The Seasonal Cycle of pCO$_2$ and CO$_2$ fluxes in the Southern Ocean: Diagnosing Anomalies in CMIP5 Earth Systems Models

Precious N. Mongwe $^{1,2,}$, Marcello Vichi$^{2,3}$ & Pedro M.S. Monteiro$^{1,2}$

$^1$Southern Ocean Carbon-Climate Observatory (SOCCO), CSIR, Cape Town, South Africa
$^2$Department of Oceanography, University of Cape Town, Cape Town, South Africa
$^3$Marine Research Institute, University of Cape Town, Cape Town, South Africa

pmongwe@csir.co.za

Abstract

The Southern Ocean forms an important component of the global carbon cycle as a sink of CO$_2$ and heat. Recent studies based on the Coupled Model Intercomparison Project version 5 (CMIP5) Earth System Models (ESMs) show that CMIP5 models disagree on the phasing of the seasonal cycle of the CO$_2$ flux (FCO$_2$) and poorly compare with available observations estimates in the Southern Ocean. Because the seasonal cycle is the dominant mode of CO$_2$ variability in the Southern Ocean, its proper simulation is necessary to model long-term oceanic CO$_2$ changes and their related climate impacts. Here we examine the competing roles of temperature and dissolved inorganic carbon (DIC) as drivers of the seasonal cycle of pCO$_2$ in the Southern Ocean to explain the mechanistic basis for the seasonal biases in CMIP5 models, comparing them with observational products. We find that despite significant differences in the spatial characteristics of the mean annual fluxes, models show greater zonal homogeneity in the seasonal cycle of FCO$_2$ than observational products. The CMIP5 models can be grouped into one or the other of two main categories (group-SST and group-DIC) while observational products show a modest influence of both, with a dominance of DIC changes as the main driver of seasonal FCO$_2$ variability. Group-SST models show an exaggeration of the seasonal rates of change of sea surface temperature (SST) in autumn and spring during the cooling and warming peaks. The higher-than-observed rates of SST change tip the control of the seasonal cycle of pCO$_2$ and FCO$_2$ towards SST and result in a divergence between the observed and modelled seasonal cycles, particularly in the Sub-Antarctic Zone. While almost all analysed models (9 out of 10) show these SST-driven biases, 3 out of 10 (namely NorESM1-ME, HadGEM-ES and MPI-ESM, collectively the group-DIC models) compensate the solubility bias because of their overly-exaggerated primary production, such that biologically-driven DIC changes mainly regulate the seasonal cycle of FCO$_2$. Group-DIC models reproduce the observed phasing of FCO$_2$ as a result of an incorrect scaling of the biogeochemical fluxes. In the Antarctic zone, CMIP5 models compare better with observations relative to
the Sub-Antarctic Zone. This is mostly because both the CMIP5 models and the observational product show a spatial and temporal uniformity in the characteristics of FCO$_2$ in the Antarctic zone. It is unfortunately not possible to assess if CMIP5 models effectively perform better in this region or if the observational products are limited by the lack of in situ data. The suggested mechanisms should be investigated further with CMIP6 models and new available data from autonomous platforms, and our analysis framework is proposed as a useful tool to diagnose the dominant drivers.

1. Introduction

The Southern Ocean (south of 30°S) takes up about a third of the total oceanic CO$_2$ uptake, slowing down the accumulation of CO$_2$ in the atmosphere (Fung et al., 2005; Le Quere et al., 2016; Takahashi et al., 2012). The combination of upwelling deep ocean circumpolar waters (which are rich in carbon and nutrients) and the subduction of fresh-colder mid-latitude waters makes it a key region in the role of sea-air gas exchange and heat (Barbero et al., 2011; Gruber et al., 2009; Sallée et al., 2013). The Southern Ocean supplies about a third of the total nutrients responsible for biological production north of 30°S (Sarmiento et al., 2004), and accounts for about 75% of total ocean heat uptake (Frölicher et al., 2015). Recent studies suggest that the Southern Ocean CO$_2$ sink is expected to change as result of anthropogenic warming, however, the sign and magnitude of the change is still disputed (Leung et al., 2015; Roy et al., 2011; Sarmiento et al., 1998; Segschneider and Bendtsen, 2013). While some studies suggest that the Southern Ocean CO$_2$ sink is weakening and will continue to do so (e.g. Le Quéré et al., 2007; Son and Gerber, 2010; Thompson et al., 2011), other recent studies infer an increasing CO$_2$ sink (Landschutzer et al., 2015; Takahashi et al., 2012; Zickfeld et al., 2008).

Although the Southern Ocean plays a crucial role as a CO$_2$ reservoir and regulator of nutrients and heat, it remains under-sampled, especially during the winter season (JJA, Australian annual cycle) (Bakker et al., 2014; Monteiro et al., 2010). Consequently we largely rely on Earth System Models (ESM), inversions and ocean models for both process understanding and future simulation of CO$_2$ processes in the Southern Ocean. The Coupled Model Intercomparison Project (CMIP) provides an example of such a globally organized platform (Taylor et al., 2012). Recent studies based on CMIP5 ESMs, forward and inversions models show that CMIP5 models agree on the CO$_2$ annual mean sink, they disagree with available observations on the phasing of the seasonal cycle of sea-air CO$_2$ flux (FCO$_2$) in the Southern Ocean (e.g. Anav et al., 2013; Lenton et al., 2013).
The seasonal cycle is a major mode of variability for chlorophyll (Thomalla et al., 2011) and CO₂ in the Southern Ocean (Monteiro et al., 2010; Lenton et al., 2013). The large-scale seasonal states of sea-air CO₂ fluxes (FCO₂) in the Southern Ocean comprise of extremes of strong summer ingassing with a weaker ingassing or even outgassing in winter (Metzl et al., 2006). These extremes are linked by the autumn and spring transitions. In autumn CO₂ ingassing weakens linked to the increasing entrainment of sub-surface waters, which are rich in dissolved inorganic carbon (DIC), (Lenton et al., 2013; Metzl et al., 2006; Sarmiento and Gruber, 2006). During spring, the increase of primary production consumes DIC at the surface and increases the ocean capacity to take up atmospheric CO₂ (Gruber et al., 2009; Le Quéré and Saltzman, 2013; Pasquer et al., 2015; Gregor et al., 2017). The increase of sea surface temperature (SST) in summer reduces surface CO₂ solubility, which counteracts the biological uptake and reduces the CO₂ flux from the atmosphere (Takahashi et al., 2002; Lenton et al., 2013).

FCO₂ is also spatially variable in the Southern Ocean at the seasonal scale. North of 50°S is generally the main CO₂ uptake zone (Hauck et al., 2015; Sabine et al., 2004). This region forms a major part of the sub-Antarctic zone and is characterized by the confluence of upwelled, colder and nutrient-rich deep circumpolar water and mid-latitudes warm water (McNeil et al., 2007; Sallée et al., 2006). It is characterized by enhanced biological uptake during spring and solubility driven CO₂ uptake due to cool surface waters (Marinov et al., 2006; Metzl, 2009; Takahashi et al., 2012). South of 60°S towards the marginal ice zone, CO₂ fluxes are largely dominated by outgassing, driven by the upwelling of circumpolar waters, which are rich in DIC (Matear and Lenton, 2008; McNeil et al., 2007).

The inability of CMIP5 ESM to simulate a comparable FCO₂ seasonal cycle with available observations estimates in the Southern Ocean has been the subject of recent literature (e.g. Anav et al., 2013; Kessler and Tjiputra, 2016) and the mechanisms associated with these biases are still not well understood. This model-observations disagreement highlights that the current ESMs might not adequately capture the dominant seasonal processes driving the FCO₂ in the Southern Ocean. It also questions the sensitivity of models to adequately predict the Southern Ocean century scale CO₂ sink and its sensitivity to climate change feedbacks (Lenton et al., 2013). Efforts to improve simulations of CO₂ properties with respect to observations in the Southern Ocean are ongoing using forced ocean models (e.g. Pasquer et al., 2015; Rodgers et al., 2014; Visinelli et al., 2016; Rosso et al., 2017). However it remains a challenge for fully coupled simulations. In a previous study, we developed a diagnostic framework to evaluate the seasonal characteristics of the drivers of FCO₂ in ocean biogeochemical models (Mongwe et al., 2016). We here apply this approach to 10 CMIP5 models against observation product estimates in the Southern Ocean. The subsequent analysis is divided as follows; the methods section (section 2) explains our methodological
approach, followed by results (section 3), which comprise four subjections. Section 3.1 explores the spatial
variability of the annual mean representation of FCO\textsubscript{2} in the 10 CMIP5 models against observation product
estimates; section 3.2 quantitatively the biases in the FCO\textsubscript{2} seasonal cycles in the 10 models. Section 3.3
investigates surface ocean drivers of FCO\textsubscript{2} changes (temperature driven solubility and primary production),
and finally section 3.4 examines the source terms in the DIC surface budget (primary production,
entrainment rates and vertical gradients) and their role in surface pCO\textsubscript{2} changes. The discussion (section 4)
is an examination of the mechanisms behind the pCO\textsubscript{2} and FCO\textsubscript{2} biases in the models. We conclude with a
synthesis of the main findings and implications.

2. Methods

The Southern Ocean is here defined as the ocean south of the Sub-Tropical Front (STF, defined according to
Orsi et al., (1995), 11.3°C isotherm at 100m). It is divided into two main domains, the Sub-Antarctic Zone;
between the STF and the Polar Front (PF: 2°C isotherm at 200m) and the Antarctic Zone, south of the PF.
Within the Sub-Antarctic Zone and Antarctic Zone, we further partition the domain into the three main
basins of the Southern Ocean i.e. Pacific, Atlantic and the Indian Ocean.

2.1 Observations datasets

We used the Landschützer et al (2014) data product (FCO\textsubscript{2} and partial pressure of CO\textsubscript{2} (pCO\textsubscript{2}) as the main
suite of observations-based estimates against to which compare the models throughout the analysis.
Landschützer et al (2014) dataset is synthesized from Surface Ocean CO\textsubscript{2} Atlas version 2 (SOCAT2)
observations and high resolution winds using a Self Organizing Map (SOM) through a Feed Forward Neural
Network (FNN) approach (Landschützer et al., 2013). While Landschützer et al (2014) dataset is based on
more \textit{in situ} observations (SOCAT2, 15 million source measurements Bakker et al., 2014) in comparison to
Takahashi et al., 2009 (3 million surface measurements), used in Mongwe et a., (2016). We are nevertheless
mindful that due to paucity of observations in Southern Ocean, this data product is still subject to
significant uncertainties discussed in Ritter et al., (2018). To evaluate the uncertainty between data
products we compare the Landschützer et al (2014) data with Gregor et al (2017) data product, which is
based on two independent empirical models:Support Vector Regression (SVR) and Random Forest
Regression (RFR) as well as against Takahashi et al (2009) for pCO\textsubscript{2} in the Southern Ocean. We compare
pCO\textsubscript{2} instead of FCO\textsubscript{2} firstly, because Gregor et al., (2017) only provided fugacity and pCO\textsubscript{2}, and being
mindful that the choice of wind product and transfer velocity constant in computing FCO\textsubscript{2} would increase
the level of uncertainty (Swart et al., 2014). Secondly, while the focus of the paper is on the examination of biases in the air-sea flux of CO₂, the major part of our diagnostic analysis is based on pCO₂, which primarily determines the direction and part of the magnitude of the fluxes. We find that the three data products agree on the seasonal phasing of pCO₂ in the Sub-Antarctic zone, but they show differences in the magnitudes (Fig. S1). In the Antarctic zone, all three datasets agree in both phasing and amplitude (Fig. S1). At this stage it is not clear whether this agreement is due to all the methods converging even with the sparse data or the reason for agreement is the lack of observations is reason for the agreement. Nevertheless more independent in situ observations will be helpful to resolve this issue. In this regard float observations from the SOCCOM program (Johnson et al., 2017) and glider observations (Monteiro et al., 2015) for example are likely to become helpful in resolving these data uncertainties in addition to ongoing ship based measurements.

We also used the Takahashi et al. (2009) in situ FCO₂ dataset as a complementary source for comparison of spatial FCO₂ properties in the Southern Ocean. Takahashi et al., (2009) data estimates are comprised of a compilation of about 3 million surface measurements globally, obtained from 1970 – 2000 and corrected for reference year 2000. This dataset is used, as provided, on a 4° (latitude) x 5° (longitude) resolution. Using monthly mean sea surface temperature (SST) and salinity from the World Ocean Atlas 2013 (WOA13) dataset (Locarnini et al., 2013), we reconstructed total alkalinity (TAiK) using the Lee et al. (2006) formulation. We also use this dataset as the main observations platform in section 2.3. To calculate the uncertainty of the computed TAiK, we compared the calculated total alkalinity (TAiKcalc) based on ship measurements of SST and surface salinity dataset with actual observed TAiKobs of the same measurements for a set of winter (August) data collected in the Southern Ocean. We found that TAiKcalc compares well with TAiKobs (R² = 0.79) (Fig. S2, Supplementary). We then used this computed monthly TAiK and pCO₂ from Landschützer et al (2014) to compute DIC using CO2SYS (Pierrot et al., 2006, http://cdiac.ornl.gov/ftp/co2sys/CO2SYS_calc_XLS_v2.1), using K1, K2 from Mehrbach et al., 1973 refited by Dickson and Millero, 1987. For interior ocean DIC, we used the Global Ocean Data Analysis Project version 2 (GLODAP2) annual means dataset (Lauvset et al., 2016). The Mixed Layer Depth (MLD) data was taken from de Boyer Montégut et al. (2004), on a 1° x 1° grid, the data is provided as monthly means climatology and was used as provided. We also use satellite chlorophyll dataset from Johnson et al., (2013).

2.2 CMIP5 Model data
We used 10 models from the Coupled Model Intercomparison Project version 5 (CMIP5) Earth System Models (ESM) shown in Table 1. The selection criterion for the models was based on the availability of essential variables for the analysis in the CMIP5 data portal (http://pcmdi9.llnl.gov) at the time of writing: i.e. monthly FCO$_2$, pCO$_2$, chlorophyll, net primary production (NPP), surface oxygen, surface Dissolved Inorganic Carbon (DIC), MLD, Sea Surface Temperature (SST), vertical temperature fields and annual DIC for the historical scenario. The analysis is primarily based on the climatology over 1995 – 2005, which was selected to match a period closest to the available observational data product (Landschützer et al (2014), 1998 – 2011). However we do examine the consistency of the seasonality of FCO$_2$ over periods longer than 10 years by comparing the seasonal cycle of FCO$_2$ and temporal standard deviation of 30 years (1975 – 2005) vs 10 years (1995 – 2005) for HadGEM2-ES and CanESM2. We find that the seasonal cycle of FCO$_2$ remains consistent (R = 0.99) in both HadGEM2-ES and CanESM2 over 30 year (Fig. S3). All CMIP5 model outputs were regridded into a common 1°x1° regular grid throughout the analysis, except for annual CO$_2$ mean fluxes, which were computed on the original grid for each model.

Table 1: A description of the 10 CMIP5 ESMs that were used in this analysis. It shows the ocean resolution, atmospheric resolution, and available nutrients for the biogeochemical component, sea-ice model, vertical levels and the marine biogeochemical component for each ESM.
2.3 Sea-Air CO₂ Flux Drivers: The Seasonal Cycle Diagnostic Framework

The seasonal cycle diagnostic framework was developed as a way of scaling the rates of change of SST and the total DIC driven changes to the seasonal cycle of pCO₂ on a common DIC scale (Mongwe et al., 2016).

We use the framework to explore how understanding differences emerging from the temperature and DIC driven CO₂ variability could be helpful as a diagnostic of the apparent observations–model seasonal cycle biases in the Southern Ocean.

The total rate of change of DIC in the surface layer consists of the contribution of air-sea exchanges, biological, vertical and horizontal transport-driven changes (Eq. 1).

\[
\frac{\partial \text{DIC}}{\partial t}_{\text{Tot}} = \left(\frac{\partial \text{DIC}}{\partial t}\right)_{\text{air-sea}} + \left(\frac{\partial \text{DIC}}{\partial t}\right)_{\text{Bio}} + \left(\frac{\partial \text{DIC}}{\partial t}\right)_{\text{Vert}} + \left(\frac{\partial \text{DIC}}{\partial t}\right)_{\text{Hor}}
\] (1)

Because we used zonal means from medium resolution models, we assume that the horizontal terms are negligible, which leaves air-sea exchange, vertical fluxes (advection and diffusion) and biological processes as the dominant drivers of DIC. In order to constrain the contribution of temperature on the seasonal variability of pCO₂ and FCO₂ we derived a new “DIC equivalent term” (DICₜ) defined as the magnitude of DIC change that would correspond to a change in pCO₂ driven by a particular temperature change. In this way the ΔpCO₂, driven solely by modelled or observed temperature change, is converted into equivalent DIC units, which allows its contribution to be scaled against the observed or modelled total surface DIC change (Eq.1).
This calculation of DIC is done in two steps: firstly, the temperature impact on pCO$_2$ is calculated using the Takahashi et al., (1993) empirical expression that linearizes the temperature dependence of the equilibrium constants.

\[
\left( \frac{\partial pCO_2}{\partial T} \right)_{SST} = 0.0423 \times pCO_2 \times \left( \frac{\partial pCO_2}{\partial SST} \right)
\]  

(2)

Though this relationship between dSST and dpCO$_2$ is based on a linear assumption (Takahashi et al., 1993), this formulation has been shown to hold and has been widely used in literature (e.g. Bakker et al., 2014; Feely et al., 2004; Marinov and Gnanadesikan, 2011; Takahashi et al., 2002; Wanninkhof et al., 2010; Landschützer et al., 2018). We show in the supplementary material that the extension of this expression into polar temperature ranges (SST < 2$^\circ$C) only introduces a minor additional uncertainty of 4 -5% (SM Fig. S4)

Secondly, the temperature driven change in pCO$_2$ is converted to an equivalent DIC using the Revelle factor.

\[
\left( \frac{\partial DIC_T}{\partial T} \right)_{SST} = \frac{\gamma_{DIC} \times pCO_2 \times \left( \frac{\partial pCO_2}{\partial SST} \right)}{\frac{\partial DIC}{\partial pCO_2}}
\]  

(3)

Here we also used a fixed value for the Revelle Factor ($\gamma_{DIC}$=14), typical of polar waters the Southern Ocean but in order to assess the error linked to this assumption. We recomputed the Revelle factor in the Sub-Antarctic and Antarctic zones using annual mean climatologies of TAlk, salinity, sea surface temperature and nutrients. Firstly we examined DIC changes for the nominal range of pCO$_2$ change (340 – 399 μatm:1 μatm intervals) and then used this dataset to derive the Revelle factor. The range of calculated Revelle factors in the Southern Ocean was between $\gamma_{DIC}$ ~ 12 – 15.5 with an average of $\gamma_{DIC}$ = 13.9±1.3. This justifies our use of $\gamma_{DIC}$ = 14 for the conversion of the solubility driven pCO$_2$ change to an equivalent DIC (DICT) throughout the analysis. We have provided the uncertainty that this conversion makes into the temperature constraint DIC$_T$, by using the upper and lower limits of the Revelle factor ($\gamma_{DIC}$ = 12 – 15.5) in the model framework. In the Supplementary Material (Fig. S5) we show examples for observations in the Sub-Antarctic and Antarctic zones, which show that the extremes of the Revelle factor values ($\gamma_{DIC}$ = 12 – 15.5) do not alter the phasing or magnitude of the relative controls of temperature or DIC on the seasonal cycle of pCO$_2$.

The rate of change of DIC was discretized on a monthly mean as follows:
\[
\left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{SST}} \approx \left( \frac{\partial \text{DIC}}{\partial t} \right)_{n,l} = \frac{\text{DIC}_{n+1,l} - \text{DIC}_{n,l}}{1 \text{ month}}
\]  

(4)

235

236

Where \( n \) is time in month, \( l \) is vertical level (in this case the surface, \( l=1 \)). We here take the forward derivative such that November rate is the difference between December the 15\textsuperscript{th} and November the 15\textsuperscript{th}, thus being centered at the interval between the months.

Finally, to characterize periods of temperature or DIC dominance as main drivers of the instantaneous (monthly) pCO\(_2\) change we subtract Eq. 4 from Eq. 1, which yields a residual indicator \( M_{T-\text{DIC}} \). Eq. 5. \( M_{T-\text{DIC}} \) is then used as indicator of the dominant driver of instantaneous pCO\(_2\) changes, in this scale monthly time scale.

\[ M_{T-\text{DIC}} = \left| \left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{SST}} \right| - \left| \left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{Tot}} \right| \]  

(5)

236

237

\( M_{T-\text{DIC}} > 0 \) indicates that the pCO\(_2\) variability is dominated by the temperature driven solubility and when \( M_{T-\text{DIC}} < 0 \), it indicates that pCO\(_2\) changes are mainly modulated by DIC processes (i.e. Biological CO\(_2\) changes and vertical scale physical DIC mechanisms). We also the following DIC processes; i.) Biological DIC changes using chlorophyll, NPP, export carbon, surface oxygen, and ii.) . Physical DIC mechanisms using estimated entrainment rates at the base of the mixed layer: details of this calculation are in section 2.4.

In the Southern Ocean, salinity and TAlk are considered lower order drivers of the seasonal cycle of pCO\(_2\) (Takahashi et al., 1993). In the supplementary material (Fig. 6), we show that salinity and TAlk do not play a major role as drivers of the local seasonal cycle of pCO\(_2\). We do so by computing the equivalent rate of change of DIC resulting from seasonal variability of salinity and TAlk as done for temperature (Eq. 2), i.e. still assuming empirical linear relationships from Takahashi et al (1993): \( \left( \frac{\ln (\text{pCO}_2)}{\ln (\text{TAlk})} \right) \approx -9.4 \) and \( \left( \frac{\ln (\text{pCO}_2)}{\ln (\text{Sal})} \right) = 0.94 \). By applying these relationships to the model data, we confirmed that indeed salinity and TAlk are secondary drivers of pCO\(_2\) changes i.e. \( \left[ \left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{Tot}} \right]_{\text{average}} \approx 5 \mu\text{mol} \text{ kg}^{-1} \text{ month}^{-1} \), while \( \left[ \left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{TAlk}} \right]_{\text{average}} \approx 0.6 \mu\text{mol} \text{ kg}^{-1} \text{ month}^{-1} \) and \( \left[ \left( \frac{\partial \text{DIC}}{\partial t} \right)_{\text{TAlk}} \right]_{\text{maximum}} \approx 0.4 \mu\text{mol} \text{ kg}^{-1} \text{ month}^{-1} \).

The seasonal cycle of the ocean-atmosphere pCO\(_2\) gradient (\( \Delta \text{pCO}_2 \)) is the main driver of the variability of FCO\(_2\) over comparable periods (Sarmiento and Gruber, 2006; Wanninkhof et al., 2009; Mongwe et al., 2016). Wind speed plays a dual role as a driver of FCO\(_2\): it drives the seasonal evolution of buoyancy-mixing dynamics, which influences the biogeochemistry and upper water column physics but these
processes are incorporated into the variability of the DIC. Wind speed also drives the rate of gas exchange across the air-seas interface (Wanninkhof et al., 2013) however, because winds in the Southern Ocean do not have large seasonal variation (Young, 1999), for this analysis we neglect the role of wind as secondary driver of the seasonal cycle of FCO₂. Consequently, the seasonal cycle of FCO₂ is directly linked to surface pCO₂ are driven by changes in temperature, salinity, TAlk and DIC (Sarmiento and Gruber, 2006; Wanninkhof et al., 2009). In this analysis we use this assumption as a basis to explore how the seasonal variability of temperature and DIC regulate the seasonal cycle of pCO₂ in CMIP5 models relative to observational product estimates.

2.4 Entrainment mixing

CO₂ uptake by the Southern Ocean has been shown to weaken during winter in the Southern Ocean linked to the entrainment of sub-surface DIC as the MLD deepens (e.g. Lenton et al., 2013; Metzl et al., 2006; Takahashi et al., 2009). Here we estimate this rate of entrainment (RE) using Eq. 6, which estimate the advection of preformed DIC at the base of the mixed layer:

\[
RE = U_e \left( \frac{\Delta DIC}{\Delta z} \right)_{MLD} \tag{6}
\]

\[
RE_n = \left( \frac{\Delta MLD_n}{\Delta t} \right) \left( \frac{\Delta DIC}{\Delta z} \right)_{n,MLD} \tag{7}
\]

\[
\left( \frac{\Delta DIC}{\Delta z} \right)_{n,MLD} = \frac{DIC_{n,MLD_{n+1}} - DIC_{n,MLD_n}}{\Delta z} \tag{8}
\]

In which \( U_e \) is an equivalent entrainment velocity based on the rate of change of the MLD. This approximation of vertical entrainment is necessary as it is not possible to compute this term from the CMIP5 data because the vertical DIC distribution is only available as annual means. We use the entrainment rates to estimate the influence of subsurface/bottom DIC changes on surface DIC changes driven and subsequently pCO₂ and FCO₂. Because we are mainly interested in the period autumn – winter, where the MLD \( \geq 60 \) m in the Sub-Antarctic zone and \( \geq 40 \) m in the Antarctic zone, at this depth seasonal variations in DIC are anticipated to be minimal thus these estimates can be used. The monthly and annual mean DIC from a NEMO PISCES 0.5 x 0.5o model output was used to estimate the uncertainty by comparing RE computed from both (Dufour et al., 2013). We found that the annual and monthly estimates to be indeed comparable with minimal differences (not shown). It is noted as a caveat that this rate of entrainment is only a coarse estimate because we were using annual means, and is intended only for the autumn-winter period when MLDs are deepen.
3. Results

3.1 Annual climatological sea-air CO$_2$ fluxes

The annual mean climatological distribution of FCO$_2$ in the Southern Ocean obtained from observational products is spatially variable but mainly characterized by two key features: (i) CO$_2$ in-gassing north of 50°S and 55°S (Polar Frontal zone, PFZ) within and north of the Sub-Antarctic Zone, and (ii), CO$_2$ out-gassing between the PF (~58°S) and the Marginal Ice Zone (MIZ, ~ 60° - 68°S) (Fig. 1a-b). Most CMIP5 models broadly capture these features, however, they also show significant differences in space and magnitude between the basins of the Southern Ocean (Fig. 1). With the exception of CMCC-CESM, which shows a northerly-extended CO$_2$ outgassing band between about 40°S and 50°S, CMIP5 models generally show the CO$_2$ outgassing zone between 50°S – 70°S in agreement with observational estimates (Fig. 1).

The analyzed 10 CMIP5 models show a large spatial dispersion in the spatial representation of the magnitudes of FCO$_2$ with respect to observations (Fig. 1, Table 2). They generally overestimate the upwelling-driven CO$_2$ outgassing (55°S -70°S) in some basins relative to observations. IPSL-CM5A, CanESM2, MPI-ESM, GFDL-ESM2M and MRI-ESM, for example, show CO$_2$ outgassing fluxes reaching up to 25 g m$^{-2}$ yr$^{-1}$, while observations only show a maximum of 8 g m$^{-2}$ yr$^{-1}$ (Fig. 1). Between 40°S - 56°S (Sub-Antarctic zone), observations and CMIP5 models largely agree, showing a CO$_2$ in-gassing feature, which is mainly attributable to biological processes (McNeil et al., 2007; Takahashi et al., 2012). South of 65°S, in the MIZ, models generally show an excessive CO$_2$ ingassing with respect to observations (with the exception of CanESM2, IPSL-CM5A-MR and CNRM-CM5). Note that as much as this bias south of the MIZ might be a true divergence of CMIP5 models from the observed ocean, it is also possibly due to the lack of observations in this region, especially during the winter season (Bakker et al., 2014; Monteiro, 2010).

Table 2 shows the Pattern Correlation Coefficient (PCC) and the Root Mean Square Error (RMSE), which are here used to quantify the model spatial and magnitude performances against Landschützer et al (2014) data product. Out of the 10 models, 6 show a moderate spatial correlation with Landschützer et al (2014) (PCC = 0.40 – 0.60), i.e. CNRM-CM5, GFDL-ESM2M, HadGEM2-ES, IPSL-CM5A-MR, CESM1-BGC, NorESM-ME and CanESM2. While MPI-ESM-MR (PCC = 0.37), MRI-ESM (PCC = 0.36) and CMCC-CESM (PCC = -0.09) show a weak to null spatial correlation with observations, the latter mainly due to the overestimated outgassing region. Spatially, GFDL-ESM2M and NorESM1-ME are the most comparable to Landschützer et al...
NorESM1-ME and CESM1-BGC are the only 2 of the 10 models showing a consistent spatial (RMSE < 10) and magnitude (PCC = 0.50) performance. From Table 2, it is evident that an appropriate representation of the spatial properties of FCO$_2$ with respect to observations does not always correspond to comparable magnitudes. CanESM2 for example shows a good spatial comparison (PCC = 0.54), yet a poor estimation of the magnitudes (RMSE = 19.5). In this case caused by an overestimation of CO$_2$ uptake north of 55°S (= - 28 g m$^{-2}$ yr$^{-1}$) and CO$_2$ outgassing (> 25 g m$^{-2}$ yr$^{-1}$) in the Antarctic zone, resulting in a net total Southern Ocean annual weak sink (-0.05 Pg C m$^{-2}$ yr$^{-1}$). These inconsistencies in the spatial and magnitude performances highlights some of the limitations of using annual mean indicators to evaluate model performance and thus a process-based diagnostic approach could be useful in understanding the departure of models from observed estimates.

### 3.2 Sea-Air CO$_2$ Flux Seasonal Cycle Variability and Biases

The seasonal cycle of FCO$_2$ is shown in Fig. 2. The seasonality of FCO$_2$ in the 10 CMIP5 models shows a large dispersion in both phasing and amplitude, but mostly disagree with observations in the phase of the seasonal cycle as well as with each other. More quantitatively, CMIP5 models show weak to negative correlations with the Landschützer et al (2014) data product in the Sub-Antarctic Zone and have slightly higher correlations in the Antarctic Zone (see supplementary Fig. S7). This discrepancy is consistent with Anav et al., (2013) findings, who however used fixed latitude criteria. Based on the phasing, the seasonality of FCO$_2$ in CMIP5 models can be a priori divided in two main groups: group-DIC models, comprising of MPI-ESM, HadGEM-ES and NorESM1-ME, and group-SST models, the remainder i.e. GFDL-ESM2M, CMCC-CESM, CNRM-CERFACS, IPSL-CM5A-MR, CESM1-BGC, MRI-ESM and CanESM2. The naming convention is suggestive of the mechanism driving the seasonal cycle, as it will be clarified further on. A similar grouping was also identified by Kessler and Tjiputra (2016) using a different criterion. Fig. 3 shows the seasonal cycle of FCO$_2$ of an equally-weighted ensemble of the two groups compared to observations, the shaded area shows the decadal standard deviation for the models and the Landschützer et al (2014) data product for 1998 -2014 standard deviation for in the various regions.

In the Sub-Antarctic zone, the observational products show a weakening of CO$_2$ uptake during winter (less negative values in June-August) with values close to the zero at the onset of spring (September) in all three basins. Similarly, during the spring season, all three basins are seen to maintain a steady increase of CO$_2$
uptake until mid-summer (December), while they differ during autumn (March-May). The Pacific Ocean shows an increase in CO₂ uptake during autumn that is not observed in the other basins (only marginally in the Indian Ocean). In the Antarctic zone, the observed FCO₂ seasonal cycle is mostly similar in all three basins (Fig. 3d-f). While this seasonal cycle consistency may suggest a spatial uniformity of the mechanisms of FCO₂ at the Antarctic, we are also mindful that this may be due to a result of the paucity of observations in this area. In the Antarctic zone, all three basins show a weakening of CO₂ uptake from the onset of autumn (March) until mid-winter (June–July) when it outgasses. The winter CO₂ outgassing is followed by a strengthening of the CO₂ uptake throughout spring to summer, when it reaches a CO₂ ingassing peak.

The differences in the seasonal cycle of FCO₂ across the three basins of the Sub-Antarctic zone found in the observational product (Fig. 2), likely resemble the differences in the spatial behavior seen in Fig. 1. To verify this, we correlate the seasonal cycles from the Landschützer et al (2014) observational product in the three basins (Fig. 4). The FCO₂ seasonal cycle in the Sub-Antarctic Atlantic and Indian basins is the only one that is similar (R = 0.8), while the other basins are quite different to each other (R = -0.1 for Pacific – Atlantic and R ~ 0.4 for Pacific – Indian). Contrary to the observational product, CMIP5 models show the same seasonal cycle phasing across all three basins in the Sub-Antarctic zone (basin – basin correlation coefficients are always larger than 0.50 in Fig. 4), with the exception of three models (i.e. CMCC-CESM, CESM-BGC1 and GFDL-ESM2M). In the Antarctic zone, CMIP5 models agree with observations in the spatial uniformity of the seasonal cycle of FCO₂ among the three basins.

Group-DIC models are characterized by an exaggerated CO₂ uptake during spring-summer (Fig. 3) with respect to observations estimates and CO₂ outgassing during winter. These models generally agree with observations in the phasing of CO₂ uptake during spring, but overestimate the magnitudes. It is worth noting that the seasonal characteristics of group-DIC models are mostly in agreement with the observations in the Atlantic and Indian basin in Sub-Antarctic zone (R > 0.5 in Fig. 4). The large standard deviation (~ 0.01 g C m⁻² day⁻¹) during the winter and spring-summer seasons in the Atlantic Ocean shows that though group-DIC models agree in the phase, magnitudes vary considerably (Fig. 3b). For example MPI-ESM reach up to 0.06 g C m⁻² day⁻¹ outgassing during winter, while HadESM2-ES and NorESM2 peak only at ~ 0.03 g C m⁻² day⁻¹. Group-SST models on the other hand are characterized by a CO₂ outgassing peak in summer (Dec-Feb) and a CO₂ in-gassing peak at the end of autumn (May) and their phase is opposite to the observational estimates in the Atlantic and Indian basins (Fig. 3b,c). Group-SST models only show a strengthening of CO₂ uptake during spring in the Indian Ocean. Interestingly, group-SST models compare relatively well with the observed FCO₂ seasonal cycle in the Pacific Ocean, whereas group-DIC models disagree the most with the observed estimates (Fig. 3a). This phasing differences within models and against observed estimates
probably suggests that the disagreement of CMIP5 models FCO₂ with observations is not a matter of a relative error/constant magnitude offset, but likely point to differences in the seasonal drivers of FCO₂.

In the Antarctic zone (Fig. 3d-f), both group-DIC and group-SST models perform better than the Sub-Antarctic, also in more quantitative terms as shown by the correlation analysis in Fig. S7. However, the similarity in the seasonality of the different basins found in the observational product is now properly simulated by the models (Fig. 4, with the exception of MRI-ESM and CanESM2 where R < 0 for all three basins). Here FCO₂ magnitudes oscillate around zero with the largest disagreements occurring during mid-summer, where observations estimates shows a weak CO₂ sink (≈ -0.03 gC m⁻² day⁻¹), group-SST showing a zero net CO₂ flux and a strong uptake in group-DIC shows (e.g. ≈ -0.12 gC m⁻² day⁻¹ in the Pacific Ocean). The large standard deviation (≈ 0.01 gC m⁻² day⁻¹) here indicates considerable differences among models (Fig. 3d-f).

### 3.3 Seasonal Scale Drivers of Sea-Air CO₂ Flux

We now examine how changes in temperature and DIC regulate FCO₂ variability at the seasonal scale following the method described in Sec. 2.3. Fig. 5 shows the monthly rates of change of SST (dSST/dt) for the 10 models compared with WOA13 SST. CMIP5 generally shows agreement in the timing of the switch from surface cooling (dSST/dt < 0) to warming (dSST/dt > 0) and vice versa; i.e. March (summer to autumn), and September (winter to spring) respectively. In both the Sub-Antarctic and Antarctic zone CMIP5 models agree with observations in this timing (Fig. 5). However, while they agree in phasing, the amplitude of these warming and cooling rates are overestimated with respect to the WOA13 dataset with exception of NorESM1-ME. Subsequently these differences in the magnitude of dSST/dt have important implications for the solubility of CO₂ in seawater; larger magnitudes of |dSST/dt| are likely to enhance the response of the pCO₂ to temperature through CO₂ solubility changes. For example, because the observations in the Indian Ocean shows a warming rate of about 0.5°C month⁻¹ lower compared to the other two basins, we expect a relatively weaker role of surface temperature in this basin.

As described in sec. 2.3, the computed dSSt/dt magnitudes were used to estimate the equivalent rate of change of DIC driven by CO₂ solubility using Eq. 2. The seasonal cycle of |(dDIC∥/dt)SST| vs |(dDIC/dt)Tot|, for the 10 models and observations is presented in the supplementary material (Fig. S8), here we show the seasonal mean of M₇DIC Eq. 3. As articulated in sec. 2.3, M₇DIC (Fig. 6) is the difference between the total surface DIC rate of change of DIC (Eq. 1) and the estimated equivalent temperature driven solubility DIC changes Eq. 3, such that when |(dDIC∥/dt)SST| > |(dDIC/dt)Tot|, temperature is the dominant driver of the
instantaneous pCO$_2$ changes, and conversely when $| (d\text{DIC}/dt)_{\text{SST}} | < | (d\text{DIC}/dt)_{\text{Tot}} |$, DIC processes is the dominant mode in the instantaneous pCO$_2$ variability. The models showing the former feature are SST-driven and belong to group-SST, while the models showing the latter are DIC-driven and belong to group-DIC.

According to the $M_{\text{T-DIC}}$ magnitudes in Fig. 6, the seasonal cycle of pCO$_2$ in the observational estimates is predominantly DIC-driven most of the year in both the Sub-Antarctic and Antarctic zone. Note that, however, during periods of high $|d\text{SST}/dt|$, i.e. autumn and spring, observations show a moderate to weak DIC control ($M_{\text{T-DIC}} \approx 0$). The Antarctic zone is mostly characterized by a stronger DIC control (mean Annual $M_{\text{T-DIC}} > 3$) except for the spring season (Fig. 6). Consistent with the similarity analysis presented in Fig. 4, the Antarctic zone shows coherence in the sign of the temperature –DIC indicator ($M_{\text{T-DIC}} > 0$) within the three basins.

3.4 Source terms in the DIC surface budget

To further constrain the surface DIC budget in Eq. 1, we examine the role of the biological source term using chlorophyll and Net Primary Production (NPP) as proxies. Fig. 8 shows the seasonal cycle of chlorophyll, NPP and the rate of surface DIC changes (dDIC/dt). The observed seasonal cycle of chlorophyll (Johnson et al., 2013) shows a similar seasonal cycle within the three basins during the spring – summer seasons (autumn-winter data are removed due to the satellite limitation) in both Sub-Antarctic and Antarctic zone. Magnitudes are however different in the Sub-Antarctic zone; the Atlantic basin shows larger chlorophyll magnitudes (Chlorophyll reach up to 1.0 mg m$^{-3}$) compared to the Pacific and Indian basins (Chl < 1 mg m$^{-3}$).

CMIP5 models here show a clear partition between group-DIC and group-SST models. While they mostly maintain the same phase, group-DIC shows larger amplitudes of chlorophyll relative to group-SST and observed estimates in the Sub-Antarctic zone. This difference is even clearer in NPP magnitudes, where group-DIC models show a maximum of NPP > 1 mmol m$^{-2}$ s$^{-1}$ in summer, while group-SST magnitudes shows about half of it. Except for CESM1-BGC and CMCC-CESM (and NorESM1-ME for NPP), each CMIP5 model generally maintains a similar chlorophyll seasonal cycle (phase and magnitude) in all three basins of the Southern Ocean. This is contrary to the observations, which show differences in the magnitude. Consistently with the observational product, CESM1-BGC simulates larger amplitude in the Atlantic basin. While CMCC-CESM also has this feature, it also shows an overestimated chlorophyll peak in the Indian
Ocean. In the Antarctic zone both observations and CMIP5 models generally agree in both phase and magnitude (except for CanESM2) of the seasonal cycle of chlorophyll in all three basins.

We now examine the influence of the vertical DIC rate in Eq. 1, using estimated entrainment rates (RE, Eq. 5) based on MLD and vertical DIC gradients (see sec. 2.3). Fig. 7 shows the seasonal changes of MLD compared with the rate from the observational product. CMIP5 models largely agree on the timing of the onset of MLD deepening (February in the Pacific Ocean, and March for the Atlantic and Indian Ocean) and shoaling (September) in the Sub-Antarctic zone (with the exception of NorESM1-ME and IPSL-CMSA in the Pacific Ocean). The Indian Ocean generally shows deeper winter MLD in both observations and CMIP5 models in the Sub-Antarctic zone. Note that while CMIP5 models generally show the observed deeper MLDs in the Indian Ocean, they show a large variation; for example, the winter maximum depth range from 100 m (CMCC-CESM, Pacific Ocean) to 350 m (CanESM2, Indian Ocean) in the Sub-Antarctic zone. In the Antarctic zone CMIP5 models are largely in agreement on the timing of the onset of MLD deepening (February), but also variable in their winter maximum depth. It is worth noting that the observed MLD seasonal cycle might be biased due to limited in situ observations particularly in the Antarctic zone (de Boyer Montégut et al., 2004).

The estimated RE values in Fig. 10 show that almost all CMIP5 (with the exception of NorESM1-ME) entrain subsurface DIC into the mixed layer during autumn–winter in agreement with the observational estimates. In the Sub-Antarctic zone, the estimates using the observational products show the strongest entrainment in the Atlantic Ocean in May (RE reaches up to 10 \( \mu \text{mol kg}^{-1} \text{ month}^{-1} \)), while it is lower in the other basins. In the Antarctic zone, observed RE conversely shows stronger entrainment rates in the Pacific and Indian Ocean (RE > 15 \( \mu \text{mol kg}^{-1} \text{ month}^{-1} \)) in comparison to the Atlantic basin (RE = 11 \( \mu \text{mol kg}^{-1} \text{ month}^{-1} \)). CMIP5 models entrainment rates are variable but not showing any particular deficiency when compared with the observational estimates. Also, the group-DIC and group-SST models show no clear distinction, the major striking features being the relatively stronger entrainment in MPI-ESM and CanESM2 across the three basins in the Sub-Antarctic zone in mid to late winter (RE = 15 \( \mu \text{mol kg}^{-1} \text{ month}^{-1} \)) and the large winter entrainment in IPSL-CM5A-MR in the Antarctic Pacific Ocean. The supply of DIC to the surface due to vertical entrainment is therefore generally comparable between model simulations and the available estimate.

However, our RE estimates are estimated at the base of the mixed layer, which is not necessarily a complete measure of the vertical flux of DIC at the surface. We therefore investigate the annual mean vertical DIC gradients in Fig. 10 as an indicator of where the surface uptake processes occur. The simulated CMIP5 profiles are similar to GLODAP2, but some differences arise. In the Sub-Antarctic zone, GLODAP2
shows a shallower surface maximum in the Atlantic basin consistent with higher biomass in this basin (Fig. 8) \((d\text{DIC}/dz)_{\text{smax}} = 0.55 \text{ μmol kg}^{-1} \text{ m}^{-1}\), at 50 m) compared to the Pacific \((d\text{DIC}/dz)_{\text{smax}} = 0.60 \text{ μmol kg}^{-1} \text{ m}^{-1}\), at 80 m) and Indian basin \((d\text{DIC}/dz)_{\text{smax}} = 0.40 \text{ μmol kg}^{-1} \text{ m}^{-1}\), at 80 m). CMIP5 models generally do not show this feature in the Sub-Antarctic zone, except for CESM1-BGC1 \((d\text{DIC}/dz)_{\text{smax}} = 0.50 \text{ μmol kg}^{-1} \text{ m}^{-1}\), at 50 m). Instead, they show the surface maxima at the same depth in all three basins. In the Antarctic zone both CMIP5 models and observations shows larger \((d\text{DIC}/dz)_{\text{smax}}\) magnitudes and nearer surface maxima (with the exception of CanESM2 and CESM1-BGC). This difference in the position and magnitude of the DIC maxima between the Sub-Antarctic and Antarctic zone has important implications for surface DIC changes and subsequently pCO$_2$ seasonal variability. Because of the nearer surface DIC maxima in the Antarctic zone, surface DIC changes are mostly influenced by these strong near surface vertical gradients than MLD changes. This implies that even if the entrainment rates at the base of the MLD are comparable between the Sub-Antarctic and the Antarctic, the surface supply of DIC may be larger in the Antarctic zone.

4. Discussion

Recent studies have highlighted that important differences exist between the seasonal cycle of pCO$_2$ in models and observations in the Southern Ocean (Lenton et al., 2013; Anav et al., 2015; Mongwe, 2016). Paradoxically, although the models may be in relative agreement for the mean annual flux, they diverge in the phasing and magnitude of the seasonal cycle (Lenton et al., 2013; Anav et al., 2015; Mongwe, 2016). These differences in the seasonal cycle raise questions about the climate sensitivity of the carbon cycle in these models because they may reflect differences in the process sensitivities to drivers that are themselves climate sensitive.

In this study we expand on the framework proposed by Mongwe et al. (2016), which examined the competing roles of temperature and DIC as drivers of pCO$_2$ variability and the seasonal cycle of pCO$_2$ in the Southern Ocean, to explain the mechanistic basis for seasonal biases of pCO$_2$ and FCO$_2$ between observational products and CMIP5 models. This analysis of 10 CMIP5 models and one observational product (Landschutzer et al., 2014) highlighted that although the models showed different seasonal modes (Fig. 2), they could be grouped into two categories (SST- and DIC-driven) according to their mean seasonal bias of temperature or DIC control (Fig. 3 & 6).

A few general insights emerge from this analysis. Firstly, despite significant differences in the spatial characteristics of the mean annual fluxes (Fig. 1), models show unexpectedly greater inter-basin coherence,
in the phasing seasonal cycle of FCO$_2$ and SST-DIC control than observational products (Fig. 3 & 6). Clear inter-basin differences have been highlighted in studies on the climatology and interannual variability that examined pCO$_2$ and CO$_2$ fluxes based on data products (Landschutzer et al., 2015; Gregor et al., 2017) as well as phytoplankton chlorophyll based on remote sensing (Thomalla et al., 2011; Carranza et al., 2016). Briefly, the Atlantic Ocean shows the highest mean primary production in contrast to the Pacific Ocean, which has the lowest (Thomalla et al., 2011). Similarly, strong inter-basin differences for pCO$_2$ and FCO$_2$ have been highlighted and ascribed to SST control (Landschutzer et al., 2016) and wind stress - mixed layer depth (Gregor et al., 2017). The combined effect of these regional differences in forcing of pCO$_2$ and FCO$_2$ would be expected to be reflected in the CMIP5 models as well. A quantitative analysis of the correlation of the phasing of the seasonal cycle of FCO$_2$ between basins for different models shows that all the models except 3 (CMCC-CESM, GFDL-ESM2M CESM1-CESM) are characterized by strong inter-basin correlation in both the SAZ and the AZ (Fig. 4). This suggests that the carbon cycle in these CMIP5 models is not sensitive to inter-basin differences in the drivers as is the case for observations.

Secondly, an important part of this analysis is based on the assumption that the observational products that are used to constrain the spatial and temporal variability of pCO$_2$ and FCO$_2$ reflect the correct seasonal modes of the Southern Ocean. This assumption requires significant caution not only due to the limitations in the sparseness of the in situ observations but also due to limitations of the empirical techniques in overcoming these data gaps (Landschutzer et al., 2014; Rödenbeck et al., 2015; Gregor et al., 2017a,b; Ritter et al., 2018). The uncertainty analysis from these studies suggests that, while the seasonal bias in observations may be less in the SAZ and PFZ, it is the highest in the AZ where access is limited mostly to summer, and winter ice cover result in uncertainties that may limit the significance of the data - model comparisons. It is important to note that though the observation product we use here (Landschutzer et al., 2014) is based on more surface measurement (10 millions, SOCAT v3) compared to previous datasets (e.g. Tahakahashi et al., 2009, 3 millions), the data are still sparse in time and space in the Southern Ocean. Thus using this data product as our main observational estimates for this analysis we are mindful of the limitations in its discussion below.

Thirdly, the seasonal cycle of ΔpCO$_2$ is the dominant mode of variability in FCO$_2$ (Mongwe et al., 2016; Wanninkhof et al., 2009). Though winds provide the kinematic forcing for air-sea fluxes of CO$_2$ and indirectly affect FCO$_2$ through mixed layer dynamics and associated biogeochemical responses (Mahadevan et al., 2012; du Plessis et al., 2017), ΔpCO$_2$ sets the direction of the flux. Surface pCO$_2$ changes are mainly driven by DIC and SST (Hauck et al., 2015; Takahashi et al., 1993). Subsequently the sensitivity of CMIP5 models to how changes in DIC and SST regulates seasonal cycle of FCO$_2$ is fundamental to the model’s
ability to resolve observed FCO$_2$ seasonal cycle. Thus here we examined the influence of DIC and SST on
FCO$_2$ at seasonal scale for 10 CMIP5 models with respect to observed estimates. But because temperature
does not directly affects DIC changes, we first scaled up the impact of SST changes on pCO$_2$ through surface
CO$_2$ solubility to equivalent DIC units using the Revelle factor (section 2.3). In this way we can distinguish
the influence of surface solubility and DIC changes (i.e. biological and physical) on pCO$_2$ and hence then
FCO$_2$.

Fourthly, using this analysis framework (sec 2.3, summarized in Fig. 6) we found that CMIP5 models FCO$_2$
biases cluster in two groups, namely group-DIC ($M_{T,DIC}<0$) and group-SST ($M_{T,DIC}>0$). Group-DIC models are
characterized by an overestimation of the influence of DIC on pCO$_2$ with respect to observations estimates,
which instead indicate that physical and biogeochemical changes in the DIC concentration mostly regulate
the seasonal cycle of FCO$_2$ (in short, DIC control). Group-SST models show an excessive temperature
influence on pCO$_2$; here surface CO$_2$ solubility biases are mainly responsible for the departure of modeled
FCO$_2$ from the observational products. While CMIP5 models mostly show a singular dominant influence of
these extremes, observations show a modest influence of both, with a dominance of DIC changes as the
main driver of seasonal FCO$_2$ variability. Below we discuss the seasonal cycle characteristics and possible
mechanisms for these two groups of CMIP5 models in the Sub-Antarctic and Antarctic Zones of the
Southern Ocean.

4.1 Sub-Antarctic Zone (SAZ)

Our diagnostic analysis indicates that the seasonal cycle of pCO$_2$ in the observational product (Landschützer
et al., 2014) is mostly DIC controlled across all three basins of the SAZ ($M_{T,DIC}<0$ in Fig. 6). The Atlantic
Ocean shows a stronger DIC control (Annual mean $M_{T,DIC} \geq 2$) compared to the Pacific and Indian Ocean
(Annual mean $M_{T,DIC} \approx 1$). This stronger influence of DIC on pCO$_2$ in the Atlantic Ocean is consistent with
higher primary production in this basin (Graham et al., 2015; Thomalla et al., 2011), here shown by the
larger mean seasonal chlorophyll from remote sensing in the Atlantic basin with respect to the Pacific and
Indian basin (Fig. 8). This significant basin difference is most likely linked to a number of factors: the
Atlantic basin has longer periods of shallow MLD compared to the Pacific and Indian basins (Fig. 7a-c, Nov –
Mar & Nov - Feb respectively) and has been shown to have higher supplies of continental shelves and land
based iron (Boyd and Ellwood, 2010; Tagliabue et al., 2012; 2014). These conditions are more likely to
enhance primary production that translates into a higher rate of change of surface DIC (Fig. 8), which
becomes the major driver of FCO$_2$ variability. In contrast, shorter periods of shallow MLD and lower iron
inputs in the Pacific Ocean (Tagliabue et al., 2012), likely account for lower chlorophyll biomass and hence
the weaker DIC control evidenced in our analysis ($M_{\text{TR-DIC}} \approx 0$ in Fig. 6). In the Indian Ocean, the winter mixed layer is deeper than in the Atlantic and deepens earlier in the season (Fig. 7c). These conditions limit chlorophyll concentration (Fig. 8) and possibly contribute to the lower rates of surface temperature change because of the enhanced mixing (cf Fig. 5a-c). As a consequence the resulting net driver in the Indian and Pacific basins is a weaker DIC control, because both biological DIC and solubility changes are relatively weaker and they oppose each other. Because of this, when the magnitudes of the rate of change of SST are larger during cooling and warming seasonal peaks (autumn and spring respectively), DIC control is weaker ($M_{\text{TR-DIC}} \approx 0$) during these seasons.

CMIP5 models do not capture these basin-specific features as demonstrated with the correlation analysis in Fig. 4, with the exception of three group-SST models (i.e. CESM1-BGC, GFDL-ESM2M and CMCC-CESM). These, in contrast, mostly show comparable FCO$_2$ phasing in the three basins. This spatial uniformity of CMIP5 models is both zonal and meridional for most models in the Southern Ocean (Fig. 3, 4), which is in contrast to observation products (Fig. 3). This suggests that CMIP5 models show equal sensitivity to basin scale FCO$_2$ drivers, suggesting that pCO$_2$ and FCO$_2$ driving mechanisms are less local than for observations.

The major feature of group-SST models in the SAZ is the outgassing during summer and ingassing in winter (Fig. 3a-c, Dec-Feb), which our diagnostics in Fig. 6 attribute to temperature (solubility) control. The summer period coincides with the highest warming rates ($dSST/dt$, Fig 5a-c), and associated reduction in solubility of CO$_2$. Similarly, exaggerated cooling rates at the onset of autumn (Fig. 5a-c) enhance CO$_2$ solubility causing a change in the direction of FCO$_2$ into strengthening CO$_2$ ingassing (Fig 3a-c). Thus, while group-SST models have seasonal amplitude of FCO$_2$ comparable to observations, they are out of phase (Fig. 3) as was the case in a previous analysis of a forced ocean model (Mongwe et al., 2016).

In addition to increasing CO$_2$ solubility, the rapid cooling at the onset of autumn also deepens the MLD (March-June, Fig. 7), which induces entrainment of DIC, increasing surface CO$_2$ concentration and weakening the ocean-atmosphere gradient and, in some instances, reversing the air-sea flux to outgassing (Lenton et al., 2013a; Mahadevan et al., 2011; Metzl et al., 2006). While these processes (cooling and DIC entrainment) are likely to co-occur in the Southern Ocean, in CMIP5 models they are characterized by their extremes: temperature impact of solubility exceeds the rate of entrainment (Fig. 6 & 10). Because of the dominance of the solubility effect in group-SST models, the impact of DIC entrainment on surface pCO$_2$ changes, the weakening of CO$_2$ ingassing / outgassing only happens in mid-late winter (June-July-August) when entrainment fluxes peak (Fig. 10) and the SST rate approaches zero (Fig. 5).
In the spring-summer transition, primary production is anticipated to enhance the net CO₂ uptake (Thomalla et al., 2011; Le Quéré and Saltzman, 2013). However, the elevated surface warming rates during spring reduces CO₂ solubility in group-SST models and overwhelms the role of primary production in the seasonal cycle of pCO₂ and FCO₂ (atmospheric CO₂ uptake). As a consequence, these group-SST models mostly show a constant or weakening net CO₂ uptake flux during spring in the Pacific and Atlantic Ocean even though primary production is occurring and is relatively elevated (Fig. 3 & 8). Though some models show chlorophyll concentrations comparable to observations (e.g. GFDL-ESM2M, CNRM-CMS, CanESM2), and sometimes greater (e.g. MRI-ESM), the impact of temperature driven solubility dominates due the phasing of the rates of the two drivers (Fig. 2a-c). The Indian Ocean however shows the only exception to this phenomenon. Here, the amplitude of the seasonal surface warming is relatively smaller (~ 0.5 °C⁻¹ month⁻¹ lower than the Pacific and Atlantic basins), and the biologically driven CO₂ uptake becomes notable and show a net strengthening of the sink of CO₂ during spring (Fig. 3c).

Though almost all analysed CMIP5 models (with the exception of NorESM1-ME) exaggerate the warming and cooling rates in autumn and spring, group-DIC models do not manifest the expected temperature-driven solubility impact on pCO₂ and FCO₂ (Fig. 2). Instead, the seasonal cycle of pCO₂ and FCO₂ are controlled by DIC changes. However, this is driven by an overestimated seasonal primary production and the associated carbon export fluxes (Fig. 8). It is striking how in these models the seasonal cycle of chlorophyll and FCO₂ are in phase (Fig 3a-c, 8a-c, with linear correlation coefficients always larger than 0.9, not shown) but, as we discuss below, this is not because the temperature rates of change are correctly scaled but because the biogeochemical process rates are exaggerated (Fig. 8).

Because of the particularly enhanced production in group-DIC models, the CO₂ sink is stronger (Fig. 8) with respect to observation estimates during spring. This is visible in the reduction of surface DIC (negative dDIC/dt in Fig. 8a, g-i), which can only be explained by drawdown due to the formation and export of organic matter (Le Quéré and Saltzman, 2013). However, note that in the same way, after the December production peak, both CMIP5 models and observations show an increase of surface DIC concentrations (positive dDIC/dt) until March (Fig. 8, g-i). These DIC growth rates are particularly enhanced in group-DIC models compared to some group-SST and observations (Fig. S9). The onset of these DIC increases also coincides with the depletion of surface oxygen (Fig. S9), which we makes us speculate that this is due to the remineralisation of organic matter to DIC through respiration. Unfortunately, only a few models have stored the respiration rates, therefore the ultimate reason for this DIC rebound remains to be examined at a later stage. We would however tend to exclude other processes, because the onset of CO₂ outgassing
seen in March in group-DIC models occurs prior to significant MLD deepening (Fig. 7) and entrainment fluxes, therefore remineralization is likely be a key process here (Fig. 8).

### 4.2 Antarctic Zone (AZ)

The seasonal cycle framework summarized in Fig. 6 shows that the variability of FCO$_2$ and pCO$_2$ in the Landschützer et al. (2014) product is characterized by a stronger DIC control (annual mean $M_{T,DIC} < -2$) relative to the Sub-Antarctic ($M_{T,DIC} = -1$), except in the spring season ($M_{T,DIC} > -1$). This DIC control is spatially uniform in the Antarctic zone across all three basins (Fig. 4). The available datasets indicate that the combination of weaker SST rates due to lower solar heating fluxes (Fig. 5), and stronger shallower vertical DIC maxima (Fig. 10) favour a stronger DIC control through larger surface DIC rates. The spatial uniformity in the seasonality of FCO$_2$ is also evident in the satellite chlorophyll and calculated dDIC/dt from GLODAP2 in Fig. 9. Contrary to the Sub-Antarctic this might be suggesting that FCO$_2$ mechanisms are here less local. It could be hypothesized that the seasonal extent of sea-ice, deeper mixing and heat balance differences affect this region more uniformly compared to the Sub-Antarctic zone, and hence the mechanisms of FCO$_2$ are spatially homogeneous. However, we cannot forget that sparseness of observations in this region is a known key limitation to data products (Bakker et al., 2014; Gregor et al., 2017; Monteiro et al., 2010; Rödenbeck et al., 2013) that might hamper the emergence of basin specific features. Consequently, this highlights the importance and need to prioritize independent observations in the Southern Ocean south of the polar front and in the Marginal Ice Zone. Increased observational efforts should also include a variety of platforms such as autonomous vehicles like gliders (Monteiro et al., 2015) and biogeochemical floats (Johnson et al., 2017) in addition to ongoing ship-based measurements.

In general terms, CMIP5 models are mostly in agreement (with an exception of MRI-ESM) with the observational product on the dominant role of DIC to regulating the seasonal cycle of FCO$_2$ (Fig. 6d-f), though not all models agree in the phase of the seasonal cycle of FCO$_2$ (e.g. CanESM2, Fig. 2). Though CMIP5 models still mostly show the SST rates biases in autumn and spring with respect to observed estimates, the stronger and near surface vertical DIC maxima (Fig. 10), likely favor DIC as a dominant driver of FCO$_2$ changes. Differences between group-SST and group-DIC models are only evident in mid-summer when SST rates heighten and primary production peaks (Fig. 3 & 9). Probably because of sea ice presence, the onset of SST warming is a month later (November) here in comparison to the Sub-Antarctic (October). This subsequently allows the onset of primary production before the surface warming, which then permits the biological CO$_2$ uptake to be notable in group-SST models. We notice here that the reason why CMIP5 models develop a winter bloom in the AZ requires further investigation (Hague and Vichi, submitted). Thus
the two model groups here agree in the FCO₂ ingassing during spring with group-SST models being the closest to the observational product. The MRI-ESM is the only model showing anomalous solubility dominance during autumn and spring as in the Sub-Antarctic zone.

This coherence of CMIP5 models and observations in the Antarctic zone, may suggest that CMIP5 models compare better to observations in this region (Fig. 4). However, because CMIP5 models also show this spatial homogeneity in the Sub-Antarctic Zone (contrary to observational estimates), it not clear whether this indicates an improved skill in CMIP5 model to the mechanisms of FCO₂ in this region, or both CMIP5 models and observational product lacks spatial sensitivity to the drivers of FCO₂. The sparseness of observations in the AZ points to the latter.

5. Synthesis

We used a seasonal cycle framework to highlight and examine two major biases in respect of pCO₂ and FCO₂ in 10 CMIP5 models in the Southern Ocean.

Firstly, the general exaggeration of the seasonal rates of change of SST in autumn and spring seasons during peak cooling and warming respectively with respect to available observations. These elevated rates of SST change tip the control of the seasonal cycle of pCO₂ and FCO₂ towards SST from DIC and result in a divergence between the observed and modelled seasonal cycles, particularly in the Sub-Antarctic Zone.

While almost all analysed models (9 of 10) show these SST-driven biases, 3 of the 10 (namely NorESM1-ME, HadGEM-ES and MPI-ESM) don’t show these solubility biases because of their overly exaggerated primary production (and remineralization) rates such that biologically driven DIC changes mainly regulate the seasonal cycle of FCO₂. These models reproduce the observed phasing of FCO₂ as a result of an incorrect scaling of the biogeochemical fluxes. In the Antarctic zone, CMIP5 models compare better with observations relative to the Sub-Antarctic Zone. This is mostly because both CMIP5 models and observational product estimates show a spatial and temporal uniformity in the characteristics of FCO₂ in the Antarctic zone. However, it is not certain if this is because model process dynamics perform better in this high latitude zone or that the observational products variability is itself limited by the lack of in situ data. This remains an open question that needs to be explored further and highlights the need for increased scale sensitive and independent observations south of the Polar Front and into the sea ice zone.

The second major bias is that contrary to observational products estimates, CMIP5 models generally show an equal sensitivity to basin scale FCO₂ drivers (except for CMCC-ESM, GFDL-ESM2M and CESM1-BGC) and
hence the seasonal cycle of FCO$_2$ has similar phasing in all three basins of the Sub-Antarctic zone. This is in contrast to observational and remote sensing products that highlight strong seasonal and interannually varying basin contrasts in both pCO$_2$ and phytoplankton biomass. It is not clear if this is due to inadequate carbon process parameterization or gaps in the dynamics of the physics. This should be investigated further with CMIP6 models and our analysis framework is proposed as a useful tool to diagnose the dominant drivers. Contrary to observed estimates, CMIP5 models simulate FCO$_2$ seasonal dynamics that are zonally homogeneous and for this reason it is suggested that any investigation of local (basin scale) mechanisms, dynamics and long term trends of FCO$_2$ using CMIP5 models should be cautious. This highlights a key area of development for CMIP6.

Acknowledgements

This work was undertaken with financial support from the following South African institutions: CSIR Parliamentary Grant, National Research Foundation (NRF SANAP programme), Department of Science and Technology South Africa (DST), and the Applied Centre for Climate and Earth Systems Science (ACCESS). We thank the CSIR Centre for High Performance Computing (CHPC) for providing the resources for doing this analysis. We also want to thank Peter Landschützer, Taro Takahashi and Luke Gregor for making their data products available as well as the three reviewers for their productive comments that we think have strengthened the paper.

References


Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H. and Núñez-Riboni, I.: Global ocean


Matear, R. J. and Lenton, A.: Impact of Historical Climate Change on the Southern Ocean Carbon Cycle, J.


Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., Hales, B.


**Figures**

![Figure 1: Annual climatological Sea-Air CO₂ Flux (FCO₂, in gC m⁻² yr⁻¹) for observations (L14:Landschützer et al., 2014 and T09: Takahashi et al., 2009) and 10 CMIP5 models over 1995 – 2005.](image)
**Fig. 2**: Seasonal cycle of Sea-Air CO$_2$ Flux (FCO$_2$, in gC m$^{-2}$ yr$^{-1}$) in observations and 10 CMIP5 models in the Sub-Antarctic and Antarctic zones of the Pacific Ocean (first column), Atlantic Ocean (second column) and Indian Ocean (third column). The shaded area shows the temporal standard deviation over the considered period (1995 – 2005).
Fig. 3. Seasonal cycle of the equally weighted ensemble means of FCO₂ (gC m⁻² yr⁻¹) from Fig. 2 for group DIC models (MPI-ESM, HadGEM-ES and NorESM), and group SST models (GFDL-ESM2M, CMCC-CESM, CNRM-CERFACS, IPSL-CM5A-MR, CESM1-BGC, NorESM2, MRI-ESM and CanESM2). The shaded areas show the ensemble standard deviation. The black line is the Landschützer et al., (2014) observations.
Fig. 4: The correlation coefficients (R) of basin–basin seasonal cycles of FCO$_2$ for observations (Landschützer et al., 2014) and 10 CMIP5 models in the three basins of the Southern Ocean i.e. Pacific, Atlantic and Indian basin.
Fig. 5: Mean seasonal cycle of the estimated rate of change of sea surface temperature (dSST/dt, °C month⁻¹) for the Sub-Antarctic and Antarctic zones of the Pacific Ocean (first column), Atlantic Ocean (second column) and Indian Ocean (third column).
Fig. 6: Mean seasonal and annual values of the DIC–temperature control index ($M_{1;DIC}$). The increase in the red color intensity indicates increase in the strength of the temperature driver and the blue intensity shows the strength of the DIC driver. The models are sorted according to the annual mean value of the indicator presented in the last column ($A_{mean}$)
Fig. 7: Seasonal cycle of the Mixed Layer Depth (MLD) in the Sub-Antarctic and Antarctic zones of the Pacific Ocean (first column), Atlantic Ocean (second column) and Indian Ocean (third column).
Fig. 8: The seasonal cycle of chlorophyll (mg m$^{-3}$), Net Primary Production (mmol m$^{-2}$ s$^{-1}$) and the surface rate of change of DIC (μmol kg$^{-1}$ month$^{-1}$) in the Sub-Antarctic zone of the Pacific Ocean (first column), Atlantic Ocean (second column) and Indian Ocean (third column).
Fig. 9 Same as Fig. 8 for the Antarctic zone.
**Fig. 10:** (a-f) Estimated DIC entrainment fluxes (mol kg month$^{-1}$) at the base of the mixed layer and (g-i) vertical DIC gradients (μmol kg$^{-1}$ m$^{-1}$) in the Sub-Antarctic and Antarctic zone of the Pacific Ocean (first column), Atlantic Ocean (second column) and Indian Ocean (third column).
Table 2: Sea-Air CO₂ fluxes (Pg C yr⁻¹) annual mean uptake in the Southern Ocean (first column), here defined as south of the Sub-tropical front, Sub-Antarctic zone (second column) and Antarctic zone (third column). The third and forth column shows the Pattern Correlation Coefficient (PCC) and Root Mean Square Error (RMSE) for the whole Southern Ocean for each model.
<table>
<thead>
<tr>
<th>Model</th>
<th>Southern Ocean</th>
<th>Sub-Antarctic zone</th>
<th>Antarctic zone</th>
<th>PCC</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNRM-CM5</td>
<td>-0.823 ± 0.003</td>
<td>-0.682 ± 0.002</td>
<td>-0.122 ± 0.001</td>
<td>0.44</td>
<td>17.9</td>
</tr>
<tr>
<td>GFDL-ESM2M</td>
<td>-0.161 ± 0.005</td>
<td>-0.074 ± 0.004</td>
<td>-0.077 ± 0.002</td>
<td>0.43</td>
<td>8.47</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>-0.489 ± 0.005</td>
<td>-0.284 ± 0.003</td>
<td>-0.197 ± 0.001</td>
<td>0.55</td>
<td>10.9</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>-0.496 ± 0.003</td>
<td>-0.582 ± 0.006</td>
<td>0.101 ± 0.003</td>
<td>0.53</td>
<td>10.5</td>
</tr>
<tr>
<td>MPI-ESM-MR</td>
<td>-0.870 ± 0.006</td>
<td>-0.530 ± 0.002</td>
<td>-0.326 ± 0.002</td>
<td>0.37</td>
<td>9.87</td>
</tr>
<tr>
<td>MRI-ESM</td>
<td>-0.048 ± 0.002</td>
<td>0.022 ± 0.003</td>
<td>-0.070 ± 0.001</td>
<td>0.36</td>
<td>15.6</td>
</tr>
<tr>
<td>NorESM1</td>
<td>-0.699 ± 0.004</td>
<td>-0.412 ± 0.003</td>
<td>-0.270 ± 0.002</td>
<td>0.60</td>
<td>8.96</td>
</tr>
<tr>
<td>CESM1-BGC</td>
<td>-0.532 ± 0.006</td>
<td>-0.132 ± 0.003</td>
<td>-0.385 ± 0.004</td>
<td>0.47</td>
<td>9.15</td>
</tr>
<tr>
<td>CMCC-CESM</td>
<td>0.121 ± 0.006</td>
<td>0.367 ± 0.004</td>
<td>-0.225 ± 0.003</td>
<td>-0.09</td>
<td>17.9</td>
</tr>
<tr>
<td>CanESM2</td>
<td>-0.058 ± 0.008</td>
<td>-0.720 ± 0.006</td>
<td>0.661 ± 0.004</td>
<td>0.54</td>
<td>19.5</td>
</tr>
<tr>
<td>Observations</td>
<td>-0.253 ± 0.3</td>
<td>-0.296 ± 0.3</td>
<td>0.053 ± 0.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>