We thank Referee#2 for his/her positive and constructive feedback on our study. In the revised manuscript, we have addressed most or all of the comments raised and we think that the manuscript has been significantly improved since. Below, please find point-by-point response to each of Referee#2 comment. They are constructed as follows (1) original comments from the referee in **bold**, (2) our response in *italics*, and (3) description of changes applied in the revised manuscript in blue.

Referee#2: My main concern is on the use of a positive inter-model correlation between present day and future behaviour of models as an indicator of the Southern Ocean as a constrain to reduce future uncertainty. Here's what I don't get: all models are wrong (G1/G2 too strong/little CO2 uptake) and for those that are less wrong (G2) it is for the wrong reason (opposite seasonal cycle). The link between present and future behaviour is not evidence of models becoming right in the future. It does not give more credibility to the projected sustained growth of CO2 uptake in the SO because this growth is still a result based on present-day biases (the authors show that these biases persist in the future). The sustained CO2 uptake growth in the SO is the reason for which this region is selected for the analysis (and because of the significance of the correlation)

Response: The Referee#2 is correct in that both G1/G2 models simulate relatively wrong seasonality in the CO_2 uptake, either in the seasonal phase (G2) or amplitude (G1). As the referee pointed out, we found it encouraging that all models consistently show increasing CO_2 uptake for the SO region, and not in other regions, hence the strong inter-model relationships. As also shown is the study, these discrepancies in amplitude and/or phase are due to the respective simulated SST and NPP seasonal cycles. To address this comment, we did literature research to find which bio-physical processes is potentially most important in determining the correct long-term CO_2 fluxes in the SO. We found that the non-thermal component of the pCO_2 variability is the dominant determining factor (Figs. 2d and 2f of Landschützer et al., 2015). And based on this we have added a new figure 10 (see below) showing the anomaly of non-thermal pCO_2 seasonal cycle (i.e., the sum of ALK and DIC components). It shows that there are two models, CanESM2 and GFDL, which simulate comparable seasonal phase and amplitude with the observations. The GFDL model coincidently simulates contemporary annual carbon uptake that is very close to the observation-based estimate (see also Fig. 4a in the revised manuscript). On the other hand the CanESM2 simulates too strong oceanic pCO_2 compare to the atmospheric pCO_2, which explains the simulated outgassing.

Revision: In the revised manuscript, we have added the new figure (Fig. 10) and add a new paragraph in the discussion section (4th paragraph) to address the referee#2 concern as follows:
"Based on the linear inter-model relationship presented in this study, the GFDL-ESM2G, MIROC-ESM, and HadGEM2-ES models simulate contemporary CO_2 fluxes in the SO closest to the observational-based estimate (see for example Fig. 5a), and therefore are likely to have more credibility in their future projections. Nevertheless, from our seasonal cycle analysis it is not clear if these models simulate the observed mechanisms governing the CO_2 fluxes. According to Landschutzer et al. (2015), the non-thermal component of the pCO_2 variation is an important driver for the long-term CO_2 fluxes in the SO. Figure 10 shows the seasonal anomaly of non-thermal pCO_2 seasonal cycle in the SO from models and observation-based estimate. The CanESM2 and GFDL-ESM2G simulate comparable amplitude and seasonal phase with the observation-based estimate, but the former model has anomalously high surface pCO_2 (i.e., it simulates a net source of CO_2 to the atmosphere in the SO). Taking this as an additional constrain, our analysis suggests that the GFDL-ESM2G performs best in capturing the observed CO_2 fluxes in the Southern Ocean."

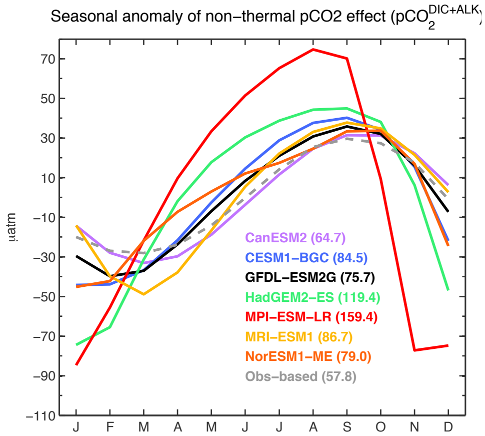

Figure 10. Anomalies of non-thermal pCO_2 seasonal cycle, ($pCO_2^{\{DIC+ALK\}}$) as simulated by seven ESMs for the 2001-2010 period. The gray dashed line indicates the observation-based estimate of $pCO_2^{\{DIC+ALK\}}$ seasonal cycle. The numbers within the parentheses represent the amplitude for each model.

Referee#2: Said so, I am still convinced that the SO indeed is a constrain to improve future projections but I am not sure the inter-model correlation is evidence of it. Perhaps a more explicit explanation on the meaning of this correlation could help.

Response: We agree that explicit statement on the meaning of this correlation including its limitation would be useful to avoid misunderstanding from the readers.

Revision: In the abstract, we have added the following statement:
"This strong correlation suggests that models with low carbon uptake rate in the contemporary SO tend to simulate low uptake rate in the future and vice versa. Nevertheless, our analysis also shows that none of the models fully capture the observed bio-physical mechanisms governing the CO_2 fluxes in the SO."

Referee#2: The use of 45 S as a limit for the Southern Ocean is likely to cut out in some places, depending on the model, part of the region of high CO2 uptake associated with the winter deepening of the mixed layer and the formation of subantarctic mode water (e.g. Sallee et al., 2012). I wonder how sensitive are results on this limit and whether a more dynamic limit based, for example, on outcrop surfaces of isopycnals for SAMW or using Ekman divergence as a separation between Antarctic and subantarctic zones could change results in any way. Have the authors carried out any complementary analysis on this issue? Another choice for the SO limit could shed some light on the negative (although not significant) correlation for the mid-latitude SO. As it is now, this region includes part of the subtropical gyre and part of the deep winter mixer layer area forced by the westerlies. These are likely to evolve in opposite ways in the future with a strengthening of the westerlies due to the increase in the meridional temperature gradient, as stated by the authors at page 2659 (lines 8-13). Stronger winds could enhance intermediate water formation but also SAMW formation which is split between the two regions in the current separation. If further analysis is not possible I suggest at least an expanded discussion on this aspect.

Response: We thank the referee for this suggestion. We have not performed any complementary analysis, specifically using dynamic boundaries in our analysis. The issue pertaining the use of 45°S as boundary is also raised by referee#1. As pointed out the boundary of the uptake region in the SO is *model dependent. Our motivation for selecting 45°S in the SO was for simplicity and following the SO separation presented in Mikaloff-Fletcher et al. (2007). To address this comment, we have followed the referee suggestion to perform a new analysis applying a dynamic boundary. For this we have selected boundary that separates surface water density greater and less than sigma 26.5 kg m^{-3}, which is a density that separate the SAMW and Subtropical Mode Water (TMW) following in Séférian et al. (2012). In this analysis, the boundaries were computed monthly for each individual model. Our new analysis is consistent with and supports our earlier result (i.e., similar inter-model correlation, see also new Fig. 5 below). We think that this is important and have decided to add this (also the new Fig. 5) to the revised manuscript.*

Revision: The following new statements have been added in the end of subsection 2.6:
"We note that the selection of the 45°S as boundary between the mid- and high-latitude SO, could pose issues since the SO region has a sophisticate dynamics and, dependent on the models, the 45°S latitude could cut into regions of dominant carbon sources or sinks. To address this issue, we perform additional analysis where we use a dynamic boundary separating the mid- and high-latitude Southern Ocean applying a surface density of 26.5 kg m^{-3}. For instance, Séférian et al. (2012) apply this density line to separate the Subtropical Mode Water (TMW, region of weak increase in future CO_2 uptake) and the Subantarctic Model Water (MW, region of strong increase in future CO_2 uptake)."

In addition, we have also added a new Fig. 5 illustrating these dynamic boundaries as simulated by the models as the corresponding inter-model relationships when this boundary is used. The following paragraph has been added to section 3.1.
"As stated in subsection 2.6, we also compute the correlation coefficient metrics for the SO region using a dynamic boundary (instead of a fixed 45°S latitude). Figure 5c illustrate the model-dependent dynamic boundaries as simulated for August 2005. Figure 5a and b show that the linear inter-model relationships remain strong (correlation coefficient of at least 0.76) when the dynamic boundary is used, suggesting that the inter-model relationships in the SO is relatively robust."

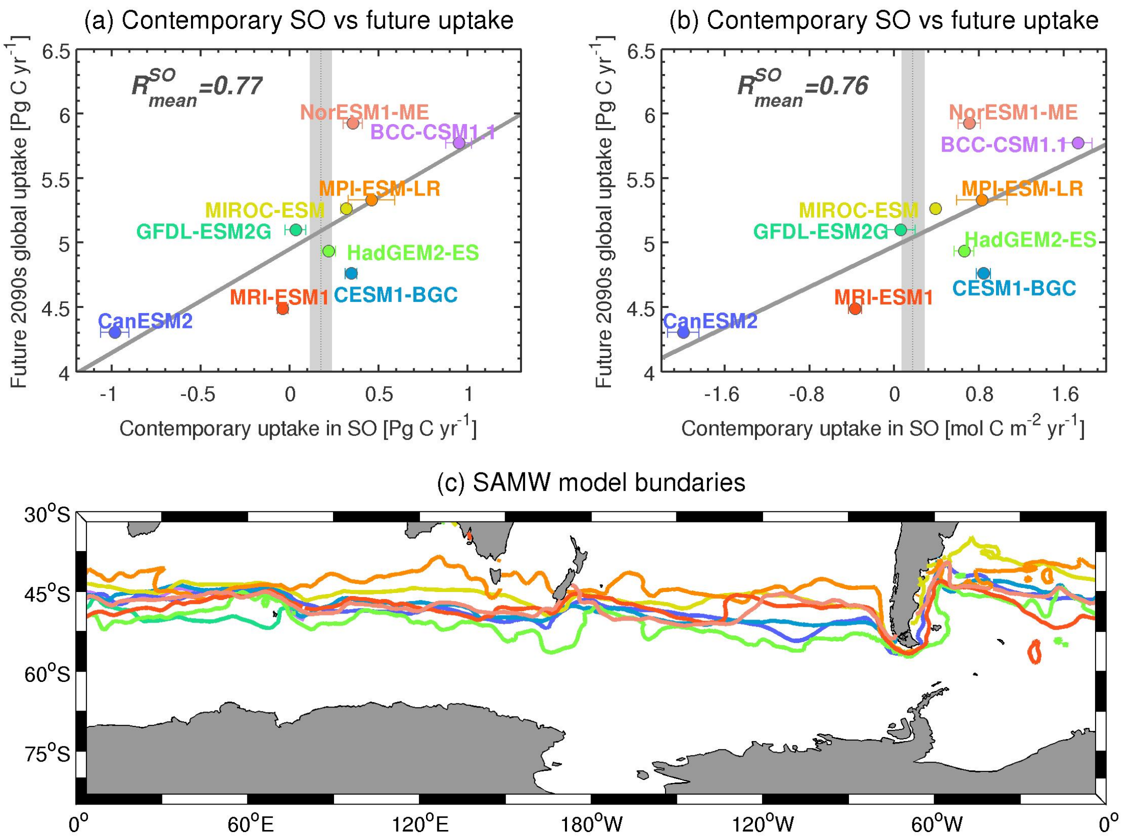

Figure 5. Annual contemporary carbon uptake vs. global uptake rate projected in the last decade of the 21st century by CMIP5 models. Here the SO is defined using dynamic boundary separated by the surface water density of 26.5 kg m^{-3}. Panels (a) and (b) show the contemporary SO carbon uptake on the x-axes in Pg C yr^{-1} and mol C m^{-2} yr^{-1} units, respectively. Panel (c) illustrates the 26.5 kg m^{-3} density lines that separate the MW from the TMW for the month of August 2005 as simulated by the different models (same color convention as in panels a and b).

Referee#2: Also, likely less important but still interesting is the uptake of CO2 due to the overestimated open sea convection in the SO. Most CMIP5 models form AABW through unrealistic extended open sea convection in the subpolar SO (Heuze et al., 2013). This is mostly because of still too-coarse resolution and thus the difficulty to resolve the complex formation processes occurring on the continental shelf. Convection regime is however, very variable across models and so it is its response to climate change, with a general reduction of convection area and duration but with large variability of the timing across models (deLavergne et al., 2014). The impact of the reduction and shutdown of convective area on the uptake of anthropogenic carbon can be important, specially when considered in terms of its contribution to the total SO CO2 uptake trend (Bernardello et al., 2014). The authors mention the importance of deep winter mixing in polar regions as an efficient way to transport anthropogenic carbon from surface to depth (page 2659, Lines 6-8). In light of the above I wonder if perhaps considering mixed layer depth, in addition to SST and NPP, could give new insights on the processes involved in determining the inter-model differences in CO2 uptake.

Response: We agree that analyzing the mixed layer depth could add new insights on the uncertainty associated with the physical processes in the model. However, when we did the analysis one year ago, we discovered that only two models provide the field monthly average mixed layer depth under the fully *interactive esmRCP8.5 experiments. Nevertheless, six models provide the field maximum mixed layer thickness as shown in Fig. R1 below (shown are seasonal fields for the contemporary period). In general, we did not find clear distinct pattern between G1 and G2 models that would fit our analysis, nor did we find any significant changes in the 21st century. We decided not to include this in the paper. But following the referee#2 suggestion, we have added a paragraph discussing the uncertainty the projected CO_2 uptake related to the caveat in model convective processes.*

Revision: The following paragraph has been added into the revised manuscript (Section 4, paragraph 6):
"In the SO, the CO_2 flux and its evolution in response to climate change also depend critically on the spatial and temporal variation of convection processes (e.g., Sallée et al., 2012). Due to the coarse spatial resolution in CMIP5 models, convection processes along the continental margin that form the AABW (Antarctic Bottom Water) are not well reproduced (Heuzé et al., 2013). Similarly, Bernadello et al. (2014) suggests that the anthropogenic CO_2 uptake in the Weddell Sea is closely linked to the size and timing of deep-water convection. It remains to be investigated how these uncertainties contribute to the inter-model spread of the projected CO_2 uptake in the SO shown here, especially with the next round of CMIP6, which includes models with higher resolution."

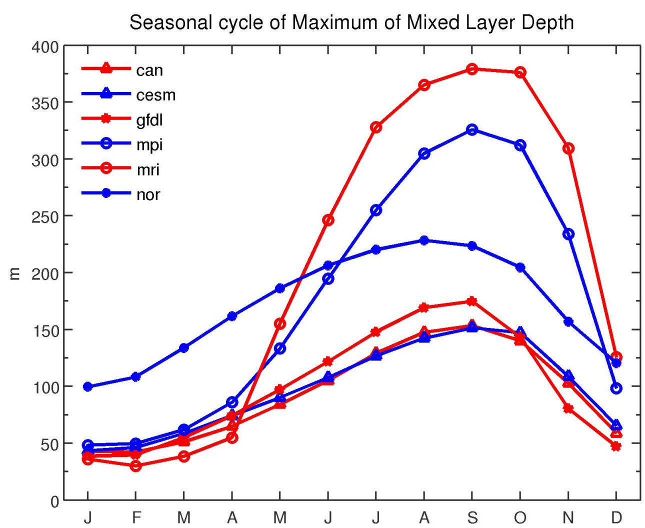

Figure R1. Monthly maximum mixed layer thickness in the SO (south of 45°S) as simulated by six CMIP5 models for the 2006-2010 period under the esmRCP8.5 experiment.

Referee#2: It's not explained why only fully-interactive simulations are considered. The same processes responsible for the seasonal pCO2 cycle biases described should be active also in simulations with prescribed atmospheric CO2. If so, maybe more models would be available. Is there a motivation behind this choice?

Response: We chose the fully interactive simulations since they include all the associated changes in atmospheric CO_2 with the evolving oceanic uptake, allowing for a more realistic spatially varying atmospheric CO_2 concentration. Hence, we think that the 'fully-interactive' simulations are more representative of the real world than the non-interactive, with prescribed atmospheric CO_2 concentration, simulations. The referee is correct that similar processes should be responsible for the seasonal pCO_2 cycle. Nevertheless, in the earlier stage of our analysis, we have also looked into the non-interactive simulations (seven models), but we only found a weak relationship, and there is no *obvious reason for this discrepancy. To keep our analysis straight forward, we decided to only analyze the fully interactive simulations.*

Revision: In the method section (2.3, paragraph 1), we have revised the sentence:
"These 'esm' simulations take into account carbon fluxes between the land-atmosphere and ocean-atmosphere interfaces to prognostically simulate the atmospheric CO_2 concentration."
with
"These 'esm' simulations take into account carbon fluxes between the land-atmosphere and ocean-atmosphere interfaces to prognostically simulate the atmospheric CO_2 concentration, thus they include more realistic spatially varying atmospheric CO_2 concentration."

Referee#2: Put in Figure 6 panel titles also G1 and G2 to facilitate the comprehension of the Figure.

Response: We have added "G1" and "G2" into the figure legend.

Revision: The figure (now Figure 7 in the latest manuscript) has been revised as suggested.

Referee#2: Page 2655, Lines 19-22: The numbers given here for the highest uptake estimates do not seem to coincide with the values in Figure 1. For example, NorESM1-ME in the plot does not go below 2.4 Pg C/yr so the average uptake for the period 2001-2010 should be higher than that.

Response: We thank referee#2 for noticing this error. The values in Fig. 1 are correct. But there was a small bug in our script that computes the CO_2 uptake values used in the text. The values have been updated and are now consistent with Fig. 1.

Revision: Following the correction in our script, we recomputed some of the quantities in the manuscript. Thus there are minor changes in the numbers, which do not change the main finding of our study.