AGU PUBLICATIONS

Geophysical Research Letters

RESEARCH LETTER
10.1002/2015GL067584

Key Points:
- Seasonal cycles in atmospheric potential oxygen (APO) are useful for evaluating ocean models
- CMIP5 models show clear differences in their skill at reproducing observed APO seasonal cycles
- CMIP5 models that perform poorly on APO tend to project a larger future Southern Ocean carbon sink

Supporting Information:
- Figures S1–S4

Correspondence to:
C. D. Nevison,
cynthia.nevison@colorado.edu

Citation:
Nevison, C. D., et al. (2016), Evaluating CMIP5 ocean biogeochemistry and Southern Ocean carbon uptake using atmospheric potential oxygen: Present-day performance and future projection, *Geophys. Res. Lett., 43*, 2077–2085, doi:10.1002/2015GL067584.

Received 28 DEC 2015
Accepted 15 FEB 2016
Accepted article online 18 FEB 2016
Published online 5 MAR 2016

Evaluating CMIP5 ocean biogeochemistry and Southern Ocean carbon uptake using atmospheric potential oxygen: Present-day performance and future projection

C. D. Nevison[1], M. Manizza[2], R. F. Keeling[2], B. B. Stephens[3], J. D. Bent[3], J. Dunne[4], T. Ilyina[5], M. Long[3], L. Resplandy[2], J. Tjiputra[6], and S. Yukimoto[7]

[1]Institute for Arctic and Alpine Research, University of Colorado Boulder, Boulder, Colorado, USA, [2]Scripps Institution of Oceanography, La Jolla, California, USA, [3]National Center for Atmospheric Research, Boulder, Colorado, USA, [4]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, New Jersey, USA, [5]Max Planck Institute for Meteorology, Hamburg, Germany, [6]Uni Climate, Uni Research and Bjerknes Centre for Climate Research, Bergen, Norway, [7]Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan

Abstract Observed seasonal cycles in atmospheric potential oxygen (APO ~ O_2 + 1.1 CO_2) were used to evaluate eight ocean biogeochemistry models from the Coupled Model Intercomparison Project (CMIP5). Model APO seasonal cycles were computed from the CMIP5 air-sea O_2 and CO_2 fluxes and compared to observations at three Southern Hemisphere monitoring sites. Four of the models captured either the observed APO seasonal amplitude or phasing relatively well, while the other four did not. Many models had an unrealistic seasonal phasing or amplitude of the CO_2 flux, which in turn influenced APO. By 2100 under RCP8.5, the models projected little change in the O_2 component of APO but large changes in the seasonality of the CO_2 component associated with ocean acidification. The models with poorer performance on present-day APO tended to project larger net carbon uptake in the Southern Ocean, both today and in 2100.

1. Introduction

The Southern Ocean is an important sink for atmospheric CO_2 that accounts for more than 40% of oceanic anthropogenic carbon uptake [*Khatiwala et al.*, 2009; *Sallée et al.*, 2012; *Frölicher et al.*, 2015]. Some modeling studies have suggested that the growth of the Southern Ocean carbon sink expected during the 1980s and 1990s may have been offset by increased wind-driven upwelling of deep waters naturally enriched in carbon, which counteracts the air-sea pCO_2 gradient caused by the fossil fuel-driven buildup of atmospheric CO_2 [*Le Quéré et al.*, 2007; *Lovenduski et al.*, 2008]. However, more recent papers have found that the expected increasing strength of the Southern Ocean carbon sink resumed in the 2000s [*Landschutzer et al.*, 2015; *Munro et al.*, 2015a]. In this paper, we use "sink" to refer to the net flux of carbon over the Southern Ocean, including both the anthropogenic flux related to rising CO_2 and the natural flux due to the circulation-driven redistribution of carbon in the ocean.

Evaluation of trends in the Southern Ocean carbon sink based on ocean mixed layer observations of pCO_2 have been complicated by data gaps and natural interannual and decadal variability, which hinder efforts to detect long-term secular trends [*Lovenduski et al.*, 2015; *Munro et al.*, 2015b]. Observations of atmospheric CO_2 provide only a weak constraint on the air-sea flux of CO_2 because of large confounding signals from the land biosphere. Land exchanges can dominate the cycle even in the Southern Hemisphere due to long-range transport of signals from the north [*Nevison et al.*, 2008a]. The seasonal cycle in atmospheric CO_2 is an especially challenging metric for assessing ocean processes, due to carbonate buffering, competing thermal and biological influences, and the dominant influence of terrestrial photosynthesis and respiration [*Heimann et al.*, 1989; *Cadule et al.*, 2010].

The seasonal cycle in atmospheric potential oxygen (APO ~ O_2 + 1.1 CO_2) provides information about the carbon cycle that is not resolvable from atmospheric CO_2 alone [*Keeling et al.*, 1996; *Stephens et al.*, 1998; *Manning and Keeling*, 2006]. APO filters out terrestrial influences by assuming a compensating −1.1:1 O_2:CO_2 stoichiometry between terrestrial photosynthesis and respiration, and thereby isolates oceanic biogeochemical signals. A key advantage of APO is that O_2 air-sea fluxes, in contrast to CO_2 air-sea fluxes, are not

damped by carbonate chemistry but otherwise respond to similar processes. APO varies seasonally mainly due to air-sea exchanges of O_2 associated with oceanic export production, net community production, deep ventilation, and heating and cooling, all of which are key processes controlling air-sea CO_2 exchanges.

Earth System Models (ESMs) are widely used tools for projecting the future evolution of ocean carbon uptake. Output from a range of ESMs, forced with common historical and future scenarios, is available through the Coupled Model Intercomparison Project phase 5 (CMIP5). Many recent studies have tested the skill of the CMIP5 ocean models against hydrographic measurements and remotely sensed ocean color products [*Anav et al*., 2013; *Bopp et al*., 2013; *Henson et al*., 2013]. One recent study evaluated six CMIP5 ocean models against observed APO seasonal cycles [*Nevison et al*., 2015]. Here we build upon that study, focusing on the Southern Ocean and examining APO seasonal cycles predicted by an expanded set of eight CMIP5 models for both the present-day and the "business as usual" Representative Concentration Pathway (RCP8.5) scenario. We examine the sensitivity of APO to the model air-sea CO_2 flux, which exerts a substantial influence on our results and explore correlated seasonal patterns in model O_2 and CO_2 fluxes. Finally, we relate performance on several metrics of the APO cycle to the model Southern Ocean carbon sink, a variable of key interest in CMIP5.

2. Methods

2.1. CMIP5 Ocean Biogeochemistry Models

The eight CMIP5 models analyzed in this study include the ocean general circulation biogeochemical components of the Geophysical Fluid Dynamics Laboratory (GFDL) Earth System Models (depth-based ESM2M and density-based ESM2G vertical oceans) [*Dunne et al*., 2012, 2013]; the Institut Pierre-Simon Laplace Coupled Model 5, version IPSL-CM5A-LR (IPSL) [*Aumont and Bopp*, 2006; *Séférian et al*., 2012]; the Community Earth System Model (CESM) [*Moore et al*., 2002, 2004, 2013]; the Max Planck Institute for Meteorology (MPI-ESM) Earth System Model, version MPI-ESM-LR [*Ilyina et al*., 2013]; the Norwegian Earth System Model-ME (NorESM1) [*Tjiputra et al*., 2013]; the Hadley Global Environmental Model 2-Carbon Cycle (HadGEM2) [*Palmer and Totterdell*, 2001]; and the Meteorological Research Institute Earth System Model 1 (MRI) [*Yukimoto*, 2011].

For each model, the following output fields were obtained from the CMIP5 historical simulation (middle 1800s–2005) and the RCP8.5 simulation (2006–2100): net air-sea O_2 and CO_2 fluxes, net surface heat flux (Q), and sea surface salinity (SSS) and temperature (SST). These simulations used prescribed atmospheric CO_2 values, guaranteeing that all the ocean models "saw" the same atmospheric concentration. The eight models above encompass all available CMIP5 models that reported the fields needed to compute APO (not counting the multiple spatial resolutions and large ensembles available for some models). The SSS, SST, and Q fields were used to estimate air-sea N_2 fluxes based on the $Q(dS/dT)_{N2}/C_p$ equation [*Keeling et al*., 1993; *Manizza et al*., 2012] with modifications from *Jin et al*. [2007]. In this equation, Q is heat flux, $(dS/dT)_{N2}$ is the temperature derivative of the solubility coefficient for N_2, and C_p is the heat capacity of sea water.

2.2. Atmospheric Transport Model Simulations With GEOS-Chem

The ocean model O_2, N_2, and CO_2 air-sea fluxes were translated into atmospheric potential oxygen (APO) using forward simulations with the GEOS-Chem atmospheric transport model (ATM) [*Suntharalingam et al*., 2004; *Nassar et al*., 2010]. GEOS-Chem was run at $2° \times 2.5°$ horizontal resolution, with 47 sigma levels, driven by MERRA (Modern Era Retrospective-Analysis) meteorological forcing. The GEOS-Chem simulations were forced with ocean model air-sea fluxes from 1997 to 2004 for the historical case and 2095–2098 for the RCP8.5 scenario.

For each CMIP5 model, separate passive tracer GEOS-Chem simulations were performed for each gas to create separate time series of atmospheric O_2, N_2, and CO_2. Anomalies of O_2, N_2, and CO_2 were simulated as equivalent deviations in trace gas mole fractions in micromole per mole of dry air and combined to calculate a 3 year (excluding year 1 of the simulation) model APO time series in per meg units [*Stephens et al*., 1998; *Naegler et al*., 2007; *Nevison et al*., 2008b, 2012]:

$$\mathrm{APO_{model}} = \frac{1}{X_{O2}}(O_2) - \frac{1}{X_{N2}}(\mathrm{N_2}) + \frac{1.1}{X_{O2}}(\mathrm{CO_2}) \quad (1)$$

where X_{O2} and X_{N2} are the dry air mole fractions of O_2 and N_2, treated here as constants (0.2094 and 0.7808, respectively) and (O_2), (N_2), and (CO_2) are the equivalent trace gas mole fraction deviations. Here N_2 is needed

because atmospheric O_2 actually is measured as the O_2/N_2 ratio, as described below. The mean seasonal cycle was estimated by fitting a third-order polynomial plus first two harmonics function to the daily GEOS-Chem station-specific output. The harmonic components of the fit were used to construct the mean seasonal cycle and to identify the day of the seasonal maximum.

An additional product, APO_{Taka}, was computed with equation (2).

$$APO_{Taka} = \frac{1}{X_{O2}}(O_2) - \frac{1}{X_{N2}}(N_2) + \frac{1.1}{X_{O2}}\left(CO_{2,Taka}\right) \quad (2)$$

where all components are the same as those used in equation (1), except for the oceanic CO_2 component. Here ($CO_{2,Taka}$) is the atmospheric time series resulting from a GEOS-Chem simulation forced with the monthly mean climatological air-sea CO_2 fluxes of *Takahashi et al.* [2009]. APO_{Taka} was computed to evaluate the sensitivity of APO_{model} to the CMIP5 air-sea CO_2 fluxes. In all APO_{model} and APO_{Taka} calculations, fossil CO_2 emissions were assumed to make a negligible contribution to the seasonal cycle [*Nevison et al.*, 2008b] and thus were not included.

2.3. APO Data

2.3.1. Surface Monitoring Sites

Atmospheric O_2 data, measured in terms of deviations in the O_2/N_2 ratio, were obtained from three surface monitoring sites from the Scripps Institution of Oceanography (SIO) network. The sites included South Pole (SPO, 89.98°S, 24.8°W), Palmer Station (PSA, 64.9°S, 64.0°W), and Cape Grim, Tasmania (CGO, 40.7°S, 144.7°E). Data are available from the early to middle 1990s through the present, depending on the station [*Manning and Keeling*, 2006].

Observed APO was calculated according to

$$APO_{obs} = \delta(O_2/N_2) + \frac{1.1}{X_{O2}}(CO_2\text{-}350) \quad (3)$$

where $\delta(O_2/N_2)$ is the relative deviation in the O_2/N_2 ratio from a reference ratio in per meg units, $X_{O2} = 0.2094$ is the O_2 mole fraction of dry air [*Tohjima et al.*, 2005], CO_2 is the dry air mole fraction of carbon dioxide in parts per million (μmol mol^{-1}), 350 is an arbitrary reference value, and 1.1 is an estimate of the $-O_2$:CO2 ratio of terrestrial respiration and photosynthesis [*Severinghaus*, 1995]. Mean seasonal cycles and day of seasonal maxima for observed APO were estimated using the methodology described above in section 2.2.

3. Results and Discussion

3.1. Evaluation of CMIP5 Ocean Models Using APO

At all three surface sites (SPO, PSA, and CGO), the APO_{obs} seasonal amplitude falls in the lower range of the CMIP5 APO_{model} results, with most models, except IPSL and MRI, tending to overestimate the observed amplitude (Figures 1 and 2). Three models (CESM, ESM2M, and ESM2G) reproduce the observed phasing of the seasonal cycle well, while the others are either too early by up to 1–2 months or too late by ~1/2 month (Figure 2). The evaluation of the models based on GEOS-Chem forward simulations is generally consistent with the results of *Nevison et al.* [2015], who used a matrix method to translate air-sea fluxes from a subset of the CMIP5 models featured here into APO seasonal cycles at surface sites. That study reported higher confidence in the phasing of model APO results and somewhat lower confidence in the amplitude, due to model-data mismatch errors associated with ATM uncertainty.

ATM uncertainty has long been an important issue when using surface level APO data to evaluate coupled ocean model/ATM simulations [*Battle et al.*, 2006; *Naegler et al.*, 2007] and remains a concern in the current analysis. Model surface-level APO is known to be particularly sensitive to trapping of strong local air-sea fluxes at sites like PSA [*Blaine*, 2005]. Among the three SIO Southern Ocean stations, SPO is the farthest removed from local sources and the least sensitive to ATM uncertainty [*Blaine*, 2005]. We therefore focus on SPO in our evaluation below.

Among the eight CMIP5 models examined here, four fall within the uncertainty of either the observed APO phase (CESM, ESM2M, and ESM2G) or amplitude (IPSL) at SPO while MPI-ESM, NorESM1, and HadGEM2, lie well outside the uncertainty on both metrics. MRI lies well outside the observed APO phase uncertainty and marginally outside the observed amplitude uncertainty (Figure 2). The evaluation based on surface-level

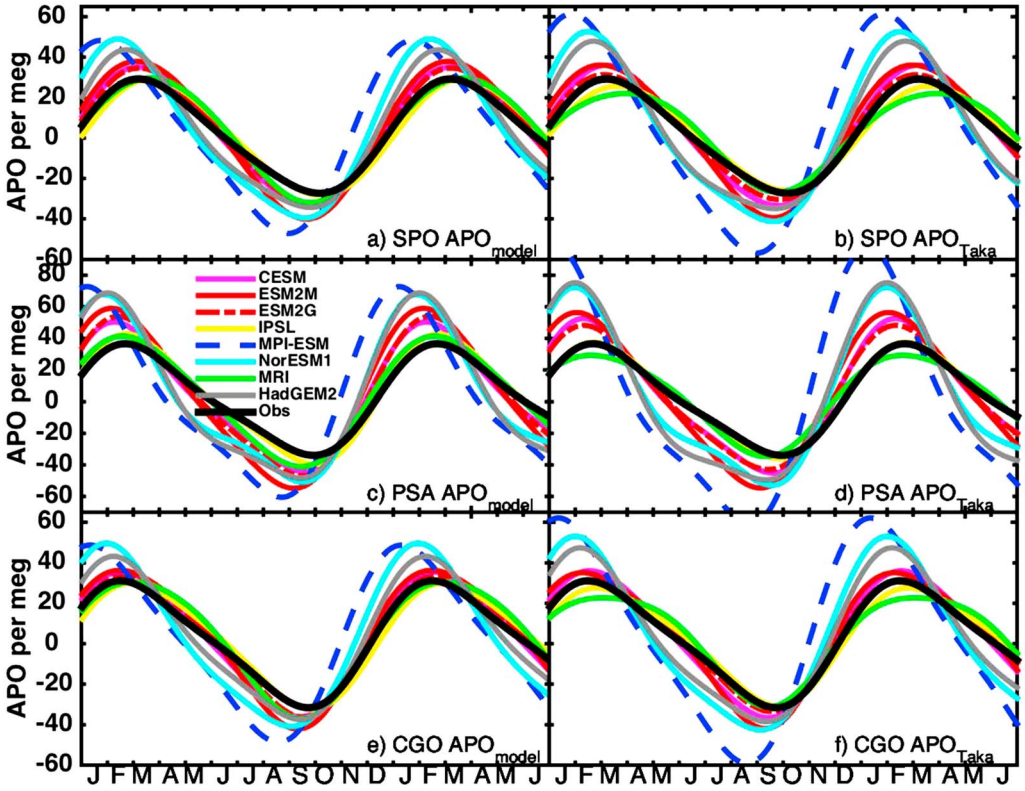

Figure 1. APO mean annual cycle for 1997–2004 simulated by GEOS-Chem forced with air-sea fluxes from eight CMIP5 models. Heavy black line shows the SIO mean annual APO cycle, (a, b) South Pole, (c, d) Palmer Station, and (e, f) Cape Grim. Note the different *Y* axis scale for the middle panels. (Figures 1a, 1c, and 1e) GEOS-Chem forced by CMIP5 O_2, N_2, and CO_2 air-sea fluxes. (Figures 1b, 1d, and 1f) Same as left, except the CMIP5 CO_2 fluxes are replaced with climatological air-sea CO_2 fluxes [*Takahashi et al.*, 2009].

results at SPO agrees well with GEOS-Chem column average results from 300 to 1000 mbar compared to aircraft data over the Southern Ocean (Figure S1 in the supporting information) [*Bent*, 2014]. Column averages eliminate a major source of ATM uncertainty associated with different vertical mixing schemes, in which some ATMs trap fluxes more strongly near the surface than others [*Stephens et al.*, 2007].

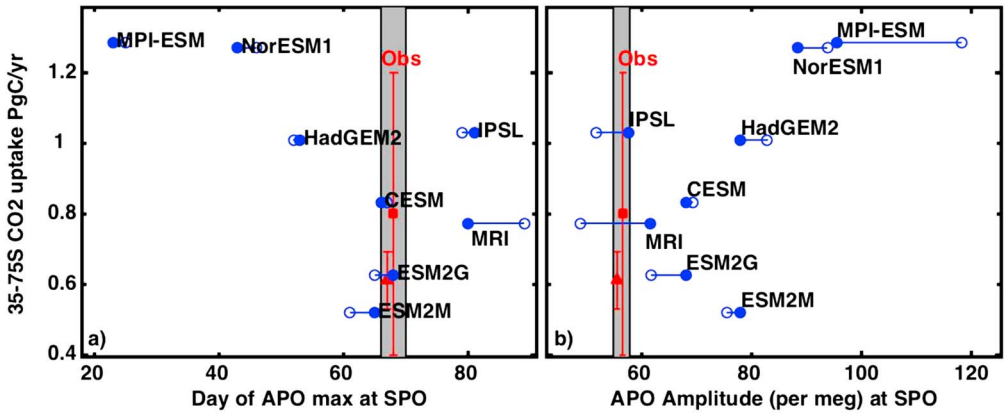

Figure 2. Annual mean CO_2 uptake in the Southern Ocean for 1997–2004 integrated over 75°−35°S as simulated by eight CMIP5 models plotted versus two metrics of the present-day APO (solid circles) and APO_{Taka} (open circles) seasonal cycle at South Pole: (a) day of seasonal maximum, (b) meanAPO amplitude. The grey shaded area denotes the estimated uncertainty range in the APO observations. Two observation-based estimates of the carbon sink integrated over 75°–35°S are shown as red squares [*Takahashi et al.*, 2009] and red triangles [*Landschutzer et al.*, 2015], with error bars estimated assuming 50% uncertainty [*Lenton et al.*, 2013] or from the reported annual mean standard deviation, respectively.

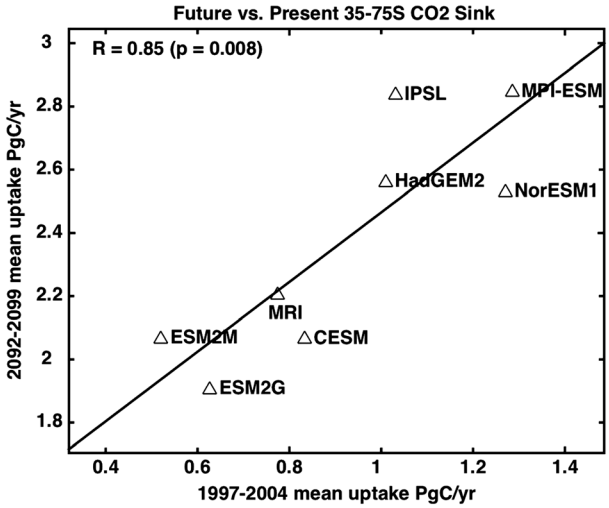

Figure 3. Annual mean CO_2 uptake in the Southern Ocean integrated over 75°S–35°S for 8 CMIP5 models, comparing present-day 1997–2004 results and 2092–2099 projections under the RCP8.5 scenario.

Since many ocean models do not reliably simulate the seasonal phase and amplitude of air-sea CO_2 fluxes [*Jiang et al.*, 2014; *Majkut et al.*, 2014], we performed a sensitivity test in which the CMIP5 CO_2 fluxes were replaced with climatological CO_2 fluxes [*Takahashi et al.*, 2009]. The resulting APO_{Taka} seasonal cycles show a clearer inter-model spread, with APO_{obs} tending to fall somewhere between IPSL and CESM and the GFDL models (Figures 1b, 1d, and 1f and S1, right). These results illustrate that APO_{model} is sensitive to the air-sea CO_2 flux component, as discussed further below.

3.2. Present and Future APO and the Southern Ocean Carbon Sink

The four top-performing models (CESM, ESM2M, ESM2G, and IPSL) with respect to the APO_{model} seasonal cycle at SPO predict a present-day carbon sink of 0.75 ± 0.2 Pg C/yr in the Southern Ocean (defined from 35°S–75°S). In comparison, the full eight CMIP5 model set predicts a larger uptake of 0.92 ± 0.3 Pg C/yr. The two models (MPI-ESM and NorESM1) that predict the largest Southern Ocean carbon uptake (1.3 Pg C/yr) are the least consistent with the observed seasonal cycle in APO (Figure 2). These results hold when plotting the CMIP5 carbon sink against either the phase or amplitude of the APO seasonal cycle.

The present-day CMIP5 Southern Ocean carbon sink is a statistically significant predictor ($R = 0.85$, $p = 0.008$) of the future Southern Ocean carbon sink in 2100 projected under RCP8.5 (Figure 3). Among CESM, ESM2M, and ESM2G, the three top-performing models with respect to present-day APO seasonal phasing, the mean future Southern Ocean sink circa 2100 under RCP8.5 is 2.0 ± 0.1 Pg C/yr. In contrast, the mean future sink projected by MPI-ESM, NorESM1, and HadGEM2 (models that perform less well with respect to present-day APO) is 2.6 ± 0.2 Pg C/yr. However, IPSL, the top-performing model on APO amplitude, also projects a relatively large future sink of 2.8 Pg C/yr.

While the CMIP5 models forecast large changes in the future Southern Ocean carbon sink, they project little change in the O_2 or N_2 components of the APO seasonal cycle under RCP8.5 (Figure S2) The GFDL models project the largest decrease in the O_2 seasonal amplitude, but this corresponds to only a −8% change relative to the present day. In all models, there is little change in the O_2 cycle phasing. The lack of strong changes in the O_2 component may reflect future productivity increases in some regions of the Southern Ocean but decreases in others [*Steinacher et al.*, 2010; *Bopp et al.*, 2013].

In contrast to the O_2 component, many models project large changes in the phase and amplitude of the oceanic CO_2 flux component, which in turn has a substantial impact on APO. These models project a breakdown in carbonate buffering as the ocean acidifies, with an associated decrease in the timescale of air-sea CO_2 exchange [*Doney et al.*, 2009; *Hauck and Völker*, 2015]. However, as discussed below, the models projecting the strongest amplification of the oceanic CO_2 seasonal cycle in the future (Figure S2) are those that exaggerate the seasonal cycle in present-day surface ocean pCO_2 relative to observations.

3.3. Seasonality in CO_2 Fluxes and Relevance to APO and the Carbon Sink

As shown in Figures 1 and S1, the model APO seasonal cycle is sensitive to the CMIP5 air-sea CO_2 fluxes. In some models with small O_2/N_2 seasonal amplitudes (IPSL and MRI), the oceanic CO_2 term is in phase with O_2/N_2 and thus acts to increase APO (Figure S3). Conversely, in models with large O_2/N_2 seasonal amplitudes (MPI-ESM, NorESM1, and HadGEM2), the oceanic CO_2 seasonal cycle is also large and out of phase with O_2/N_2, acting to decrease APO (Figure S3). The net effect leads to convergence in the APO amplitude among the models, despite large differences in underlying CO_2 and O_2 components. As discussed below, the CMIP5

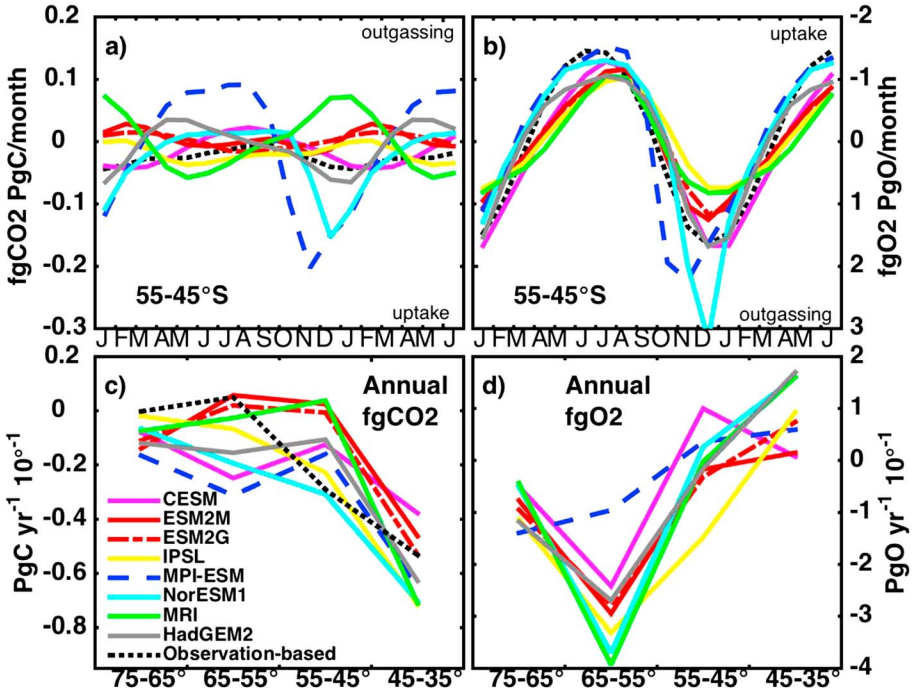

Figure 4. (top row) Mean seasonal cycle in the air-sea (a) CO_2 flux and (b) O_2 flux, integrated over the 55–45°S latitude band for 1997–2004. For both CO_2 and O_2, positive values reflect outgassing from ocean to atmosphere. The seasonal O_2 fluxes are plotted with a reverse *Y* axis to help visualize their relationship with the CO_2 fluxes. (bottom row) Annual mean air-sea (c) CO_2 flux and (d) O_2 flux from eight CMIP5 models, integrated over 10° latitude bands from 75°S–35°S. Dotted black lines in Figures 4a–4c are observation-based climatological CO_2 fluxes of *Takahashi et al.* [2009] or seasonal O_2 flux anomalies of *Garcia and Keeling* [2001] (which are zero in the annual mean and thus not shown in Figure 4d).

air-sea CO_2 flux differences reflect issues with model biogeochemistry and ocean dynamics that are relevant both to APO_{model} and the strength of the model Southern Ocean carbon sink.

Four CMIP5 models (MPI-ESM, NorESM1, CESM, and HadGEM2) predict an inverse seasonal relationship between air-sea CO_2 and O_2 fluxes in Southern Ocean (Figure 4). This relationship is most pronounced in late spring and early summer and is particularly striking for MPI-ESM and NorESM1. Both those models predict strong austral spring/early summer O_2 outgassing that appears to be driven mainly by the biology/ventilation component of the O_2 flux, since the thermal component is relatively similar among all eight CMIP5 models (Figure S4). MPI-ESM and NorESM1 show a correspondingly strong biology-driven uptake of CO_2 in spring and early summer, respectively. In contrast, four CMIP5 models (ESM2G, ESM2M, IPSL, and MRI) predict that the seasonal air-sea CO_2 and O_2 fluxes are in phase, with either neutral or net CO_2 release to the atmosphere in summer. In these models, the biological uptake is small enough that the summertime CO_2 flux is balanced, or even dominated by the thermal outgassing. For comparison, the observation-based estimate of *Takahashi et al.* [2009] shows net CO_2 uptake during summer in the Southern Ocean, although considerably weaker than that predicted by MPI-ESM, NorESM1, HadGEM2, and CESM, especially south of 55°S (Figures 4 and S4). Effectively, none of the models captures the observed CO_2 flux pattern, yet representatives from both the net summer uptake and net summer outgassing model groups are among the top performers on APO. This suggests that any of these models might give the right answer on APO at least in part for the wrong reasons.

Few previous CMIP5 intercomparison papers have discussed APO and air-sea O_2 fluxes, but several have examined air-sea CO_2 fluxes and corresponding surface pCO_2. A prominent finding from those studies is that most models tend to predict a substantially stronger (up to a factor of 10) seasonal cycle in Southern Ocean surface pCO_2 than that inferred from observations [*Anav et al.*, 2013; *Jiang et al.*, 2014], a conclusion supported by our own analysis of air-sea CO_2 fluxes (Figures 4a and S4). The large seasonal pCO_2 swings have been linked in some models to excessive seasonal amplitude in mixed layer depth (MLD) [*Anav et al.*, 2013; *Jiang et al.*, 2014]. The large pCO_2 seasonal amplitudes are associated with negative annual mean surface

ocean pCO_2 biases relative to observations in the Southern Ocean, especially south of the Polar Front (Figures 4c and S4), which imply an overly large CO_2 sink. The annual mean bias is driven by summertime values that are strongly negative relative to observed surface and atmospheric pCO_2 and thus associated with strong summertime carbon uptake.

More generally, the excessive model pCO_2 seasonal amplitudes are linked to the models' inability to capture the appropriate balance between seasonality in pCO_2 associated with SST-driven changes in solubility, and DIC-driven seasonality associated with spring/summer biological uptake and fall/winter upwelling and ventilation of DIC-enriched deep water. For CO_2, these thermal and biology-driven cycles constitute large, opposing fluxes, for which models often have difficulty capturing the net balance. In contrast, thermal and biological signals tend to be in phase and mutually reinforcing for O_2 fluxes [*Stephens et al*., 1998].

Two of the CMIP5 models with the largest Southern Ocean carbon uptake, MPI-ESM and NorESM, show CO_2 flux patterns that are consistent with a DIC-driven seasonality. The large springtime O_2 fluxes and early APO rise in MPI-ESM suggest excessive high-latitude biological production that ramps up too early in spring (Figures 1 and 2). The NorESM1 temporal patterns indicate some of the same problems in spring/summer biological production.

IPSL captures the observed APO amplitude reasonably well but predicts a too late spring rise, contrasting with MPI-ESM and NorESM1 in both these respects (Figures 1 and 2). However, like MPI-ESM and NorESM1, IPSL predicts a large present-day Southern Ocean carbon sink of >1 PgC/yr. IPSL is characterized by the smallest O_2 outgassing in spring and summer of all eight models (Figure 4) and also predicts the weakest NPP, both globally and in the Southern Ocean, of the CMIP5 models evaluated in *Nevison et al*. [2015]. Compared to the Takahashi climatology, IPSL tends to underestimate carbon uptake in summer and overestimate it in fall and winter (Figures 4a and S4). These reversed seasonal patterns suggest that IPSL may have unrealistically weak biological production in summer and correspondingly weak natural outgassing of CO_2 from winter deep water ventilation.

We note finally that anthropogenic CO_2 uptake is influenced by global circulation patterns in both the oceanic models and in the atmospheric winds that drive them [*Toggweiler et al*., 2006; *Frölicher et al*., 2015]. Net CO_2 uptake in CMIP5 also may be influenced by model drift [*Frölicher et al*., 2015], which we did not control in this study. In future work, a deeper understanding of how models differ in representing MLD and winter mixing could help elucidate whether the emergent links between performance on APO and net CO_2 uptake are truly mechanistic or merely coincidental.

4. Conclusions

Seasonal cycles in APO at Southern Hemisphere sites were used to test the ocean model components of eight Earth system models participating in CMIP5. The model/data comparisons revealed that four of the models tested reproduce either the phase or amplitude of the observed cycles reasonably well while four lie outside the range of observational uncertainty on both metrics. The APO seasonal cycle is sensitive to the air-sea CO_2 flux component, which acts to homogenize APO across the CMIP5 models, making it appear more similar than it would be based on the O_2/N_2 component alone.

The models that perform more poorly on the APO seasonal cycle tend to predict a larger carbon sink in the Southern Ocean, defined from 35° to 75°S. However, different and complex factors may influence the relationship between APO and the net carbon sink for any individual model. In some models, most carbon uptake occurs in fall and winter, with net summer CO_2 outgassing that enhances APO seasonal amplitude. In other models, most carbon uptake occurs in spring and summer, with net ingassing that reduces APO amplitude. Overall, the APO-based analysis suggests that more conservative estimates of the present-day and future Southern Ocean carbon sink may be more realistic.

Acknowledgments
The authors gratefully acknowledge the CMIP5 ocean modelers for providing the output that made this project possible. In particular, we thank Keith Lindsay, Paul Halloran, and Christoph Heinze for their assistance. We are also grateful to Anna Cabre, Mo Green, and Irina Marinov for help in obtaining CMIP5 output, to Andrew Schuh for help with the GEOS-Chem simulations, and to two anonymous reviewers for their helpful comments. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver, and the National Center for Atmospheric Research. The National Center for Atmospheric Research is sponsored by the National Science Foundation. C.D.N. and M.M. acknowledge support from NASA Ocean Biology and Biogeochemistry grant NNX11AL73G and J.T. acknowledges Norwegian Research Council-funded ORGANIC project (239965). This study is a contribution to the EU H2020 project CRESCENDO (grant 641816). The data used are listed in the references and supporting information. The APO observations are available from the Scripps Institution of Oceanography under http://scrippso2.ucsd.edu and http://scrippsco2.ucsd.edu. The CMIP5 output fields are available from http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html.

References

Anav, A., P. Friedlingstein, M. Kidston, L. Bopp, P. Ciais, P. Cox, C. Jones, M. Jung, R. Myneni, and Z. Zhu (2013), Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth system models, *J. Clim.*, *26*, 6801–6843.

Aumont, O., and L. Bopp (2006), Globalizing results from ocean in situ iron fertilization studies, *Global Biogeochem. Cycles*, *20*, GB2017, doi:10.1029/2005GB002591.

Battle, M., et al. (2006), Atmospheric potential oxygen: New observations and their implications for some atmospheric and oceanic models, *Global Biogeochem. Cycles*, *20*, GB1010, doi:10.1029/2005GB002534.
Bent, J. D. (2014), Airborne oxygen measurements over the Southern Ocean as an integrated constraint of seasonal biogeochemical processes, PhD thesis, Univ. of Calif., San Diego, La Jolla.
Blaine, T. W. (2005), Continuous measurements of atmospheric Ar/N_2 as a tracer of air-sea heat flux: Models, methods, and data PhD thesis, Univ. of Calif., San Diego, La Jolla.
Bopp, L., et al. (2013), Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models, *Biogeosciences*, *10*, 6225–6245, doi:10.5194/bg-10-6225-2013.
Cadule, P., P. Friedlingstein, L. Bopp, S. Sitch, C. D. Jones, P. Ciais, S. L. Piao, and P. Peylin (2010), Benchmarking coupled climate carbon models against long term atmospheric CO_2 measurements, *Global Biogeochem. Cycles*, *24*, GB2016, doi:10.1029/2009GB003556.
Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas (2009), Ocean acidification: The other CO_2 problem, *Annu. Rev. Mar. Sci.*, *1*, 169–192.
Dunne, J. P., et al. (2012), GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics, *J. Clim.*, doi:10.1175/JCLI-D-11-00560.1.
Dunne, J. P., et al. (2013), GFDL's ESM2 global coupled climate-carbon Earth System Models Part II: Carbon system formulation and baseline simulation characteristics, *J. Clim.*, doi:10.1175/JCLI-D-12-00150.1.
Frölicher, T. L., J. Sarmiento, D. Pynter, J. Dunne, J. Krasting, and M. Winton (2015), Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models, *J. Clim.*, *28*, 862–886, doi:10.1175/JCLI-D-14-00117.1.
Garcia, H. E., and R. F. Keeling (2001), On the global oxygen anomaly and air-sea flux, *J. Geophys. Res.*, *106*(31), 155–31,166, doi:10.1029/1999JC000200.
Hauck, J., and C. Völker (2015), Rising atmospheric CO_2 leads to large impact of biology on Southern Ocean CO_2 uptake via changes of the Revelle factor, *Geophys. Res. Lett.*, *42*, 1459–1464, doi:10.1002/2015GL063070.
Heimann, M., C. D. Keeling, and C. J. Tucker (1989), A three-dimensional model of atmospheric CO_2 transport based on observed winds: 3. Seasonal cycle and synoptic time scale variations, in *Aspects of Climate Variability and the Western Americas*, vol. 55, edited by D. H. Peterson, pp. 277–303, AGU, Washington, D. C.
Henson, S., H. Cole, C. Beaulieu, and A. Yool (2013), The impact of global warming on seasonality of ocean primary production, *Biogeosciences*, *10*, 4357–4369, doi:10.5194/bg-10-4357-2013.
Ilyina, T., K. D. Six, J. Segschneider, E. Maier-Reimer, H. Li, and I. Núñez-Riboni (2013), Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations, *J. Adv. Model. Earth Syst.*, *5*, 287–315, doi:10.1029/2012MS000178.
Jiang, C., S. Gille, J. Sprintall, and C. Sweeney (2014), Drake Passage oceanic *p*CO2: Evaluating CMIP5 coupled carbon-climate models using in situ observations, *J. Clim.*, *27*, 76–100.
Jin, X., R. G. Najjar, F. Louanchi, and S. C. Doney (2007), A modeling study of the seasonal oxygen budget of the global ocean, *J. Geophys. Res.*, *112*, C05017, doi:10.1029/2006JC003731.
Keeling, R. F., R. P. Najjar, M. L. Bender, and P. P. Tans (1993), What atmospheric oxygen measurements can tell us about the global carbon cycle, *Global Biogeochem. Cycles*, *7*, 37–67, doi:10.1029/92GB02733.
Keeling, R. F., S. C. Piper, and M. Heimann (1996), Global and hemispheric CO_2 sinks deduced from changes in atmospheric O_2 concentration, *Nature*, *391*, 218–221.
Khatiwala, S., F. Primeau, and T. Hall (2009), Reconstruction of the history of anthropogenic CO_2 concentrations in the ocean, *Nature*, *462*, 346–350.
Landschutzer, P., et al. (2015), The reinvigoration of the Southern Ocean carbon sink, *Science*, *349*, 6253.
Le Quéré, C., et al. (2007), Saturation of the Southern Ocean CO_2 sink due to recent climate change, *Science*, *316*, 1735–1738.
Lenton, A., et al. (2013), Sea-air CO_2 fluxes in the Southern Ocean for the period 1990–2009, *Biogeosciences*, *10*, 4037–4054.
Lovenduski, N. S., N. Gruber, and S. C. Doney (2008), Toward a mechanistic understanding of the decadal trends in the Southern Ocean carbon sink, *Global Biogeochem. Cycles*, *22*, GB3016, doi:10.1029/2007GB003139.
Lovenduski, N. S., A. R. Fay, and G. A. McKinley (2015), Observing multidecadal trends in Southern Ocean CO_2 uptake: What can we learn from an ocean model? *Global Biogeochem. Cycles*, *29*, 416–426, doi:10.1002/2014GB004933.
Majkut, J., B. R. Carter, T. L. Frölicher, C. O. Dufour, K. B. Rodgers, and J. L. Sarmiento (2014), An observing system simulation for Southern Ocean carbon dioxide uptake, *Phil. Trans. R. Soc. A.*, *372*, 20130046.
Manizza, M., R. F. Keeling, and C. D. Nevison (2012), On the processes controlling the seasonal cycles of the air-sea fluxes of O_2 and N_2O: A modeling study, *Tellus-B*, *64*, 18,429, doi:10.3402/tellusb.v64i0.18429.
Manning, A. C., and R. F. Keeling (2006), Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network, *Tellus-B*, *58B*, 95–116.
Moore, J. K., S. C. Doney, J. C. Kleypas, D. M. Glover, and I. Y. Fung (2002), An intermediate complexity marine ecosystem model for the global domain, *Deep Sea Res., Part II*, *49*, 403–462.
Moore, J. K., S. C. Doney, and K. Lindsay (2004), Upper ocean ecosystem dynamics and iron cycling in a global 3D model, *Global Biogeochem. Cycles*, *18*, GB4028, doi:10.1029/2004GB002220.
Moore, J., K. Lindsay, S. Doney, M. Long, and K. Misumi (2013), Marine Ecosystem Dynamics and Biogeochemical Cycling in the Community Earth System Model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 scenarios, *J. Clim.*, *26*, 9291–9312, doi:10.1175/JCLI-D-12-00566.1.
Munro, D. R., N. S. Lovenduski, T. Takahashi, B. R. Stephens, T. Newberger, and C. Sweeney (2015a), Recent evidence for a strengthening CO_2 sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002–2015), *Geophys. Res. Lett.*, *42*, 7623–7630, doi:10.1002/2015GL065194.
Munro, D. R., N. S. Lovenduski, B. R. Stephens, T. Newberger, K. R. Arrigo, T. Takahashi, P. D. Quay, J. Sprintall, N. M. Freeman, and C. Sweeney (2015b), Estimates of net community production in the Southern Ocean determined from time series observations (2002–2011) of nutrients, dissolved inorganic carbon, and surface ocean pCO_2 in Drake Passage, *Deep Sea Res. Part II*, *114*, 49–63.
Naegler, T., P. Ciais, J. Orr, O. Aumont, and C. Roedenbeck (2007), On evaluating ocean models with atmospheric potential oxygen, *Tellus*, *59B*, 138–156.
Nassar, R., et al. (2010), Modeling global atmospheric CO_2 with improved emission inventories and CO_2 production from the oxidation of other carbon species, *Geosci. Model Dev.*, *3*, 689–716.
Nevison, C. D., N. Mahowald, S. Doney, I. Lima, G. van der Werf, J. Randerson, D. Baker, P. Kasibhatla, and G. McKinley (2008a), Contribution of ocean, fossil fuel, land biosphere and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric CO_2, *J. Geophys. Res.*, *113*, G01010, doi:10.1029/2007JG000408.

Nevison, C. D., N. M. Mahowald, S. C. Doney, I. D. Lima, and N. Cassar (2008b), Impact of variable air-sea O_2 and CO_2 fluxes on atmospheric potential oxygen (APO) and land-ocean carbon sink partitioning, *Biogeosciences*, *5*, 875–889.

Nevison, C. D., R. F. Keeling, M. Kahru, M. Manizza, B. G. Mitchell, and N. Cassar (2012), Estimating net community production in the Southern Ocean based on atmospheric potential oxygen and satellite ocean color data, *Global Biogeochem. Cycles*, *26*, GB1020, doi:10.1029/2011GB004040.

Nevison, C. D., M. Manizza, R. F. Keeling, M. Kahru, L. Bopp, J. Dunne, J. Tiputra, T. Ilyina, and B. G. Mitchell (2015), Evaluating the ocean biogeochemical components of Earth system models using atmospheric potential oxygen and ocean color data, *Biogeosciences*, *12*, 193–208, doi:10.5194/bg-12-193-2015.

Palmer, J. R., and I. J. Totterdell (2001), Production and export in a global ocean ecosystem model, *Deep Sea Res. Part I*, *48*, 1169–1198, doi:10.1016/s0967-0637(00)00080-7.

Sallée, J. B., R. J. Matear, S. R. Rintoul, and A. Lenton (2012), Localized subduction of anthropogenic carbon dioxide in the Southern Hemisphere oceans, *Nat. Geosci.*, *5*, 579–584, doi:10.1038/NGEO1523.

Séférian, R., D. Iudicone, L. Bopp, T. Roy, and G. Madec (2012), Water mass analysis of effect of climate change on air-sea CO_2 fluxes: The Southern Ocean, *J. Clim.*, *25*, 3894–3908, doi:10.1175/JCLI-D-11-00291.1.

Severinghaus, J. P. (1995), Studies of the terrestrial O_2 and carbon cycles in sand dune gases and in Biosphere 2 PhD thesis, Columbia Univ., New York.

Steinacher, M., et al. (2010), Projected 21st century decrease in marine productivity: A multi-model analysis, *Biogeosciences*, *7*, 979–1005.

Stephens, B. B., R. F. Keeling, M. Heimann, K. D. Six, R. Murnane, and K. Caldeira (1998), Testing global ocean carbon cycle models using measurements of atmospheric O_2 and CO_2 concentration, *Global Biogeochem. Cycles*, *12*, 213–230, doi:10.1029/97GB03500.

Stephens, B. B., et al. (2007), Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO_2, *Science*, *316*, 1732–1735, doi:10.1126/science.1137004.

Suntharalingam, P., D. J. Jacob, P. I. Palmer, J. A. Logan, R. M. Yantosca, Y. Xiao, M. J. Evans, D. G. Streets, S. L. Vay, and G. W. Sachse (2004), Improved quantification of Chinese carbon fluxes using CO2/CO correlations in Asian outflow, *J. Geophys. Res.*, *109*, D18S18, doi:10.1029/2003JD004362.

Takahashi, T., et al. (2009), Climatological mean and decadal changes in surface ocean pCO_2, and net sea-air CO_2 flux over the global oceans, *Deep Sea Res., Part II*, 554–577, doi:10.1016/j.dsr2.2008.12.009.

Tjiputra, J. F., C. Roelandt, M. Bentsen, D. M. Lawrence, T. Lorentzen, J. Schwinger, Ø. Seland, and C. Heinze (2013), Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), *Geosci. Model Dev.*, *6*, 301–325.

Toggweiler, J. R., J. L. Russell, and S. R. Carson (2006), Midlatitude westerlies, atmospheric CO_2, and climate change during the ice ages, *Paleoceanography*, *21*, PA2005, doi:10.1029/2005PA001154.

Tohjima, Y., H. Mukai, T. Machida, Y. Nojiri, and M. Gloor (2005), First measurements of the latitudinal atmospheric O_2 and CO_2 distributions across the western Pacific, *Geophys. Res. Lett.*, *32*, L17805, doi:10.1029/2005GL023311.

Yukimoto, S., et al. (2011), MRI-ESM1 model description Technical reports of the Meteorological Research Institute, 64. [Available at http://www.mrijma.go.jp/Publish/Technical/DATA/VOL_64/ tec_rep_mri_64. pdf.]