in the Southern Ocean (Lenton et al., 2013), but we do not fully understand their magnitudes, opposing or synergistic, in different seasons and regions (Mongwe et al., 2018). Part of this lack of understanding is due to a lack of observations throughout all seasons, though particularly acute during winter (Bushinsky et al., 2019; Gray et al., 2018; Sutton et al., 2021). Further, complex biological processes affecting pCO₂ in summer are more difficult to accurately describe in GOBMs (Mongwe et al., 2018).

The poor sampling distribution of pCO₂ in the Southern Ocean, particularly the bias toward summer sampling, could result in a misrepresentation of the seasonal cycle in pCO₂-products (Ritter et al., 2017). This may be shown by the narrow ensemble spread of the pCO₂-products during winter (Figures 7g–7i). That being said, an observing system simulation experiment (OSSE) showed that the seasonal cycle in most of the Southern Ocean is in fact well captured by one pCO₂ product (Gloege et al., 2021). The narrower GOBM spread of the non-thermal pCO₂ component during winter may also suggest that winter-time processes (circulation) are less complex than summer (circulation and biology, Figures 7g–7i).

The introduction of biogeochemical Argo floats since the mid 2010’s has increased the number of winter observations (relative to the available ship-based observations), albeit inferred from pH and estimated total alkalinity and thus associated with a higher uncertainty (Williams et al., 2017). The machine learning approaches that include float-based observations result in stronger winter outgassing (Figure 4, Bushinsky et al., 2019). Direct pCO₂ measurements showed that the years used to train the machine learning model (2015–2018) may have had anomalously high pCO₂ (Sutton et al., 2021). However, if this is in fact the case, and not related to sampling locations, this may indicate much larger interannual variability in the Southern Ocean than the majority of the pCO₂-products currently estimate (Figure 8). Incorporating these data is thus potentially an important goal for pCO₂-products, but it has proven difficult to incorporate the float-based pCO₂ estimates further back in time than 2015, the start of the BGC-float record and account for their higher uncertainty (Bushinsky et al., 2019; Williams et al., 2017).

GOBMs also have a lower pCO₂ ensemble spread during winter compared with summer and agree on the spatial location of the winter flux minimum (Figure 4). Nevertheless, the range in magnitude is still more than twice as large as those of the pCO₂-products (Figures 7g–7i). Since the thermal component is well captured in GOBMs (Figures 7d–7f), the non-thermal physical drivers (i.e., circulation) determines the uncertainty observed in winter. In summer, GOBMs have difficulty capturing the delicate balance between biological and physical processes that leads to a large spread in model pCO₂ and fluxes (Mongwe et al., 2018). GOBMs may thus benefit from more process-based studies that further our understanding of pCO₂ drivers during summer, that is, biological productivity, respiration, remineralization, and sinking of organic carbon as part of the biological carbon pump.

4.3. Temporal Variability of CO₂ Fluxes

Our analysis reduces the previously reported discrepancy in variability of Southern Ocean air-sea CO₂ fluxes between data classes (GOBMs and pCO₂-product ensemble means, Gruber, Landschützer, & Lovenduski, 2019). We relate the growing agreement to the larger ensemble of pCO₂-products in our study, with the newer additions having a substantially lower variability than the two pCO₂-products (Jena-CarboScope and SOM-FFN) used by Gruber, Landschützer, and Lovenduski (2019). A recent study using the same RECCAP data base also concluded
that there is agreement on the magnitude of interannual variability between GOBMs and \( pCO_2 \)-products (Mayot et al., 2023).

The interannual to decadal variability of Southern Ocean air-sea \( CO_2 \) fluxes was discussed extensively in the literature, and was often related to variations in the Southern Annual Mode (SAM) (Hauck et al., 2013; Lenton & Matear, 2007; Le Quéré et al., 2007; Lovenduski et al., 2007; Mayot et al., 2023; Nicholson et al., 2022). Also, regional wind variability linked to the zonal wavenumber 3 was suggested as a driver of interannual \( CO_2 \) flux variability driving both the weakening in the 1990’s and the strengthening in the 2000’s (Keppeler & Landschützer, 2019; Landschützer et al., 2015). The arguments of SAM or wave number 3 as dominant drivers of \( CO_2 \) flux interannual variability might not be fully independent from each other, as previously a wave number 3 like pattern was reported to describe MLD anomalies during positive SAM events (Saffée et al., 2010). The fact that the maximum IAV of GOBMs is found in the SPSS Indo-Pacific sector (Section 3.4, Figure S12 in Supporting Information S1) supports the argument of the above mentioned references that upwelling of carbon-rich deep water and related outgassing of natural carbon in response to a positive SAM and strengthening of westerly winds may be the dominant driver of interannual variability (DeVries et al., 2017). This is further supported by studies of atmospheric potential oxygen (APO), which can be used as a tracer of ocean-only processes from measurements of \( CO_2 \) and \( O_2 \) at atmospheric stations (Stephens et al., 1998). Nevison et al. (2020) showed that the interannual variations of APO seasonal minimum from stations in the Southern Hemisphere were strongly correlated with the SAM during years of positive phase. Further, they showed that GOBMs (as analyzed in this study) can capture the variability of \( CO_2 \) and APO fluxes driven by the SAM variations during the austral winter months. However, the study of Nevison et al. (2020) also illustrated that the SAM index variability cannot fully explain the changes in APO seasonal winter minima suggesting that other factors or modes of variability such as ENSO could impact the \( CO_2 \) and \( O_2 \) air-sea fluxes of the Southern Ocean as also previously suggested in an ocean modeling study (Verdy et al., 2007).

On top of the interannual variability, on which \( pCO_2 \) products and GOBMs seem to reach reasonable agreement, discrepancies in the \( CO_2 \) flux trend since 2000 have emerged (Figure 8, Friedlingstein et al., 2022). These discrepancies highlight a major knowledge gap and urgently need to be resolved by critical analysis of potential biases in \( pCO_2 \)-products as well as GOBMs (see Section 4.1). While there is no evidence so far that adjustments of \( CO_2 \) fluxes based on model evaluation of interfrontal salinity and Revelle factor affect the trend (Terhaar et al., 2023), data sparsity tends to lead to an overestimation of decadal variability and trend in at least two of the \( pCO_2 \)-products (Gloege et al., 2021; Hauck et al., 2023). Hence, both data classes need to be inspected for deficiencies.

### 4.4. Zonal Asymmetry of the Fluxes

While the primary spatial mode of variability in the Southern Ocean is from north to south, zonal variability in the dynamics, biogeochemistry, and carbon fluxes have been reported in the literature (Landschützer et al., 2015; Prend et al., 2022; Rintoul, 2018; Tamsitt et al., 2016). Similarly, we find substantial zonal asymmetry in both the mean states, and seasonal and interannual variability of the Southern Ocean \( CO_2 \) fluxes (Figures S10 and S12 in Supporting Information S1); yet many of our results have been presented in a zonally averaged perspective for the sake of brevity.

In this work, we find that the largest zonal asymmetries in the Southern Ocean mean air-sea \( CO_2 \) flux occur in the SPSS biome (Figures 4b–4e and Figure S12 in Supporting Information S1). Here, the Pacific and Indian sectors are larger sources (or weaker sinks) of \( CO_2 \) to the atmosphere than the Atlantic sector. This is consistent with the \( pCO_2 \)-based products (Figures S12d–S12f in Supporting Information S1). The float-based \( pCO_2 \)-products amplify this winter outgassing dramatically. However, the GOBMs and the assimilative model ensemble averages do not show a coherent and convincing increase in outgassing in the Indian and Pacific relative to the Atlantic. The zonal asymmetry reported in the observation-based products is consistent with a recent BGC-float-based study that reported stronger outgassing in the Indian and Pacific sectors of the Southern Ocean (Prend et al., 2022). The authors attributed this dominance to stronger winter-time entrainment of deep waters to the surface in these regions. The zonal asymmetry is also apparent in the air-sea \( CO_2 \) fluxes decomposed into natural and anthropogenic contributions (Figure S7 in Supporting Information S1). Here, too, the SPSS is the region with the greatest asymmetry. In the Indian sector, the large natural outgassing fluxes of the ensemble mean are nearly perfectly opposed by the anthropogenic uptake.
4.5. Link Large-Scale Synthesis to Observational Programs

The analysis presented here provides a synthesis of large-scale data sets with a focus on budgets, spatial, and temporal patterns of fluxes and carbon accumulation, and a first-order assessment of large-scale processes (e.g., thermal vs. non-thermal, anthropogenic vs. natural carbon fluxes). In particular, it highlights spatio-temporal windows for which discrepancies between data classes are largest (e.g., magnitude of winter outgassing, delicate balance of physical vs. biological processes in summer, magnitude of decadal trend of the Southern Ocean carbon sink). Importantly, this synthesis builds on contributions from many individual groups contributing repeat observations of surface and interior ocean biogeochemical properties from research vessels and ships of opportunity (e.g., Hoppeima et al., 1998; Metzl et al., 1999; Pardo et al., 2017; Talley et al., 2016; van Heuven et al., 2014). The ship-based observations form the cornerstone for many of the data classes in this study: observation-based ocean interior estimates of CO$_2$ storage assess changes in deep ocean measurements of CO$_2$, the surface pCO$_2$ estimates use observations from ships of opportunity, and the GOBMs are evaluated against ocean interior observations. And while sampling biases and gaps in the ship-based measurements may be filled by autonomous platforms with lower accuracy (e.g., BGC-floats), they will always require crossover validation measurements from the high-accuracy ship-board measurements. This emphasizes that the ship-based observations need to continue into the future to characterize the evolution of the Southern Ocean carbon cycle. This will only become more important, once stabilization of atmospheric CO$_2$ will lead to a larger weight on ocean processes for control of air-sea fluxes relative to the current quasi-exponential growth rate of atmospheric CO$_2$.

Further, detailed regional process studies employing a wide range of methodologies across disciplines are also important to further our holistic understanding of the Southern Ocean carbon cycle and to improve the description of biogeochemistry and ecosystem dynamics in GOBMs, particularly in summer. One example for such an inter-disciplinary field program is along the continental shelf west of the Antarctic Peninsula where shipboard observations indicate a strong, near-shore summer undersaturation of surface pCO$_2$ (Eveleth et al., 2017) and seasonal reduction in surface DIC (Hauri et al., 2015). The seasonal trends in the ocean CO$_2$ system on the shelf reflect a combination of biological net community production (Ducklow et al., 2018) and meltwater input diluting surface DIC and alkalinity (Hauri et al., 2015). Regional ocean biogeochemical models simulate similar onshore-offshore gradients in surface chlorophyll, biological productivity, DIC, and pCO$_2$ as well as the observed large interannual biophysical variability associated with year-to-year variations in seasonal sea-ice advance and retreat phenology (Schultz et al., 2021). Observed decadal trends for the region from the early 1990s to late 2010s indicate that reduced sea-ice extent associated with climate change drives an increase in upper ocean stability, phytoplankton biomass and biological DIC drawdown, resulting in a growing net downward air-sea CO$_2$ flux during summer (Brown et al., 2019). Recent year-round, autonomous mooring observations of pCO$_2$ and pH suggest a gradual increase in surface ocean pCO$_2$ and DIC over the fall and winter, with CO$_2$ outgassing during winter when pCO$_2$ is supersaturated largely blocked by sea-ice cover (Shadwick et al., 2021; M. Yang et al., 2021). Similar large-scale programs are needed in other parts of the Southern Ocean given its size and importance in the global carbon cycle. On-going research activities, also as part of the Southern Ocean Observing System, for example, in the Ross (Smith et al., 2021) and Weddell Seas (Arndt et al., 2022) have the potential of being extended.

5. Conclusions

Here, we present a schematic overview that summarizes the main characteristics of the Southern Ocean carbon cycle 1985–2018, as derived in this analysis and its Supporting Information S1 (Figure 14). In general, the sink strength for atmospheric CO$_2$ (net CO$_2$ flux, FCO$_2$) increases from South to North, but with important zonal asymmetry. The Atlantic and Indian Ocean sectors of the Subtropical Seasonally Stratified biome (STSS) are the regions that act as strongest sinks. In the Subpolar Seasonally Stratified biome (SPSS), the Atlantic sector stands out as the only sector acting as an annual mean CO$_2$ sink. Also the seasonal cycle shows a distinct north-south gradient. In the ice-covered biome (ICE) the peak uptake occurs in summer and is driven by the seasonal cycle of DIC, that is, by physical DIC transport and biological processes. In contrast, the dominant driver of the seasonal cycle of CO$_2$ uptake in the STSS is temperature, and thus the season of peak uptake occurs in winter. Trends in net CO$_2$ uptake derived from Global Ocean Biogeochemistry Models (GOBMs) and surface ocean pCO$_2$ observation based products (pCO$_2$-products) disagree in all biomes, but the discrepancy is strongest in the Pacific sector of the STSS. In terms of anthropogenic CO$_2$ (C$_{anth}$), the strongest uptake occurs in the SPSS. This is not visible in the map of net CO$_2$ flux, because the anthropogenic uptake manifests itself as a suppression of natural CO$_2$ outgassing. The largest anthropogenic carbon storage occurs in the STSS and northward.
Our analysis confirms the important role of the Southern Ocean in the global carbon cycle although RECCAP2 estimates a 50% smaller Southern Ocean CO₂ sink for the same region and timeframe as RECCAP1. We have highlighted key knowledge gaps that need to be closed through extended observation systems and augmented process descriptions in the dynamic models in order to further reduce uncertainties.

Data Availability Statement

All RECCAP2 data is available at Zenodo: Muller (2023).

The processed data for this study is available at Hauck et al. (2023a).

Code for each figure is in the GitHub repository: https://github.com/RECCAP2-ocean/Southern-Ocean, with a frozen version deposited at Hauck et al. (2023b).

References


Canadell, J., Monteiro, P., Costa, M., Cotrim da Cunha, L., Cox, P., Eliseev, A., et al. (Eds.). Climate change 2021: The physical science basis. Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change (pp. 673–816). https://doi.org/10.1038/s41467-019-05187-6


Global Biogeochemical Cycles

10.1029/2023GB007848

HAUCK ET AL.

35 of 40


Global Biogeochemical Cycles

10.1029/2023GB007848


