AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Oceans

RESEARCH ARTICLE

10.1002/2017JC013461

Special Section:
The Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) Project: Technologies, Methods, and Early Results

Metrics for the Evaluation of the Southern Ocean in Coupled Climate Models and Earth System Models

Joellen L. Russell[1], Igor Kamenkovich[2], Cecilia Bitz[3], Raffaele Ferrari[4], Sarah T. Gille[5], Paul J. Goodman[1], Robert Hallberg[6], Kenneth Johnson[7], Karina Khazmutdinova[8], Irina Marinov[9], Matthew Mazloff[5], Stephen Riser[10], Jorge L. Sarmiento[11], Kevin Speer[8], Lynne D. Talley[5], and Rik Wanninkhof[12]

[1]Department of Geosciences, University of Arizona, Tucson, AZ, USA, [2]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL, USA, [3]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [4]Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [5]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA, USA, [6]Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA, [7]Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA, [8]Geophysical Fluid Dynamics Institute, Florida State University, Tallahassee, FL, USA, [9]Department of Earth and Environmental Science, University of Pennsylvania, Philadelphia, PA, USA, [10]School of Oceanography, University of Washington, Seattle, WA, USA, [11]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, [12]Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, FL, USA

Key Points:
- Observationally based metrics are essential for assessing, comparing, and improving the heat and carbon cycles in climate simulations
- Metrics included here assess winds and heat and carbon uptake, ACC transport, sea ice extent, frontal positions, and pH
- Ocean heat and carbon uptake are strongly correlated in models and observations

Supporting Information:
- Supporting Information S1
- Figure S1
- Figure S2
- Figure S3
- Figure S4
- Figure S5
- Figure S6
- Figure S7
- Figure S8
- Figure S9
- Figure S10
- Figure S11
- Figure S12
- Figure S13
- Figure S14
- Figure S15

Correspondence to:
J. Russell,
jrussell@email.arizona.edu

Citation:
Russell, J. L., Kamenkovich, I., Bitz, C., Ferrari, R., Gille, S. T., Goodman, P. J., et al. (2018). Metrics for the evaluation of the Southern Ocean in coupled climate models and earth system models. *Journal of Geophysical Research: Oceans, 123*. https://doi.org/10.1002/2017JC013461

Received 15 SEP 2017
Accepted 13 FEB 2018
Accepted article online 16 FEB 2018

© 2018. The Authors.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Abstract The Southern Ocean is central to the global climate and the global carbon cycle, and to the climate's response to increasing levels of atmospheric greenhouse gases, as it ventilates a large fraction of the global ocean volume. Global coupled climate models and earth system models, however, vary widely in their simulations of the Southern Ocean and its role in, and response to, the ongoing anthropogenic trend. Due to the region's complex water-mass structure and dynamics, Southern Ocean carbon and heat uptake depend on a combination of winds, eddies, mixing, buoyancy fluxes, and topography. Observationally based metrics are critical for discerning processes and mechanisms, and for validating and comparing climate and earth system models. New observations and understanding have allowed for progress in the creation of observationally based data/model metrics for the Southern Ocean. Metrics presented here provide a means to assess multiple simulations relative to the best available observations and observational products. Climate models that perform better according to these metrics also better simulate the uptake of heat and carbon by the Southern Ocean. This report is not strictly an intercomparison, but rather a distillation of key metrics that can reliably quantify the "accuracy" of a simulation against observed, or at least observable, quantities. One overall goal is to recommend standardization of observationally based benchmarks that the modeling community should aspire to meet in order to reduce uncertainties in climate projections, and especially uncertainties related to oceanic heat and carbon uptake.

Plain Language Summary Observationally based metrics are essential for the standardized evaluation of climate and earth system models, and for reducing the uncertainty associated with future projections by those models.

1. Introduction

The exchanges of heat and carbon dioxide between the atmosphere and ocean are important to Earth's climate under conditions of anthropogenic forcing. Observations show that 93% of the extra energy in the Earth System due to the radiative imbalance at the top of the atmosphere has accumulated in the ocean (Fasullo & Trenberth, 2012). The Southern Ocean south of 30°S occupies 30% of the surface ocean area and exerts significant influence on the climate system. Model simulations and observational analyses of the Southern Ocean indicate that: (1) 67–98% of the excess heat that is transferred from the atmosphere into the ocean each year is stored south of 20°S (Roemmich et al., 2015); (2) the Southern Ocean south of 30°S accounts for nearly half of the annual oceanic uptake of anthropogenic carbon dioxide from the

atmosphere (43% ±3%, Frölicher et al., 2015); (3) models suggest that vertical exchange within the Southern Ocean is responsible for supplying nutrients that fertilize three-quarters of the biological production in the global ocean north of 30°S (Marinov et al., 2006; Sarmiento et al., 2004); and (4) Southern Ocean winds and buoyancy fluxes are the principal source of energy for driving the large-scale deep meridional overturning circulation throughout the ocean (e.g., Marshall & Speer, 2012; Toggweiler & Samuels, 1998).

Investigations of the processes that determine Southern Ocean heat uptake have yielded different governing mechanisms (Dalan et al., 2005; Gregory, 2000; Huang et al., 2003; Manabe et al., 1990; Tamsitt et al., 2016). Ocean heat uptake has been shown to be model and resolution-dependent, through a combination of isopycnal and diapycnal mixing, eddies, winds, tidal mixing, and topography (Exarchou et al., 2015). The uncertainty in the efficiency with which heat is transferred from the surface into the deeper ocean contributes strongly to uncertainty in ocean heat uptake, thermal expansion projections, and transient surface warming projections (Kuhlbrodt & Gregory, 2012).

The Southern Ocean carbon balance is also complex. Wind-driven intensification of Southern Ocean upwelling and overturning has been proposed as a mechanism for increased outgassing from the deep ocean and enhanced atmospheric CO_2 on long, thousand-year timescales such as those associated with deglaciations (e.g., Anderson et al., 2009; Marinov et al., 2008; Toggweiler et al., 2006). Some have claimed that the recent intensification of the westerlies is responsible for a weaker Southern Ocean carbon sink between ∼1981 and ∼2004 (e.g., Le Quéré et al., 2007; Lenton & Matear 2007; Lovenduski et al., 2008). Landschützer et al. (2015) showed a reinvigoration of the carbon sink from 2002 to 2012 related to changes in the wind zonal symmetry suggesting large decadal differences in carbon uptake.

Despite the crucial role of the Southern Ocean in the global heat and carbon cycles, intermodel disagreements cannot, as yet, be resolved due to the sparseness of the observations that are expensive and difficult to obtain in this region. In addition, mesoscale eddies, which are thought to be critical in maintaining the Southern Ocean stratification (Marshall & Raidko, 2003) and governing its response to changing winds (Hallberg & Gnanadesikan, 2006), are not resolved by the majority of climate models used for the fifth phase of the Coupled Model Intercomparison Project (IPCC-AR5/CMIP5, Taylor et al., 2012). Furthermore, while the critical importance of the deep ocean for heat and carbon storage is known (Purkey & Johnson, 2012), shortcomings of the CMIP5 models (e.g., open ocean deep convection rather than off-shelf flow) contribute to the large CMIP5 multimodel spread in carbon and heat uptake and storage (Frölicher et al., 2015). The oceanographic community has made significant progress, through new strategies, tools, simulations, and analyses, toward establishing the necessary infrastructure that will enable us to advance our understanding of the Southern Ocean's role in climate (Rintoul et al., 2012).

In order to capitalize on recent advances, the Southern Ocean Working Group (SOWG), sponsored by US CLIVAR and Ocean Carbon and Biogeochemistry (https://usclivar.org/working-groups/southern-ocean), was formed, in part, to define new observationally based data/model metrics that are able to quantify the fidelity of climate simulations and that will allow for demonstrable progress and a reduction of model uncertainty in future climate projections. Here, we investigate the causes of model uncertainty, defined by the intermodel spread related to either the magnitude of the simulated change in a given year (e.g., mean surface air temperature in 2040, Allen et al., 2000) or the year in which a benchmark is surpassed (e.g., Southern Ocean aragonite undersaturation, McNeil & Matear, 2008), so that we may resolve intermodel disagreements.

The inference that reducing model error in simulations of today will ensure that model simulations of the future are less uncertain, while intuitive, is hard to quantify. Dalmonech et al. (2014) attribute one of the causes of model uncertainty to uncertainties related to the observations and the performance metrics, implying that better observations and metrics should provide information needed to reduce model uncertainty. Shiogama et al. (2016) found that much of the uncertainty associated with model projection stemmed from uncertainty in the observations, and comparing the models to longer observational records led to significantly reduced simulation uncertainty. Knutti et al. (2010) cast doubt on the idea that better simulations today imply better simulations in the future, but nevertheless emphasized that more quantitative metrics are essential. We note that if a simulation is "right for the wrong reasons," this would suggest that good agreement between historical model metrics and observations is no guarantee of useful future projections.

This study's goal is to present what the authors consider to be representative "useful" metrics for: (1) assessing a model simulation of the Southern Ocean, especially with respect to its role in the uptake of anthropogenic heat and carbon; and (2) highlighting leading-order processes that lead to intermodel simulation differences (i.e., how are the models different from observed, and what causes these differences). A similar guide for the assessing Antarctic climate simulations was also recently published (Bracegirdle et al., 2016). That study focused on large-scale climate metrics associated with twenty-first century simulations, whereas this analysis is more narrowly focused on metrics related to heat and carbon uptake. This study is an overview of the need for, and the utility of, observationally based metrics for model intercomparisons generally, and of the Southern Ocean in particular. We present those metrics that are rich in potential analyses and provide insight into the workings of the models and the Southern Ocean. Assessment of each metric, however, is potentially worthy of its own paper discussing the history, comparative strengths and weaknesses of the observations, and of different metric formulations applied in the literature, correlations with significant ocean processes (physical, chemical, biological), etc. Forthcoming studies by us and other researchers will do just this.

The next section of this study discusses our methods, including how we define the term "metric," and to which observationally based data sets we compare the model simulations. We discuss their applicability for the climatically important parameters we want to determine (e.g., we measure temperature but want to quantify heat content anomalies, and we measure DIC and pH but want to quantify the net air/sea CO_2 flux). We also discuss some technical considerations that we encountered while formulating these metrics. Section 3 presents results of different metrics applied to six different simulations and discusses the relevance of each to the workings of the Southern Ocean in general and specifically to the effect on the uptake of heat and carbon. In section 4, we illustrate the utility of derived metrics, calculated from observational data, for assessing heat and carbon update, and we end with a summary of the utility of metrics and the potential for new ones as new observational data become available. We have included, as supporting information, a host of other metrics that we considered but determined were ancillary to our main focus on heat and carbon uptake.

2. Metrics, Data, and Models

For our purposes, a metric is *any quantity or quantifiable pattern that summarizes a particular process or the response in a model to known forcings*. We focus here on key metrics associated with the sequestration of heat and CO_2 into the ocean that are directly observable (e.g., SST) or calculable (e.g., heat content).

In this report, we use the following terms:

1. *observational data* are quantities measured with a sensor (such as temperature, carbon content, humidity, pH, wind speed and direction, etc., including measurements from satellites), or based on direct measurements (such as heat content and wind stress, etc.) that are provided as time series at individual locations, or as 3-D or 4-D gridded products;
2. *state estimation data* are the output of a data-assimilating numerical model in which the differences between the model output and the observations are minimized over a multiyear assimilation window using four-dimensional variational assimilation (e.g., the Southern Ocean State Estimate, SOSE, Mazloff et al., 2010);
3. *reanalysis data* are like state estimation data, in that they are computed using variational assimilation, but the method is adapted for numerical weather prediction and uses a time window of a few days, allowing for tighter constraint to observations than in state estimation; and
4. *simulation data* are the output of numerical simulations in which all quantities are internally generated, except for the initial conditions, the internal parameters (including the model parameterization schemes), and the external parameters and boundary conditions such as: the size and rotation rate of the planet, the solar constant, and the quantities of radiatively important gases in the atmosphere (as in CMIP1-CMIP3) or the net radiative imbalance at the top of the atmosphere in 2100 (as in CMIP5).

Wherever possible, model simulations are compared to actual observations and the atlases based on them, but when observations are too sparse in space and/or time to form a complete picture, we rely on model-assisted state estimates and reanalyses as well as metrics derived from observations through parameterizations, bulk formulae, or other empirical relationships.

2.1. Observationally Based Sources

For this study, we compare ocean temperature-based metrics of the model simulations to the 2013 World Ocean Atlas (WOA13; Garcia et al., 2014; Locarnini et al., 2013; Zweng et al., 2013). Gridded versions are available through the National Oceanographic Data Center (NODC, https://www.nodc.noaa.gov/OC5/woa13/woa13data.html). Since 1982, the World Ocean Atlas (Levitus, 1982) has provided consistent, complete, up-to-date data on the ocean's temperature, salinity, and nutrients, compiled from satellite measurements, ship-based observations, profiling floats, and other sources.

Observations of the relevant species integral to quantifying the state of the oceanic carbon system—dissolved inorganic carbon (DIC) or total carbon (TCO_2 or ΣCO_2); alkalinity; pH and pCO_2—have been synthesized, standardized, and mapped into regularly gridded data sets, first in the Global Ocean Data Analysis Project (GLODAP, Key et al., 2004) and more recently as part of GLODAPv2 (Key et al., 2015, Lauvset et al., 2016, Olsen et al., 2016). GLODAPv2 serves as our observational benchmark for pH, calculated from total alkalinity and DIC; these data are available from CDIAC (http://cdiac.ornl.gov/oceans/GLODAPv2/).

Climate and atmospheric reanalyses, like the Climate Forecast System Reanalysis (CFSR, Saha et al., 2010), the National Center for Environmental Prediction (NCEP) reanalysis (Kalnay et al., 1996), the NCEP2 reanalysis (Kanamitsu et al., 2002), and the ECMWF (European Centre for Medium-Range Weather Forecasts) Reanalysis (ERA40, Uppala et al., 2005) all provide wind speeds and wind stress over the Southern Ocean as well as other atmospheric variables like cloud cover, surface air temperature, pressure, heat, and freshwater fluxes. For this study, the surface wind-stress data (which are calculable from the directly observed wind speed) are the CFSR annual mean quantities averaged from monthly data from January 1981 through December 2010.

Sea ice data can be obtained from the National Snow and Ice Data Center (NSIDC, ftp://sidads.colorado.edu/pub/DATASETS/NOAA/G02202_v2/). Sea ice variables are also included in several of the reanalyses including CFSR, NCEP, and NCEP2. Here we compare the model simulated sea ice area (fractional coverage) in the simulations to the NCEP2 reanalysis, although a comparison with the CFSR reanalysis was not substantially different.

The Southern Ocean State Estimate (SOSE, Mazloff et al., 2010, http://sose.ucsd.edu/) provides a least-square fit to all of the available ocean observations within an ocean general circulation model. The latest generation of this state estimate called Biogeochemical-SOSE (B-SOSE, Verdy & Mazloff, 2017), includes both physical and biogeochemical properties of the ocean (between January 2008 and December 2012) to determine the "best-fit" ocean state. B-SOSE, while currently available at a lower 1/3° resolution than SOSE, continues to improve, by incorporating a rapidly growing quantity of observational data and incorporating biogeochemical processes (Galbraith et al., 2015). We use B-SOSE velocities as our "observed" benchmark, but we re-emphasize that these velocities are really model estimates, and at 1/3° resolution, B-SOSE is still coarser than eddy-simulating, so observed velocities likely differ in eddy-rich areas. Supporting information Figure S13 includes data from the original 2006–2010 SOSE at the higher (1/6°) resolution.

2.2. Model Simulations

All model output used in this study was submitted as part of the IPCC-AR5/CMIP5 intercomparisons and were downloaded from the Program for Climate Model Diagnosis and Intercomparison (PCMDI, now the Earth System Grid Federation, ESGF, https://esgf-node.llnl.gov/projects/esgf-llnl/). Each of the simulations is for the "Historical" period and includes prescribed atmospheric CO_2 and all other the forcing mechanisms (radiative, aerosol, etc.), from ~1860 through 2005 (Taylor et al., 2012). The CMIP5 models include different ozone forcing fields ranging from prescribed to prognostic stratospheric ozone changes, resulting in different responses of the SH westerlies to any changes in stratospheric ozone. The model output presented here includes the average of all simulated months from January 1986 to December 2005, a 20 year annual average. When more than one member of an ensemble was submitted, we used only the first member. Five of the simulations we examine are from Earth System Models (ESM) that include explicit carbon cycling within and between the various climate system components (atmosphere, ocean, land, vegetation). The sixth simulation is from a coupled atmosphere-ocean global climate model in which the carbon is an external parameter and only affects the radiative properties of the atmosphere. References and websites for these model simulations are shown in Table 1.

Table 1
The Earth System Models and Coupled Climate Models Used in This Study

Model name	Modeling center	Source	Reference
CanESM2	Canadian Centre for Climate Modelling and Analysis	http://ec.gc.ca/ccmac-cccma/default.asp?lang=En&n=3701CEFE-1	Flato et al. (2000)
CSIRO-Mk3.6	Commonwealth Scientific and Industrial Research Organisation	https://confluence.csiro.au/public/CSIROMk360	Jeffrey et al. (2013), Rotstayn et al. (2012)
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	http://www.gfdl.noaa.gov/earth-system-model	Dunne et al. (2012, 2013)
HadGEM2-ES	UKMO Hadley Center	https://verc.enes.org/models/earthsystem-models/metoffice-hadley-centre/hadgem2-es	Collins et al. (2008)
MIROC-ESM	JAMSTEC, U. Tokyo, NIES	http://www.geosci-model-dev.net/4/845/2011/gmd-4-845-2011.html	Watanabe et al. (2011)
MRI-ESM1	Meteorological Research Institute	http://www.mri-jma.go.jp/Publish/Technical/DATA/VOL_64/tec_rep_mri_64.pdf	Yukimoto et al. (2011)

These six models were chosen based on some simple criteria. First and foremost, each had all of the variables that were needed to calculate the metrics we wished to present. Second, during the analysis of the simulations and the writing of this study, the PCMDI database suffered an outside hack and was unavailable for almost 12 calendar months; we were satisfied that the models for which we had data represented a reasonable cross section of the available models in that, as a group, these captured some features well while other features had large intermodel differences. This paper has been written to document the metrics rather than to present a complete intercomparison. Our original intent and future goal is to assess and publish a focused, thorough intercomparison of all the available CMIP5 simulations that have reported all of the necessary variables.

Careful attention must be paid to the processing of both the observationally based data and the model output in order to ensure a clean comparison. Several of the CMIP5 models used here (GFDL-ESM2M, MIROC-ESM, MRI-ESM1) have reported their output on nonstandard grids (that is, a grid in which latitude or longitude is not uniformly represented by a single model index), most often due to a curvilinear remapping in the native ocean model grid employed to address the singularity at the North Pole. This complicates quantitative comparisons of variables along standard sections (lines of latitude or longitude) and the creation of ensembles from multiple models. Care must be taken to ensure that information is not lost—especially for velocity values near coastlines. For these models, output was linearly interpolated (from all the valid nearby points) onto a regular 1°-by-1° grid (with FERRET, http://ferret.pmel.noaa.gov). Although interpolation of simulated fields to a standard grid makes the analysis more convenient, it makes closing of heat and carbon budgets challenging. Nonlinear terms, such as advective and parameterized fluxes of heat and biogeochemical quantities, need to be calculated on the native model grids and integrated in time.

3. Results

We now present several observationally based metrics associated with the Southern Ocean circulation, heat and carbon uptake, and biological recycling. Some metrics are pattern-based: individual features or groups of features demonstrate where data/model differences lie. Some metrics can be reduced to a single number: e.g., the strength of the ACC transport at Drake Passage or the latitude of the maximum zonal mean winds over the Southern Ocean. We leave our assessment of "derived" metrics—quantities that are not measured directly, but are calculated based on parameterizations, bulk formulae, or other empirical relationships—to the Discussion section. Additional metrics are also discussed in the supporting information.

3.1. Wind Stress

The strength, extent, and latitudinal position of the Southern Hemisphere surface westerlies (Figure 1) are crucial to the simulation of the circulation, vertical exchange and overturning, and heat and carbon fluxes over the Southern Ocean. The net transfer of wind energy to the ocean depends critically on the strength and latitudinal structure of the winds. Equatorward-shifted winds are less aligned with the latitudes of the

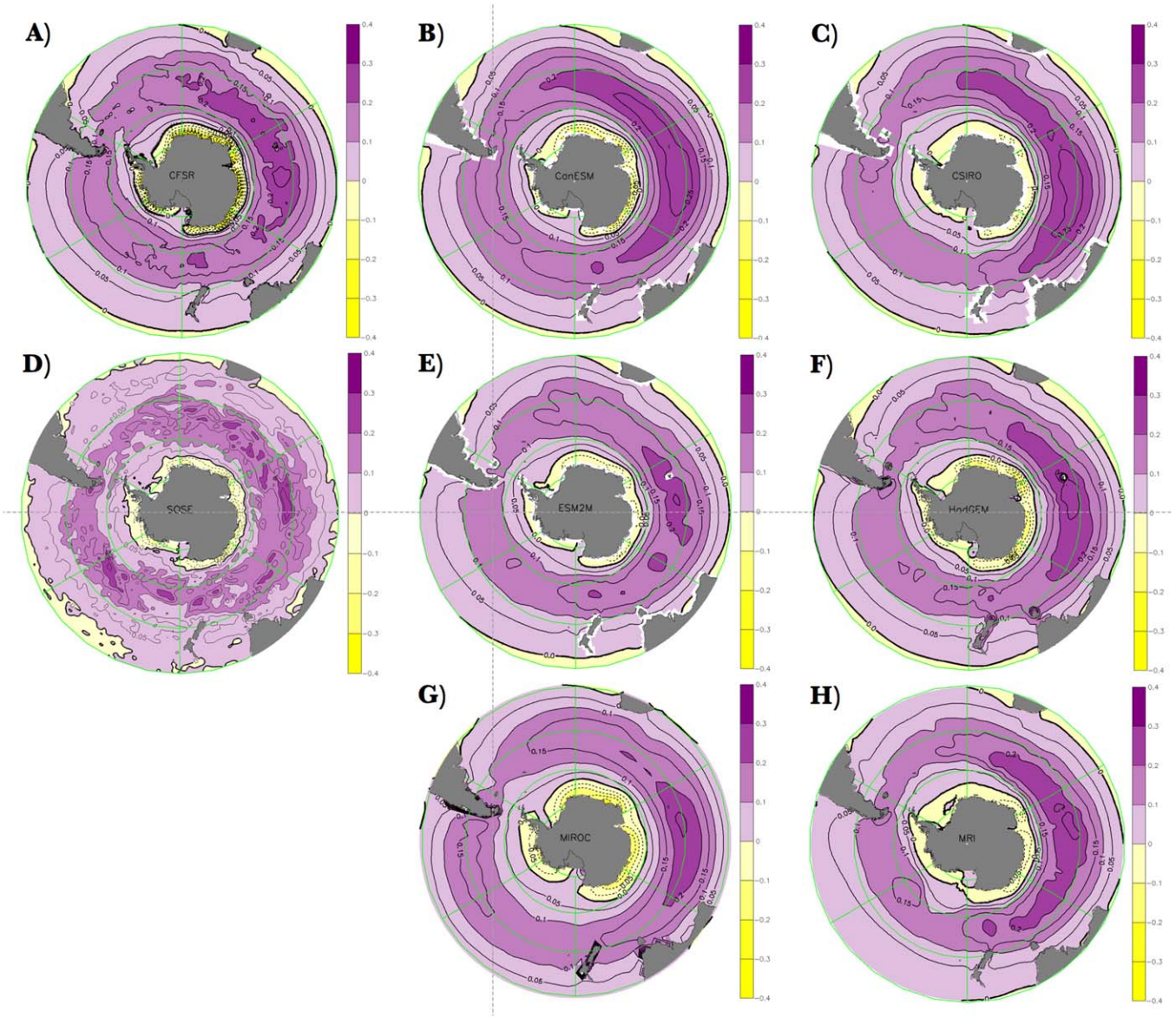

Figure 1. Annual-mean zonal wind stress (N/m^2) from reanalysis, state estimation, and several of the coupled climate models submitted as part of the CMIP5/IPCC-AR5 process. All model figures cover the simulated years 1986–2005 from the HISTORICAL forcing scenario. (a) CFSR reanalysis (1981–2010); (b) CanESM2; (c) CSIRO-Mk3.6; (d) B-SOSE (2008–2012); (e) GFDL-ESM2M; (f) HadGEM2-ES; (g) MIROC-ESM; and (h) MRI-ESM1. The MIROC-ESM model (Figure 1g) has a strong localized westerly patch in the southernmost Ross Sea at 78°S leading to the strong westerly stress at that latitude in Figure 2 (magenta line).

Drake Passage and are situated over shallower isopycnal surfaces, making them less effective at both driving the ACC and bringing dense deep water up to the surface. Although there are still competing hypotheses about what physics balance the momentum imparted by the winds over the ACC (e.g., topographic form drag, generation of upper ocean mixing, eddy radiation to other latitudes—see Ferrari & Wunsch, 2009), it is essential that simulations get the winds right. The Southern Hemisphere westerlies have intensified over recent decades, and this intensification will likely continue into the future with increasing greenhouse gas addition. Gent (2016) has recently summarized the expected impact on the meridional overturning in the Southern Ocean, pointing out the important opposing effect of eddy activity. Other previous studies that have assessed CMIP5 and/or CMIP3 simulations using various metrics related to the wind stress include Bracegirdle et al. (2013), Sen Gupta et al. (2009), and Meijers et al. (2012).

Stratification in the Southern Ocean, nominally the observed vertical density gradient, is a result of the balance between Ekman advection, surface buoyancy fluxes, and eddy-induced buoyancy fluxes (e.g., Marshall

& Radko, 2003; Marshall & Speer, 2012), so biases in any of these will lead to errors in the isopycnal slopes and water mass properties. Critically important for the heat and carbon uptake are the effects of the winds in driving the deep upwelling around Antarctica due to the surface divergence of the Ekman flow south of the wind stress maximum and surface convergence north of it. The strong upwelling and subsequent downwelling provide the pathway between the deep ocean and the atmosphere allowing for the sequestration of heat and carbon in the deep and the renewal of surface nutrients. As has been shown, the openness of this doorway to the deep ocean affects the overall sensitivity of coupled climate and earth system models to anthropogenic change (Russell et al., 2006b; Stouffer et al., 2006). Additionally, winds affect the heat and carbon uptake via sea ice export, and the strength of the Weddell and Ross gyres (Cheon et al., 2014; Holland & Kwok, 2012).

Figure 1 shows the monthly mean zonal wind stress (in N/m^2) for all months in the period of record: years January 1981 through December 2010 for the CFSR reanalysis data (our proxy for the observations); January 2008 through December 2012 for the B-SOSE assimilated data; and January 1986 through December 2005 for the CMIP5 simulations. The data are presented as a polar stereographic plot for each data set, and grid lines are added (in green) every 60° of longitude (at 0°, 60°E, 120°E, 180°, 120°W, and 60°W) and every 15° of latitude (at 75°S, 60°S, 45°S, and 30°S).

Figure 2 shows the zonal mean of the plots in Figure 1 at all longitudes over the ocean only. As noted in the fourth and fifth IPCC assessments, most of the models have their maximum zonal wind stress equatorward of the reanalysis and this systematic bias may affect the models' response to increasing radiative forcing (Flato et al., 2013; Randall et al., 2007).

As in the reanalysis (Figure 1a), all simulations have the strongest surface westerlies in the Indian sector of the Southern Ocean between 55°S and 45°S, although each simulation has its zonal-mean zonal wind stress northward (equatorward) of the observations (Figure 2). The zonal extent of the maximum zonal winds (dark purple, ≥ 0.2 N/m^2) has significant intermodel variability, although most simulations have a local

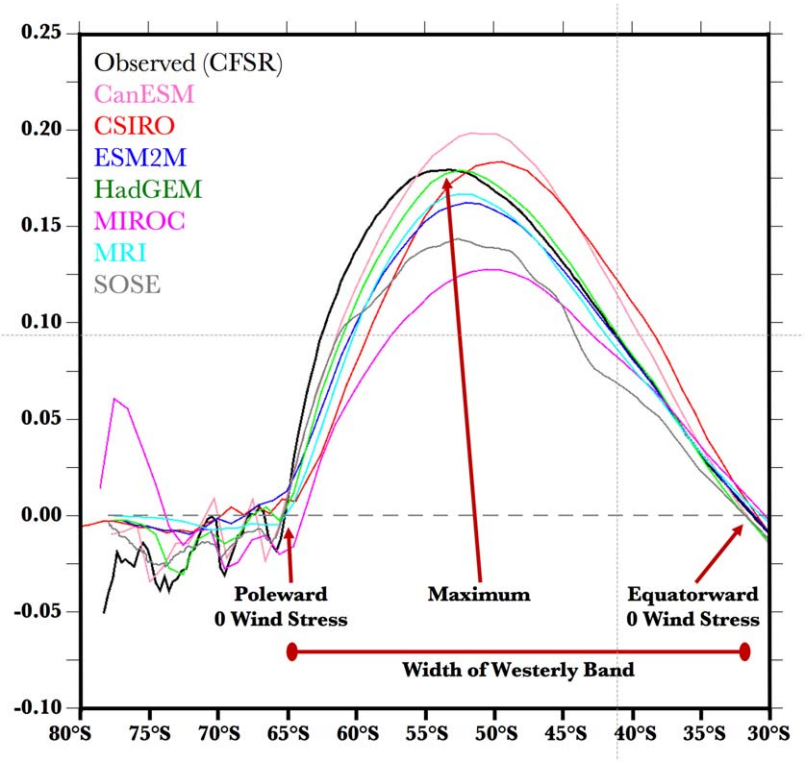

Figure 2. The zonal and annual means of the zonal wind stress (N/m^2) for each of the eight maps in Figure 1—note that each of the model simulations (colors) and B-SOSE (gray) have the peak wind stress equatorward of the observations (black). Also shown are the latitudes of the observed "poleward zero wind stress" and the "equatorward zero wind stress" which delineate the "width of the westerly band" discussed in the text and highlighted in Table 2.

Table 2
Wind, Heat Uptake, and Carbon Uptake-Related Metrics

	Maximum zonal-mean, zonal wind stress (N/m^2)	Latitude of maximum stress (°S)	Antarctic circumpolar current (Sv)	Latitude (°S) of 0 zonal wind stress (subtropical)	Latitude of (°S) 0 zonal wind stress (subtropical)	Width of SH westerly band (° latitude)	Annual heat uptake (S of 30°S, PW) (*derived*)	Annual carbon uptake (S of 30°S, Pg) (*derived*)	Annual carbon uptake (global, Pg) (*derived*)	Fraction in Southern Ocean (*derived*)
Observed	**0.1794**	**53.23**	**173.3 ± 10.7**	**67.69**	**31.26**	**36.43**	***0.19 ± 0.07***	***0.9***	***2.00***	***45.0%***
CanESM	0.1986	51.63	154.8	67.89	31.68	36.21	−0.23	0.74	1.66	44.8%
CSIRO	0.1837	49.43	110.2	69.69	30.97	38.72	−0.37	**N/A**	**N/A**	**N/A**
ESM2M	0.1623	52.00	133.7	67.98	31.17	36.81	−0.22	0.67	2.05	32.6%
HadGEM2	0.1792	52.50	172.1	67.51	31.80	33.51	0.09	1.09	2.00	54.5%
MIROC	0.1278	50.50	179.1	63.79	30.38	33.41	0.28	0.97	1.93	50.6%
MRI	0.1667	52.50	113.2	65.09	30.63	34.46	−0.06	0.88	1.98	44.7%
B-SOSE	0.1434	52.71	182.2	65.26	31.80	33.46	0.46	0.76	**N/A**	**N/A**

Note. Bold numbers are observed from direct measurements; bold and italics numbers are derived from direct observations through standard parameterizations and/or bulk formulae.

maximum in the wind stress immediately south of New Zealand (at 55°S, 170°E). Each model has easterly winds along the Antarctic coast that are less intense than in the CFSR reanalysis. The B-SOSE assimilation has a less intense maximum in both magnitude and extent than the observations or any of the model simulations except for MIROC-ESM. Note that the equatorward zero zonal wind stress line is fairly consistent across data sets, but the width of the westerly band varies due to variations in the poleward extent of this band, and this has consequences for heat and carbon uptake (see Table 2).

3.2. Fronts: Polar and Subantarctic

The main fronts in the Southern Ocean, the Polar Front in the midst of the Antarctic Circumpolar Current, and the Subantarctic Front marking its northern extent, are indicated by sharp gradients in temperature and nutrients. Most of the transport of the ACC occurs along these two frontal boundaries, and they are essential for the upwelling of nutrients (and ecosystem dynamics) and the uptake of heat and carbon. For our purposes, we chose not to assess the other main fronts in the Southern Ocean, specifically the Subtropical Front that separates the colder southern water from the subtropical gyre to the north, or the Southern ACC Front and Southern Boundary that mark the southern side of the ACC and that are coincident with the winter sea ice edge in many sectors. The Subantarctic and Polar Fronts are marked by steep isopycnal surfaces roughly separating the surface waters from the intermediate waters and the intermediate waters from the deep waters of the Antarctic (Orsi et al., 1995). The locations of these fronts are tied both to the varying sea surface height field and to the bottom topography. Frontal locations are pertinent to the ongoing discussion of the importance of eddies to the meridional heat transport across and the meridional overturning of the ACC (Marshall & Speer, 2012).

There have been several recent efforts to identify the Polar and/or Subantarctic Fronts in the Southern Ocean from satellite data of sea surface height (Gille, 2014; Shao et al., 2015; Sokolov & Rintoul, 2007, 2009) and of sea surface temperature (Dong et al., 2006; Freeman & Lovenduski, 2016; Kostianoy et al., 2004; Moore et al., 1999), with some resulting differences due to methodologies and assumptions. Data and model simulations can be compared to each other directly—once a definition of location of the front is agreed upon. As the fronts mark locations of steep temperature gradients and more intense vertical motion, any latitudinal offset will affect the local air-sea temperature differences and has the potential to bias the net heat and carbon fluxes.

The definitions of the fronts used here (see Figure 3 caption) are simplified versions derived from Orsi et al. (1995, cf.)—who used hydrographic *subsurface* measurements to identify the locations of the steeply sloping isopycnals that separate water masses. Moving northward, we define the Polar Front as the location where the minimum temperature in the top 200 m exceeds 2°C, and the Subantarctic Front as the poleward location where the temperature at 400 m exceeds 4°C. We choose these definitions as an easy way to indicate the wide disparity among the model simulations. A nonobservationally based metric, relating the position of the ACC core to flow-weighted latitude, is presented by Sen Gupta et al. (2009): while this metric does present a concise way to compare simulations to each other, we chose an observationally based metric.

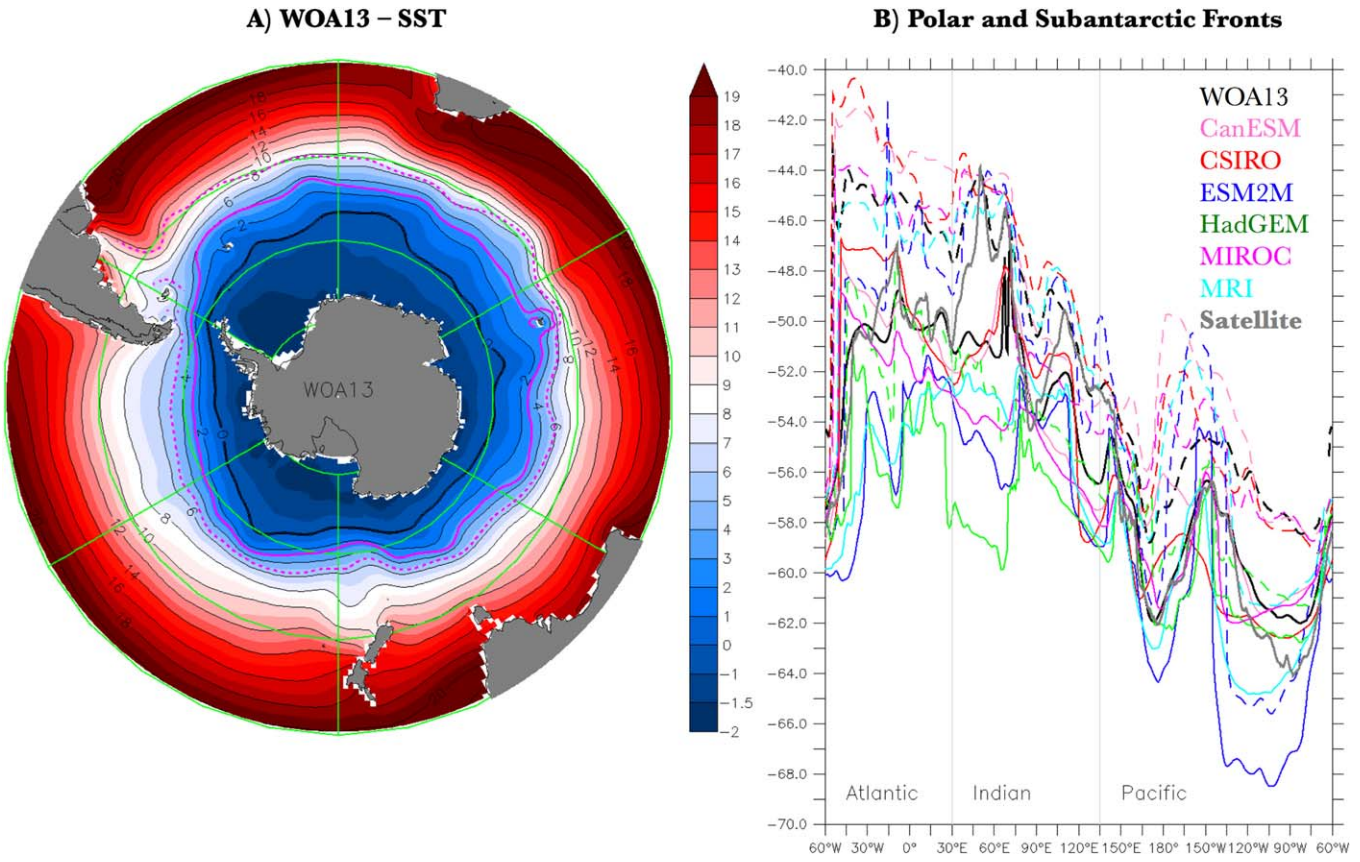

Figure 3. (a) The observed, annual mean surface temperature (WOA13, 0–50 m average); and (b) the locations of the Polar Front (solid) and the Subantarctic Front (dashed) as determined from the observations and model simulations. Also shown in Figure 3b (in gray) is the average weekly location of the Polar Front between 2002 and 2014 as determined from satellite (Freeman & Lovenduski, 2016; referred to as FL2016); their data are available at: http://store.pangaea.de/Publications/FreemanN-LovenduskiNS_2015/Polar_Front_weekly.nc). The front locations from the WOA13 are indicated in magenta in Figure 3a. All model figures cover the simulated years 1986–2005 from the HISTORICAL forcing scenario. Using definitions from Orsi et al (1995), the Polar Front is defined here as the poleward location of the 2°C isotherm of the temperature minimum between 0 and 200 m; the Subantarctic Front is defined here as the poleward location of the 4°C isotherm at 400 m. FL2016 use the observed temperature gradient at the surface to define the Polar Front. In general, the thick black line (WOA13) and the thick gray line (FL2016) are close to each other, except in the Western Indian around the Kerguelen Plateau (~60°–80°E) where the satellite data have it north of the plateau while the ocean temperature definition has it south of the plateau.

The frontal positions are clearly tied to the bathymetry (see supporting information Figure S14), most notably the southward dip just west of the Campbell Plateau (centered at ~170°E) and the northward turn just east of it. All of the data sets show the northward excursion of both fronts immediately downstream of the Drake Passage and the general shift southward across the Indian and the Pacific. There are clearly large differences between the simulations, some of which are tied to the underlying topography (CSIRO generally, see supporting information Figure SXX) and some to the temperature-based definition of the fronts: e.g., ESM2M in the South Pacific centered around 120°W where the Subantarctic Front, defined as the poleward location of the 4°C isotherm at 400 m, is significantly southward due to the overall warm bias of the simulation.

3.3. ACC Transport

Observations of the hydrographic properties along a section are the main way we determine the transport in the ocean across that section. Accordingly, measurements of the temperature and salinity throughout the water column are the "bedrock" on which our understanding of the ocean and its circulations rests. Transports, however, are not directly measurable: transport calculations rely on certain assumptions (especially a "level of known motion") that are difficult to assess in places (e.g., the ACC). The study by Donohue et al. (2016) reported the barotropic and baroclinic transport in Drake Passage derived from the CPIES array. A model-based metric of the ACC transport reported by Meijers et al. (2012) uses the barotropic transport

across Drake Passage, but the transport is not observable in most places. Generally, we rely on model simulations to test and/or verify our transport estimates. This situation is complicated by the fact that motion generally occurs along isopycnal surfaces; in locations where surfaces are sloping, the horizontal flow or property transport we calculate often has a nonzero vertical component associated with it. Last, true vertical motion due to wind-induced convergence or divergence at the surface and diapycnal mixing are even further removed from the direct observations we can make, putting even more importance on simulations.

The transport of the Antarctic Circumpolar Current (ACC) through Drake Passage is a critical metric for SO simulations. By changing the volumes and properties of ventilated waters, the strength and location of the ACC impact both the properties and the meridional overturning circulation of the Southern Ocean (Böning et al., 2008; Dufour et al., 2013; Heuzé et al., 2015). The ACC transport depends on several different simulation features: strength and position of the Southern Hemisphere westerlies, meridional isopycnal slope, salt import from the Atlantic across 30°S via North Atlantic Deep Water production, net heat flux gradient across the current, etc. (Russell et al., 2006a), so it is a reasonable measure of the overall simulation of the SO.

However, even this "simple" metric has been hard to verify observationally; previous estimates of the baroclinic ACC transport have generally been similar: 134 ± 11.2 Sv by Whitworth & Peterson (1985) (1 Sv = 10^6 m^3/s); a mean of 137 ± 8 Sv with variability of ±27 Sv by Cunningham et al. (2003) and Meredith et al. (2011); 127.7 ±1.0 Sv with a standard deviation of 8.1 Sv by Chidichimo et al. (2014). The most recent estimate (Donohue et al., 2016), based on direct measurements of the barotropic bottom transport with the cDrake array (Chereskin et al., 2012), reports a total ACC transport of 173.3 ±10.7 Sv, comprised of a

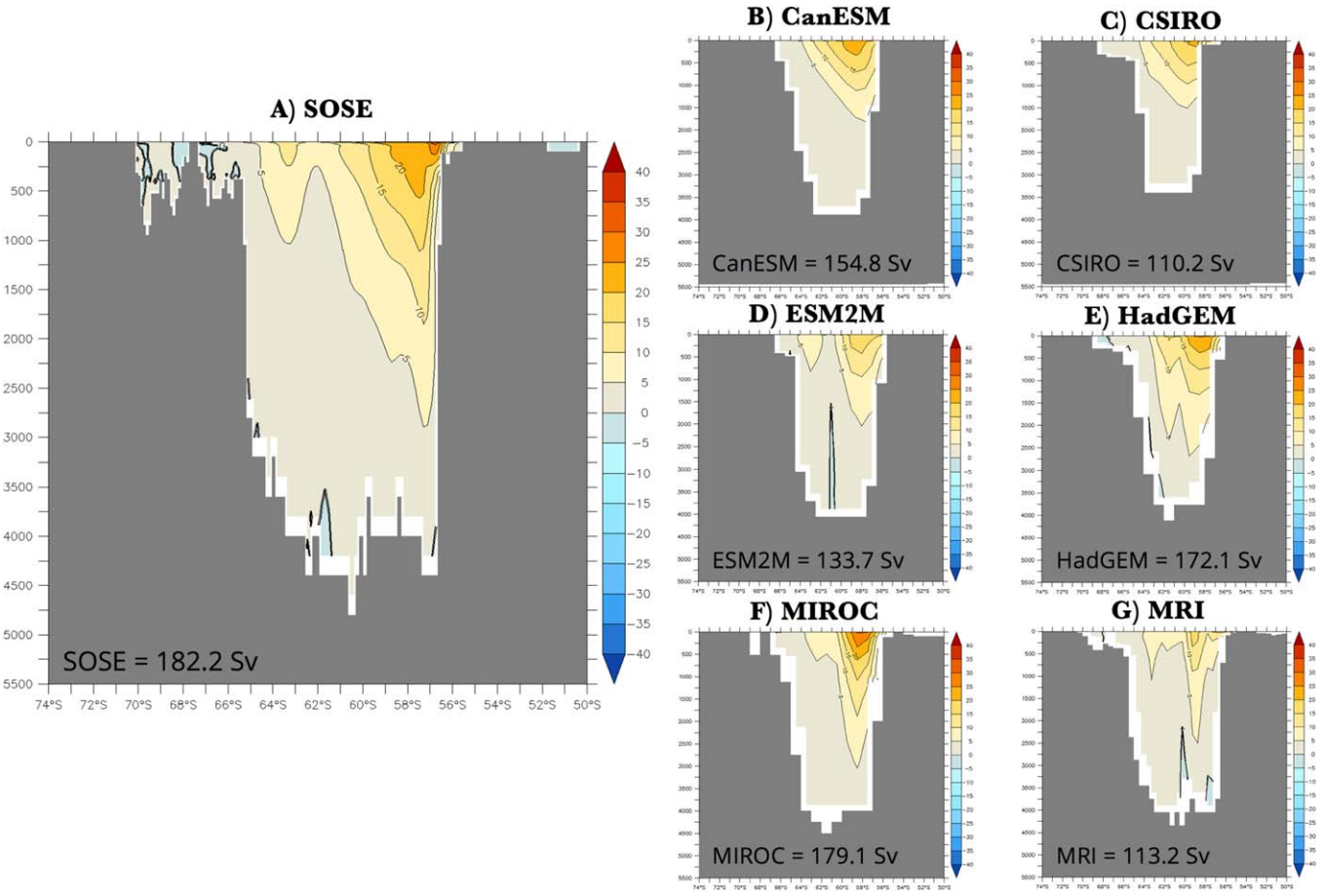

Figure 4. This figure shows the zonal velocity (in cm/s, positive is eastward—out of the page) through Drake Passage (at 69°W) from the observations and the model simulations. The standard-resolution B-SOSE reanalysis (1/3°, Verdy & Mazloff, 2017) shows the clear frontal structures that characterize the complex surface flow in the ACC. The total transport by the ACC is noted in black in the lower left portion of each panel. All model figures cover the simulated years 1986–2005 from the HISTORICAL forcing scenario. (a) B-SOSE (2008–2012); (b) CanESM2; (c) CSIRO-Mk3.6; (d) GFDL-ESM2M; (e) HadGEM2-ES; (f) MIROC-ESM; and (g) MRI-ESM1.

baroclinic transport of 127.7 ±5.9 Sv and near-bottom velocities contributing 45.6 ± 8.9 Sv distributed throughout the water column.

The net transport through Drake Passage in model simulations is simply the surface to bottom integral of the zonal velocity between the tips of the Antarctic Peninsula and South America (Figure 4). Here, for each model, we integrate along the longitude closest to 69°W (without interpolation). As noted by Russell et al. (2006a) and others, there are a host of reasons why a simulation can have the wrong ACC transport. The leading problems are: (1) winds that are too weak and too far equatorward; (2) an incorrect gradient in air-sea fluxes of heat and freshwater over the ACC; (3) not enough salt being exported from the Atlantic as part of NADW that upwells on the southern side of the ACC; (4) incorrect bottom topography; and (5) errors in the near-bottom flows. Biases in the isopycnal slopes, which are partly explained by uncertainties in eddy parameterizations, can also explain biases in the simulated ACC strength. Biases in the details of the flow into and out of Drake Passage (especially its vertical structure) and stratification will lead to biases in the zonal heat/carbon transport and interbasin communication.

In our CMIP5 model subset, models with more wind over the latitudes of Drake Passage generally have a stronger ACC, although MIROC is an outlier—it has weak, equatorward-shifted winds (Table 2 & Figure 1), yet has the strongest ACC of the models presented. The MIROC simulation highlights an important point: any individual metric cannot determine the "realism" of a simulation; only in combination can analysis of standardized metrics assure us that a model is getting "the right answer for the right reason." In coarse resolution models discussed here, the ACC transport is also sensitive to the parameterization of the eddy-induced momentum dissipation ("eddy viscosity"), which tends to be smaller in numerical simulations with higher horizontal resolution. Note that the 1/3° resolution B-SOSE reanalysis shows the clear frontal structures that characterize the complex surface flow in the ACC. This structure is poorly represented by the coarser resolution (∼1° meridional) simulations discussed here.

3.4. Sea-Ice

Sea-ice concentration and extent have a first-order role in setting ocean surface freshwater and heat fluxes, as well as the global sea ice-albedo feedback. Sea ice around Antarctica grows outward from the continent every winter and retreats back every summer. Figure 5 shows the maximum and minimum extent from the NCEP2 reanalysis and the various models. Seasonality of sea ice and its thickness influence deep and bottom water formation via brine-rejection processes and via open-ocean convection (e.g., Heuzé et al., 2013), with implications for the heat and carbon cycles (e.g., Bernardello et al., 2014; Bitz et al., 2006). Sea-ice can influence the air/sea exchange of heat and carbon by influencing the midlatitude jet stream (Kidston et al., 2011), ocean circulation or ACC transport (Hogg, 2010), mode/intermediate water mass formation, and ocean stratification and overturning circulation (Abernathy et al., 2016). More directly, sea-ice is a physical barrier to air/sea exchange, and it impacts light-limited phytoplankton photosynthesis. In Southern Ocean primary production algorithms, less surface ice coverage in a particular year results in more biological productivity (Arrigo et al., 2008), enhancing the biological ocean carbon uptake. The new generation of Argo floats with biogeochemical sensors (papers in this issue) will strongly enhance our ability to estimate the biological carbon pump and the sea-ice-biology link.

Sea-ice extent is well-measured from satellites, although there is an uncertainty associated with concentrations less than 15% per pixel. Ice thickness is much more difficult to infer from satellites. NSIDC, NASA, ASPECT all have direct observations of sea ice extent at both poles and the various reanalyses (CFSR, NCEP, and NCEP2) have reasonable representations of the annual cycle.

The quality of sea-ice models has improved dramatically over the last decades (Turner et al., 2013; Zunz et al., 2013). Biases in sea-ice can reflect wind and ocean-current-driven transport anomalies and heat flux anomalies in the ocean and atmosphere, as well as biases that are intrinsic to the sea ice component formulations. The simulation of sea ice in the models shown in Figure 5 differs widely: the ESM2M model has significantly less ice than observed in all months with virtually no ice in the summer, while the MRI model grows its ice too quickly in fall, and the CSIRO model does not melt back far enough or fast enough in spring. Ice thickness is highly correlated with ice extent (more so in winter), and most of the models have plausible distributions of ice thickness.

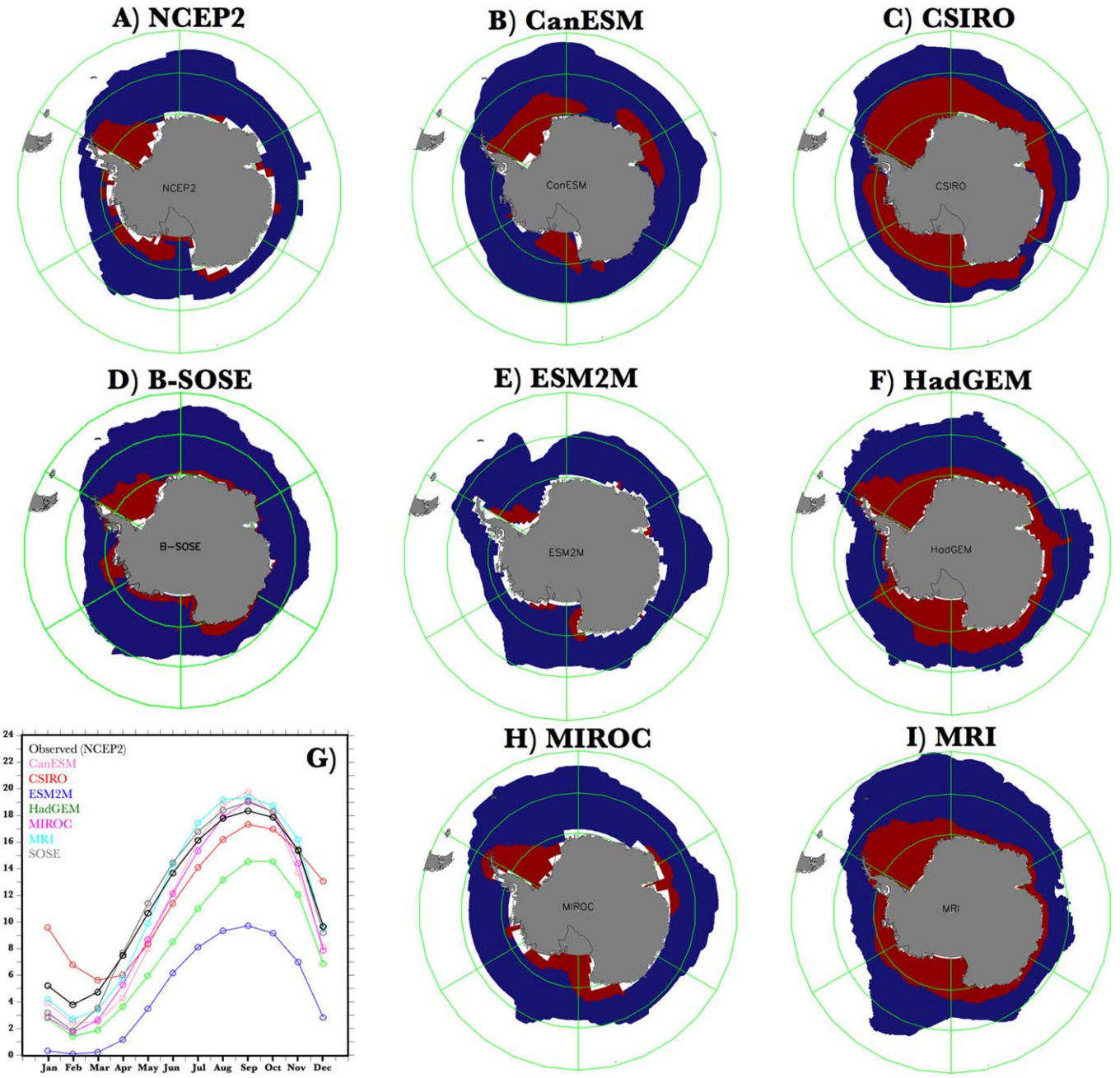

Figure 5. Annual sea ice extent (maximum in blue, minimum in red) from reanalysis, state estimation, and several of the coupled climate models submitted as part of the CMIP5/IPCC-AR5 process. The blue region is the mean extent for September, the time of maximum sea ice coverage in the observations, and the red areas are the mean extent in February, the time of minimum coverage in the observations. The edge of full coverage is defined by the 15% areal coverage contour—below 15%, the satellites have a hard time determining the difference between ice and open water (NSIDC). Several of the models have their maximum or minimum coverage either a month before or a month after the month shown here (e.g., HadGEM has its maximum coverage in October—see Figure 5g) All model figures cover the simulated years 1986–2005 from the HISTORICAL forcing scenario. (a) NCEP2 data set; (b) CanESM2; (c) CSIRO-Mk3.6; (d) B-SOSE; (e) GFDL-ESM2M; (f) HadGEM2-ES; (h) MIROC-ESM; and (i) MRI-ESM1. Also included for reference is Figure 5g, the annual cycle of sea ice area (in 10^6 km^2) from each of the eight plots. Sea ice in the CFSR reanalysis (not shown) is quite similar to NCEP2.

3.5. Lateral Exchanges at the Northern Boundary

Ideally, we want to quantify the mean and fluxes for a given property in a given region in order to create a *budget*: for the Southern Ocean, this reduces to the air/sea flux, the land/sea flux, and the transport across 30°S, along with the observed changes in the total or mean quantity (e.g., heat, fresh water, momentum,

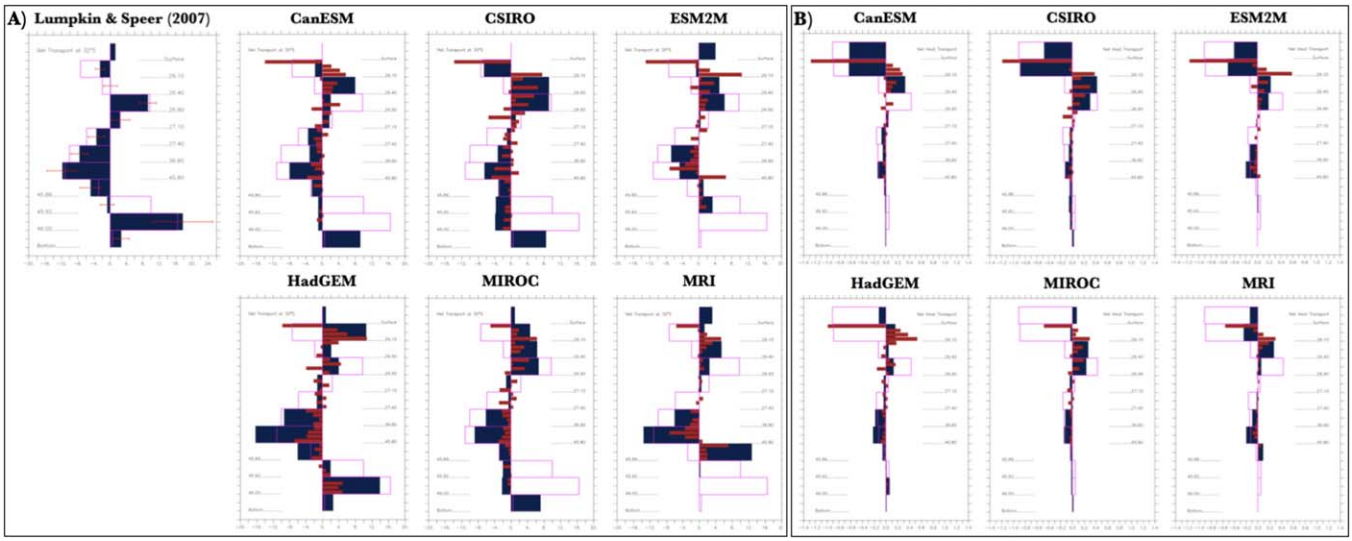

Figure 6. Density layer based (a) volume transport (in Sv = 10^6 m^3/s); and (b) heat/temperature transport (in PW/PWT) across 30°S based on the layer definitions in Talley (2008). The comparison with Talley (2008) assumes that the model's water mass characteristics are similar to those observed in the real ocean, however, in some cases this leads to a somewhat incomplete representation of the computed transports from the model simulation. The dark blue bars are the integrated totals for each layer and can be compared to the magenta lines, which are the observed values from Talley (2008). The narrower red bars are equal subdivisions of each blue layer. Unlike the Talley (2008) analysis, we have not separated out the Ekman transport (although we corrected the "observed" quantities (in purple) to reflect this, by adding the Ekman transport to the Surface Layer transport). The topmost blue bar in Figure 6a indicates the net transport across 30°S—note that both ESM2M and MRI have a significant net northward flow across 30°S. The topmost blue bar in Figure 6b indicates the total heat transport across 30°S. Positive values are northward transport (out of the Southern Ocean). Also included in Figure 6a are the transports across 32°S using the inverse methods described in Lumpkin and Speer (2007), along with their associated uncertainty for each layer. Note that there may be differences due to the exact location of the sections for each basin between the models (all at 30°S), Lumpkin and Speer (2007, all at 32°S) and Talley (2008, 32°S in the Atlantic and Indian, and 28°S in the Pacific).

carbon, etc.). Although the large-scale surface fluxes of heat and fresh water are not directly observable, the volume, heat, and tracer exchanges between the Southern Ocean and other ocean basins are powerful metrics for constraining budgets. These metrics are created with an "inverse-type" analysis where information on flow volume (shown in Figure 6a) or heat (Figure 6b) is mapped onto density layers and then calculated across various sections. An inverse analysis, where flow and tracer quantities are assigned to finely divided density layers is preferable in the climate context we are assessing because it is the net transport of heat and carbon out of the Southern Ocean that matters, as opposed to the classical water mass analysis preferred in a purely oceanographic setting in which water masses are defined based on observed hydrographic properties (e.g., defining Antarctic Intermediate Water as the salinity minimum as in Sallée et al., 2013a, 2013b) that determine how a particular water mass is being altered.

Observations at 30°S, which crosses the subtropical gyres in all three oceans, derived from hydrography and tracer studies, indicate four main depth regimes: southward flow at the surface, northward flow of mode and intermediate waters, southward transport of deep water, and northward transport of bottom water. Observations also indicate a net southward flow of heat into the Southern Ocean, primarily in the Indian Ocean sector, compensated somewhat by the northward Atlantic heat transport. This flow analysis reveals strengths and weaknesses of each simulation with respect to both flow and hydrography.

Hydrographic data offer a means to determine large-scale internal ocean velocities. Volume and heat fluxes depend on a reference velocity estimate, which is often estimated by inverse analysis (e.g., Macdonald & Wunsch, 1996; Ganachaud & Wunsch, 2000). Assimilation systems such as B-SOSE could potentially serve as a basis for these sorts of metrics, but thus far suffer from inadequacies in the models at the basis of the state estimates, particularly in the deep ocean. Calculating these metrics requires great care. Monthly nonlinear terms must be calculated from instantaneous (rather than monthly mean) velocity/temperature values. To ensure a meaningful interpretation, this metric needs to be provided as a standard model diagnostic and on the native model grid. Care must be taken when the native model grid has its temperature and salinity grid points at different locations than its velocity grid points (Arakawa B-grids and C-grids, Arakawa & Lamb, 1977): regridding of data—especially velocity—should be used as sparingly as possible, so that net transports are only minimally affected: analytical

errors are minimized when the volume and heat fluxes are computed using the same formulation that the model itself uses.

For the layer volume transports, we first "sort" the grid points at 30°S by their density (σ_θ, σ_2, and σ_4) according to the layer definitions provided by Talley (2008). This can also be carried in neutral density coordinates, as is common in recent transport analyses. Extra attention must be paid at the transitions between the density ranges to ensure that no grid points are missed and/or double counted. The layered volume transports are then just the zonal and depth integrated velocity in each layer given in 10^6 m^3/s (Sv). The layer heat/temperature transports are calculated in a similar fashion: the velocity, temperature (in °C), in situ density (in kg/m^3), and specific heat (J/kg/°) are multiplied together and integrated; the heat transport is given in PW (10^{15} J/s). In the future, it would be preferable to have the models record the simulated explicit and parameterized advective heat and carbon transports, so that true budgets could be determined.

The models simulate the four-layer flow pattern (Figure 6a), although MIROC has a net northward flow in the surface layer. The flows at all levels tend to be weaker than the observations, although this analysis could be hiding relatively strong flows in the same density class that cancel each other out.

Equally important is the realistic stratification at the northern boundary of the Southern Ocean. Biases in the isopycnal depths and in temperature at these isopycnals will lead to errors in the heat exchanges even if the circulation in pressure coordinates is generally realistic. Our analysis (see supporting information Figure S9) demonstrates that models tend to be too dense below 1,000 m and too light above 1,000 m, with the largest biases found in the Indian and Pacific sectors. These density biases complicate the definition of the four layers used in this section. Sallée et al. (2013a, 2013b) explicitly address the issue of density bias on calculated transports across 30°S using dynamically linked water mass definitions.

Northward transport within the intermediate layer is simulated reasonably well; all models, however, tend to put a large part of the transport into light density classes, which is consistent with the density bias discussed above. In contrast, all models (with the exception of HadGEM) underestimate the southward volume transport of the deep water. Given the fact that most of these simulations form a realistic amount of the North Atlantic Deep Water (NADW), this bias indicates an unrealistic "short-circuiting" of NADW due to excessive midlatitude upwelling or excessive flow of NADW directly into Indo-Pacific Ocean. ESM2M and MRI have virtually no bottom water in the right density range. Curiously, both ESM2M and MRI have northward nonzero net transports across 30°S, of 4 and 2 Sv, respectively. While a nonzero net transport could indicate an excess of precipitation over evaporation south of 30°S, flows of this magnitude are larger than expected. Parameterized transport terms, not considered here, could possibly account for the differences.

The models each underestimate the southward heat transport across 30°S: the observed transport is 0.9 ± 0.3 PW (Talley, 2008); CanESM has about 66% of that (0.6 PW), CSIRO has 50%, ESM2M has 33% and HadGEM and MRI have about 10%. Warm bias in the deep water, particularly in the 36.8–45.6 density classes, partially compensates for too weak a volume transport, reducing the biases in the heat transport. MIROC has a decidedly unrealistic net northward transport of heat out of the Southern Ocean.

3.6. CO_2 Flux

Of increased focus in the Southern Ocean are the biogeochemical responses and feedbacks to increases in atmospheric CO_2 and resulting changes in ocean physical forcing. Biogeochemical metrics that dovetail into the physical metrics described above are transport of nutrients and carbon across 30°S, CO_2 surface fluxes, and surface inorganic carbon parameters, in particular pH and calcium carbonate (aragonite/calcite) saturation states. Data-based validation of biogeochemical models is hampered by a dearth of data in this region and significant data-interpolation errors: statistical approaches to fill in the large spatial and temporal gaps have uncertainties and challenges (Palmer et al., 2010). With the increasing focus on the role of the Southern Ocean in controlling the global cycles of carbon and nutrients (e.g., Bernardello et al., 2014; Marinov et al., 2006), and feedbacks on climate, a large increase in Southern Ocean data gathering is occurring aided by autonomous measurements, primarily from profiling floats (Johnson et al., 2017), but also from ships of opportunity (Munro et al., 2015). Improvements in data-based approaches such as multilinear regressions and neural networks are providing improved fields for robust model-data comparisons. The metrics below illustrate the first steps on a path to rapidly improving assessment tools.

The CMIP5 model-average estimates that 43% (±3%) of the accumulated global oceanic uptake of anthropogenic carbon since the mid-nineteenth century has occurred in the Southern Ocean (Frölicher et al.,

2015). Note that the uptake of "anthropogenic carbon" is not an observable quantity, but the fraction of carbon entering the ocean due to the anthropogenic overburden can be derived based on other relationships and other properties, such as transient tracers (Gruber et al., 1996; Khatiwala et al., 2013; Sabine et al., 2004). In the preindustrial era, the Southern Ocean was clearly a source of carbon to the atmosphere due to the exposure of old, carbon-enriched deep water to the overlying atmosphere. Due to the increasing atmospheric burden, subduction along the Subantarctic Front (see Figure 3) has become a significant sink for carbon (anthropogenic and/or natural). Simulating this uptake accurately depends on the physical circulation, the wind position, the wind stress (vis-à-vis the gas exchange rate) and the simulated ocean carbon system: specifically, the volume of deep water that is exposed to the modern surface atmosphere and how much carbon it contains will strongly affect the magnitude of the Southern Ocean sink.

The LDEO Surface Ocean CO_2 data set, available through CDIAC or directly at (http://www.ldeo.columbia.edu/res/pi/CO2/), has nearly all of the ship-based and other measurements of ocean surface pCO_2, standardized to a common year and quality controlled. The air/sea pCO_2 difference, along with the NCEP2 wind speed climatology for the calculation of the gas exchange rate, has been used to determine the net flux of carbon through the air/sea interface for each month (Takahashi et al., 2009, gridded data downloaded from http://www.ldeo.columbia.edu/res/pi/CO2/carbondioxide/pages/air_sea_flux_2010.html). The anthropogenic CO_2 uptake from this data set is 2.0 Pg/yr of carbon globally, in line with estimates of the global

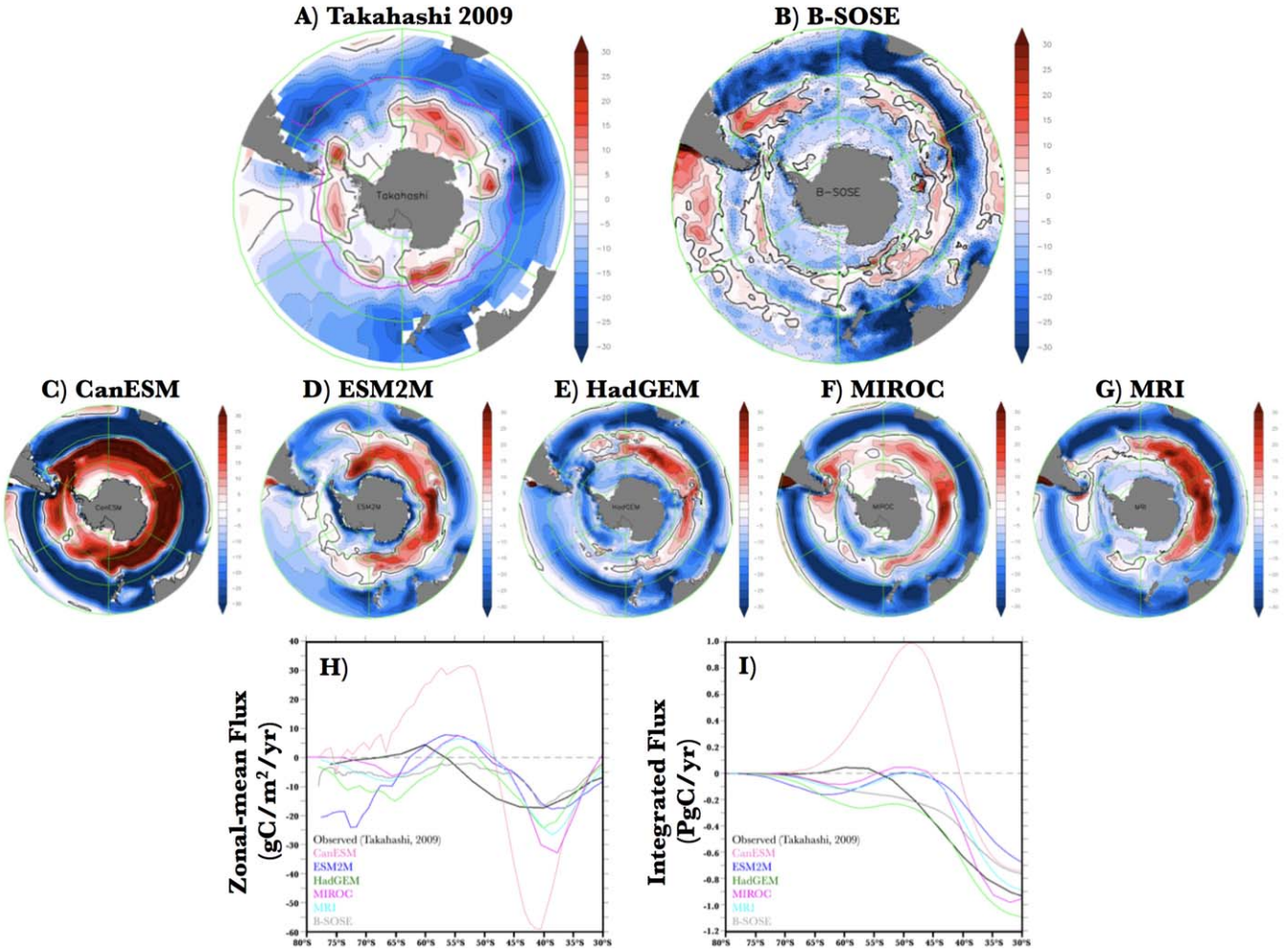

Figure 7. Annual mean CO_2 flux *(sea to air, gC/m^2/yr, positive (red) is out of the ocean)*. The observed fluxes are taken from Takahashi et al. (2009) climatology for reference year 1995, available through CDIAC, with a fairly coarse 4°-by-5° resolution. Model output for years 1986–2005 from the HISTORICAL forcing scenario. (a) Takahashi climatology (from NCEI); (b) B-SOSE; (c) CanESM2; (d) GFDL-ESM2M; (e) HadGEM2-ES; (f) MIROC-ESM; and (g) MRI-ESM1. Figure 7h is the zonal mean of each of the data sets (in gC/m^2/yr); and Figure 7i is the cumulative integral of the net CO_2 flux from 90°S to 30°S (in PgC/yr). The CanESM2 simulation has an order-of-magnitude more uptake and outgassing than the other simulations. The location of the Subantarctic Front (as seen in Figure 3) from the WOA13 is shown in magenta in Figure 7a.

oceanic uptake from the late 1990s and early 2000s (2.2 Pg/yr in the 1990's from Ciais et al., 2013, See Table 2). Interestingly, 0.9 Pg/yr (or 45%) of this uptake occurs south of 30°S, in agreement with the model based estimates (43% ± 3% for the CMIP5 ensemble mean). The climatological results in the Southern Ocean are in line with a neural network based CO_2 flux estimate of Landschützer et al. (2015).

Models simulate large areas of ingassing north of ~40°S and large areas of outgassing south of ~40°S. Relative to the Takahashi data set, the general patterns are consistent, but simulations have both too much ingassing and outgassing, with the outgassing band extending further north across models (Figure 7). The weaker signal in the Takahashi et al. (2009) data set may be due to the fact that it is the reconstructed flux in 1995; the models show the 1986–2005 average, but are more heavily weighted to the latter decade due to the accelerating rate of atmospheric CO_2 growth.

The advent of biogeochemically sensored floats is likely to increase our understanding of carbon cycling in the Southern Ocean due to both biology and solubility, and biogeochemical floats are providing unprecedented year-round data. This information will provide better targets against which to compare the simulations.

3.7. pH

Ocean acidification, the decrease in oceanic pH due to the absorption of CO_2, is an acknowledged and growing concern. Southern Ocean acidification is projected to lead to aragonite undersaturation in select

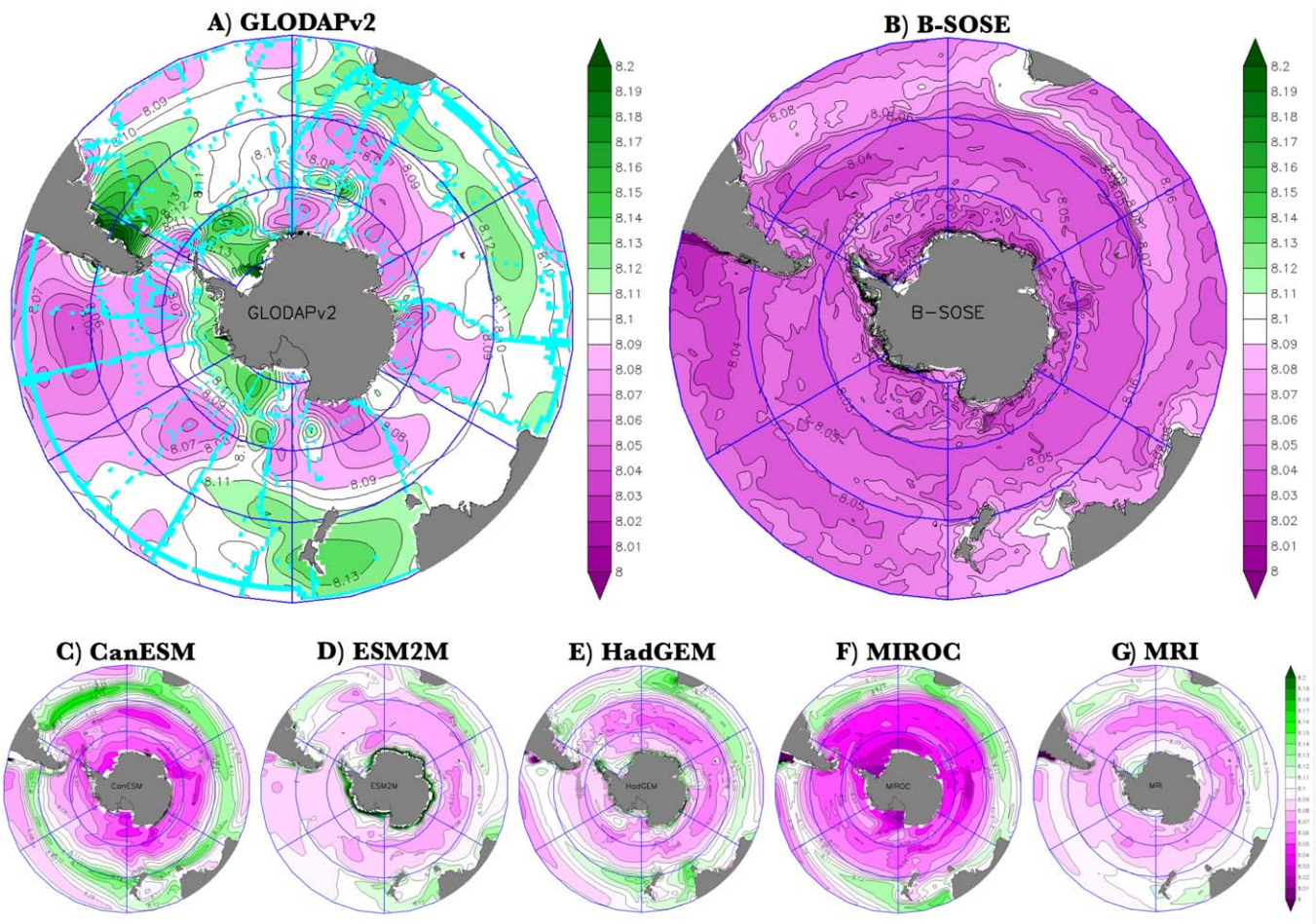

Figure 8. Surface pH (at the in situ temperature and total scale, 0–50 m average) from the GLODAPv2-gridded climatology, the state estimate (averaged from 2008 to 2012), and the model simulations. The cyan tracks in Figure 8a indicate where at least one pH observation in that 1° box was recorded. All model figures cover the simulated years 1986–2005 from the HISTORICAL forcing scenario. (a) GLODAPv2; (b) B-SOSE; (c) CanESM2; (d) GFDL-ESM2M; (e) HadGEM2-ES; (f) MIROC-ESM; and (g) MRI-ESM1.

regions of the Southern Ocean in as little as 13 years (~2030, McNeil & Matear, 2008). Monitoring and accurately simulating the Southern Ocean surface pH and its trend is critical. Small differences can potentially have large effects on simulated acidification trends as calcification rates are especially sensitive to small changes in the pH.

The GLODAPv2 data set (Key et al., 2015; Lauvset et al., 2016; Olsen et al., 2016) provides annually averaged, gridded climatological pH data calculated from the measured alkalinity and total dissolved inorganic carbon (DIC), at the in situ temperature and pressure in the Southern Ocean. GLODAP is heavily biased to summer time measurements when the pH is higher: pH at the surface in the Southern Ocean in winter is lower due to ice cover, lack of biological activity, and deep mixing. The BGC Argo floats will shortly surpass these limited observations and provide detailed depth and seasonal information as well as giving us the opportunity to spot trends as they happen. Early results indicate that the mean float pH is significantly lower than the GLODAPv2 mean, due to ocean acidification and the much older age of the GLODAP data. The difference is consistent with the expected rate of acidification (Johnson et al., 2017). Finally, the new Biogeochemical Southern Ocean State Estimate (B-SOSE, Verdy & Mazloff, 2017) shows clear signs of recent acidification at the surface.

Figure 8 shows the annually averaged pH over the top 50 m. Ekman-driven surface divergence brings old, carbon-rich, low pH water to the surface: this feature is generally captured in the model simulations with some differences between the specific pH values present. All of the simulations have lower pH water in the upwelling region than is seen in GLODAPv2 as expected. Several of the simulations have excessively alkaline waters north of the ACC (e.g., CanESM) leading to anomalously high pH and several have too acidic, low pH, water in the upwelling region (e.g., MIROC). Since the simulated pH is sensitive to the simulated DIC content and the simulated alkalinity, and the uptake of anthropogenic carbon depends, in turn, on the simulated pH (and carbon etc.). Accurate projections of future uptake by the Southern Ocean will depend on simulating a very accurate pH (to better than ~0.01) in the surface water.

Wind observations are from the CFSR reanalysis (Saha et al., 2010) and the observed ACC transport is from Donohue et al. (2016). The total annual uptakes of heat and carbon can be considered as derived rather than observed since they are based on changes in storage, rather than observations of the flux. These are also recent estimates since they are based on data collected between 1993 and 2010. The "observed" carbon uptake (normalized for 2005) is taken from Takahashi et al. (2009), and the "observed" heat uptake calculation is described in the text. The last column shows the fraction of the global uptake of carbon that is taken up south of 30°S. The CSIRO model is a coupled climate model, not an earth system model and does not explicitly simulate carbon. B-SOSE is the Biogeochemical Southern Ocean State Estimate (Verdy & Mazloff, 2017) and does not simulate the ocean north of 30°S.

4. Discussion

The geographical distribution of the surface heat flux and its net integrated value ("heat uptake") are both critical for the heat balance of the Southern Ocean and the global response to increased greenhouse forcing. But as noted in section 3.1, one of the biggest observational shortcomings is that the *net buoyancy flux—the heat flux and the fresh water flux—at the ocean's surface cannot be measured directly*. While the radiative components of the total heat uptake—net incident shortwave radiation and net outgoing longwave radiation—are routinely measured (adjusting for clouds), there are considerable uncertainties and different products disagree substantially (see Bourassa et al., 2013 for a detailed discussion). Latent and sensible heat exchanges cannot be measured from orbit and are generally *calculated* from bulk formulae or in reanalyses. The various heat flux components can be found at: OAFlux (WHOI); J-OFURO2 (Japan); ICOADS; HOAPS (Hamburg); GSSTF2c & GSSTF3 (Goddard); and FLUXNET:MTE (MPI); all are accessible through https://climatedataguide.ucar.edu/variables/atmosphere/latent-sensible-heat-flux. Also, the NCEP, NOAA-CIRES, NCEP2 data sets are available through the Physical Sciences Division at NOAA's Earth System Research Laboratory (www.esrl.noaa.gov/psd).

We can, however, estimate the longer-term net heat flux into the Southern Ocean by assessing changes in *heat storage* as follows: (1) Rhein et al. (2013) reported an annual ocean uptake of 0.257 PW between 1993 and 2010; (2) Roemmich et al. (2015) calculate from the Argo array that the net ocean heat storage above 2,000 m increased at a rate between 5.6 and 7.9 $\times$ 10^{21} (J/yr), equivalent to a global uptake of between

0.18 and 0.25 PW annually between 2006 and 2013—which would be a minimum estimate since Argo does not observe heat changes below 2,000 m which we know are warming (Purkey & Johnson, 2010, 2013) and is, therefore, consistent with Rhein et al. (2013); and (3) Frölicher et al. (2015) determined that the Southern Ocean is responsible for 75% (±22%) of the global heat uptake over the historical period across the CMIP5 model mean. If we assume that heat uptake in the Rhein et al. (2013) study and the heat uptake in the

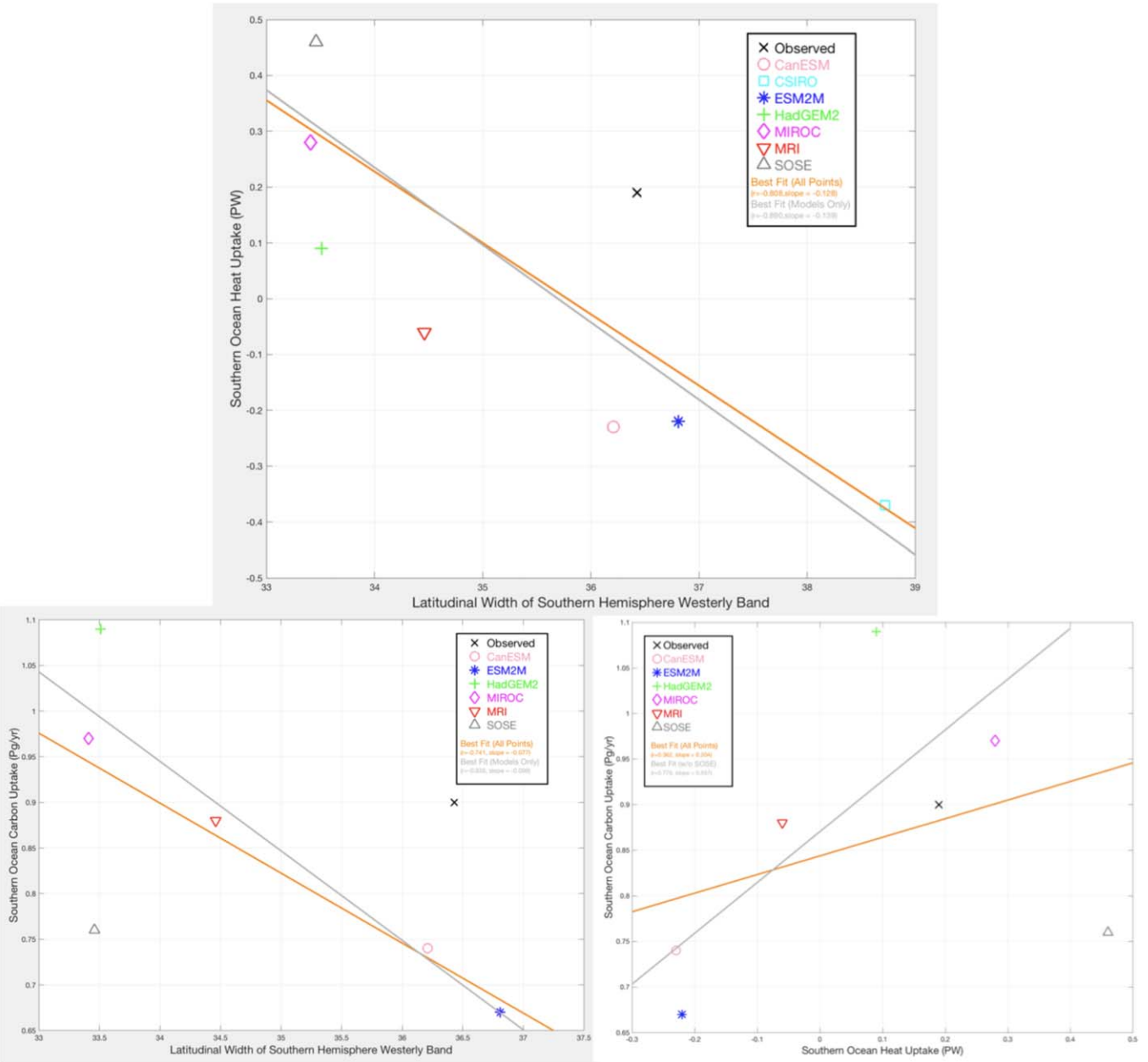

Figure 9. (a) Scatter plot of the width of the Southern Hemisphere westerly wind band (in degrees of latitude) against the annual-mean integrated heat uptake south of 30°S (in PW—negative uptake is heat lost from the ocean), along with the "best fit" linear relationship (slope = −0.128 PW/° latitude) for the models and observations shown in Figure 1. The calculation of the "observed" heat flux into the Southern Ocean is described in the text. The correlation is significant above the 98% level based on a simple *t* test. (b) Scatter plot of the width of the Southern Hemisphere westerly wind band (in degrees of latitude) against the annual-mean integrated carbon uptake south of 30°S (in Pg C/yr), along with the "best fit" linear relationship for the all the points (orange line) and excluding the B-SOSE data (gray line). Calculation of the "observed" carbon flux into the Southern Ocean is described in the text. The correlation of the linear fit to the data (without B-SOSE) is significant above the 99% level based on a simple *t* test. (c) Scatter plot of the net heat uptake south of 30°S (in PW) against the annual-mean integrated carbon uptake south of 30°S (in Pg C/yr), along with the "best fit" linear relationship for all the points (orange line) and excluding the B-SOSE data (gray line). B-SOSE's carbon uptake may be different, in part, due to the later time period over which the state estimate is determined (2008–2012).

Roemmich et al. (2015) study are consistent for their period of overlap (2006–2010), and that the Frölicher et al. (2015) relationship is relatively constant, then taken together, these imply that the Southern Ocean took up at least 0.19 $\pm$0.07 PW annually, although Meijers et al. (2011) found the net heat input over the Southern Ocean between 1992 and 2010 was significantly smaller (0.026 $\pm$ 0.005 PW annually).

Heat flux components are standard variables in the CMIP5 model simulations, and the models are quantitatively consistent with the observations, at least in the zonal mean (not shown). Table 2 and Figure 9a show the integrated heat flux south of 30°S from the model simulations, B-SOSE and as calculated above. Several models are shedding heat through the Southern Ocean, rather than taking in heat as is seen in the observations. This is not inconsistent with Frölicher et al. (2015), however, who noted that the increased *uptake* of heat by the Southern Ocean since the Industrial Revolution is often expressed as a dramatic *reduction in the heat loss*, rather than actual uptake (and this is similar to the carbon uptake assessment). Assessing the mechanisms and causes of these biases will be essential in our effort to reduce the uncertainty associated with future climate projections.

Our analysis revealed a significant correlation between the Southern Hemisphere surface westerly wind profile and the air/sea exchange of heat and carbon over the Southern Ocean in the CMIP5 simulations. The southern limit of the annual mean westerlies (defined as the latitude where the annual mean zonal wind stress shifts from westerly to easterly) is correlated with the net heat uptake over the Southern Ocean (south of 30°S) in the models at -0.796, and this correlation is significant at the 95% level (based on a simple *t* test) even though we only consider seven members. The correlation is stronger and more significant (–0.890 at 98.5%,) when we consider the entire latitudinal width of the westerly band (see Figures 2, 9a, and Table 2). Correlations of the heat uptake and carbon uptake with other wind-related metrics (e.g., the maximum zonal-mean wind stress, the latitude of the maximum wind stress, and the integral of the total wind stress between 70°S and 40°S) were all lower and not significant.

That the heat uptake by the Southern Ocean inversely scales with the location of northward surface flow and more broadly with the width of the westerlies is not surprising: a narrower band of westerlies has steeper gradients on both sides of the maximum wind stress and therefore more Ekman divergence south of the maximum and more convergence north of the maximum. As the intense divergence moves northward, where the underlying surface water becomes less dense, and is more directly over the open latitudes of Drake Passage where divergence cannot be balanced by a mean geostrophic return flow, the winds become more effective at driving a larger upwelling from deeper in the water column. This allows additional cold water to come into contact with the atmosphere and take up some heat before being subducted again as either bottom water or intermediate water.

Swart and Fyfe (2012) report a smaller equatorward bias of the mean wind position in CMIP5 relative to CMIP3 (a $\sim$1° improvement) and a significant reduction in the spread of models around the mean with respect to the mean zonal wind speed. Despite the improvement, however, models still consistently place the Southern Hemisphere surface westerly winds equatorward of their observed positions as shown in Figure 2.

This equatorward bias of the maximum zonal wind stress reduces the rate at which anthropogenic heat and carbon can be taken up in the deep ocean and would therefore be expected to result in an underestimation of the total uptake of heat and carbon by the Southern Ocean. The equatorward shift can also explain errors in simulated Antarctic Intermediate Water (AAIW), particularly its tendency to be too salty in many climate-model simulations (Russell et al., 2006a; Sloyan & Kamenkovich, 2007): the simulation of AAIW matters, especially in ESMs, since it underlies most of the global thermocline (Talley, 2013) and is the source of most of the nutrients upwelled outside of the Southern Ocean (Marinov et al., 2006; Sarmiento et al., 2004).

As is seen with heat uptake, carbon uptake is also highly correlated in the models with both the width of the Southern Hemisphere westerly band (gray line in Figure 9b, $r = -0.936$, significant at 99.4%) and also with the net heat uptake south of 30°S but with slightly lower correlation and significance (gray line in Figure 9c, $r = 0.779$, significant at $>$93%). The representation of the latitudinal structure (and strength) of the Southern Hemisphere westerlies is thus critical to the overall uptake of both heat and carbon across simulations. Although heat uptake in B-SOSE exhibits the expected relationship with the winds, the carbon uptake is significantly lower than expected. This difference in carbon uptake may be due, in part, to the later time period over which the state estimate is determined (2008–2012). Future work will expand this study to include other models and other forcing scenarios to assess the reliability of these relationships.

5. Conclusions

Observationally based metrics are a powerful and objective tool for evaluating progress in reducing uncertainties in future climate projections. The Southern Ocean Working Group, jointly sponsored by US CLIVAR and Ocean Carbon and Biogeochemistry, brought together Southern Ocean experts from diverse backgrounds to discuss which observations and variables best capture the overall quality of a Southern Ocean simulation. We have presented several of these metrics here for both simulations and observations: winds, ACC transport, frontal positions, sea-ice, heat uptake, and carbon-related metrics. These metrics are useful for reducing uncertainties in climate projections, especially those related to oceanic heat and carbon uptake. It is important to keep in mind, however, that while both are anthropogenically forced, carbon and heat do not always respond the same to changes in climate. Differences between the net CO_2 and heat uptake are caused by differences in uptake kinetics, air-sea equilibration time scales, atmospheric boundary conditions, and biology, among other things. A host of other metrics are presented as supporting information figures including upper ocean heat content, P-E fields, DIC, nitrate, zonal current velocity, and clouds. These additional metrics are important for various considerations, but were deemed supplemental to this particular analysis of heat and carbon uptake.

Most notably, we find that the uptake of both carbon and heat is highly correlated to the simulated width of the Southern Hemisphere westerlies band. The width of the westerlies band, in turn, depends largely on the latitude at which the westerlies transition to polar easterlies near Antarctica (where the annual mean zonal wind stress is zero). A narrower band and a more northward polar zero zonal wind stress line, especially when it is more directly over the Drake Passage (62°S–56°S), increases the magnitude of the simulated surface divergence and places it over less dense isopycnals, making it more effective at mixing deeper into the water column (along those isopycnal surfaces). This exposes additional cold water to the overlying warmer atmosphere and increases the overall heat uptake. The connection between the pattern of the winds and overall carbon uptake is similar, although slightly weaker.

Vertical motion (forced by the wind) has a direct effect on the simulation of several of the other metrics discussed in this paper, namely surface pH and sea ice extent, as well as playing an indirect role on the vertical structure of the exchanges across 30°S. While these variables are strong indicators of the overall quality of a simulation, each depends on both the internally generated vertical structure (tied to the physical circulation and/or the biogeochemical cycling) and on the patterns of the induced upwelling. Attention to detail is needed to avoid accepting the "right answer for the wrong reason." Important metrics relating the winds to the ACC have been thoroughly discussed elsewhere (Meijers et al., 2012; Russell et al., 2006a, 2006b; Sen Gupta et al., 2009), but it should be noted that the ACC transport through Drake Passage and the net heat uptake in the Southern Ocean are also well correlated ($r = 0.78$, significant at the 95% level), although both depend on the winds rather than directly affecting each other. While the ACC strength is better correlated with the maximum wind strength and position (not shown), the heat and carbon uptake were best correlated with the overall width.

Here we focused on the role of large-scale observations in assessing the fidelity of climate simulations in the Southern Ocean. We also must emphasize that observationally based metrics play another key role by quantifying the smaller-scale physics that must be parameterized in climate models. For example, the recent Diapycnal and Isopycnal Mixing Experiment in the Southern Ocean (DIMES) has provided the first direct estimates of lateral and vertical diffusivities: parameters that play a key role in setting the ACC zonal transport and the meridional overturning circulation (Tulloch et al., 2014; Watson et al., 2013). It is important that the modeling community adopt the results of these process studies to reduce uncertainty in parameterized physics.

At this point, our way forward requires two essential tracks: first, we, collectively, must carry out rigorous assessments of all model simulations against these and potentially *all* observationally based metrics in order to evaluate the biases in the models, reduce our intermodel differences, and reduce the uncertainty in our projections of the future. Second, we need to encourage and bring about the continued expansion of the available accurate observations: we are excited by the increasing availability of biogeochemical data from the nascent BGC-Argo efforts as well as the prospect of new data generated as part of the SOOS efforts.

In order to ensure their inclusion in the various model intercomparison projects that are part of the upcoming CMIP6, we encourage all modeling centers to make their simulations available in standard, orthogonal grids

(latitude versus longitude versus depth) and to calculate and report quantities with significant covariance (e.g., lateral heat fluxes) for better budget calculations. Toward this goal, the Earth System Model Evaluation Tool (ESMValTool; http://www.esmvaltool.org/, Eyring et al., 2016) is an invaluable resource for the climate modeling and assessment community that allows for routine comparison of single or multiple models against observations. Several of us are working on developing packages for the metrics discussed in this study to be included in the ESMValTool, and we strongly encourage other modeling groups to do the same.

Acknowledgments

The authors would like to thank US CLIVAR and OCB for their support of the Southern Ocean Working Group (SOWG). We had specifically like to thank Mike Patterson, Heather Benway, and Kristan Uhlenbrock for their contributions, guidance, and support, as well as participants of the 2014 Workshop Ocean's Carbon and Heat Uptake: Uncertainties and Metrics (Russell et al., 2015). We want to thank the members of the SOCCOM team for their participation and support. We also would like to thank J. Robbie Toggweiler and an anonymous reviewer for their critical and supportive evaluation of this study—it is significantly better due to their input. All simulation data are available through the Earth System Grid Federation (ESGF, formerly PCMDI, https://esgf-node.llnl.gov/projects/esgf-llnl/), the state estimate data are at Scripps Institution of Oceanography (http://sose.ucsd.edu/), and the observations on which the metrics are based can be found at: World Ocean Atlas 2013 (https://www.nodc.noaa.gov/OC5/woa13/), CFSR (https://rda.ucar.edu/pub/cfsr.html), NCEP/NCEP2 (https://www.esrl.noaa.gov/psd/data/gridded/reanalysis/), and GLODAPv2 (http://cdiac.ornl.gov/oceans/GLODAPv2/). The authors wish to acknowledge use of the Ferret program for analysis and graphics in this paper: Ferret is a product of NOAA's Pacific Marine Environmental Laboratory (http://ferret.pmel.noaa.gov/Ferret/). This work was sponsored, in part, by NSF's Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) Project under the NSF Award PLR-1425989, with additional support from NOAA and NASA. Logistical support for SOCCOM in the Antarctic was provided by the U.S. National Science Foundation through the U.S. Antarctic Program. Funding for this work was also provided under NSF Award PLR-1246247, "The Southern Ocean in a Warming World: Winds, Carbon and Heat." The authors are not aware of any conflicts of interest.

References

Abernathy, R. P., Cerovecki, I., Holland, P. R., Newsom, E., Mazloff, M., & Talley, L. D. (2016). Water-mass transformation by sea ice in the upper branch of the Southern Ocean overturning. *Nature Geoscience*, *9*, 596–601. https://doi.org/10.1038/ngeo2749

Allen, M. R., Stott, P. A., Mitchell, J. F. B., Schnur, R., & Delworth, T. L. (2000). Quantifying the uncertainty in forecasts of anthropogenic climate change. *Nature*, *407*, 617–620. https://doi.org/10.1038/35036559

Anderson, R. F., Ali, S., Bradtmiller, L. I., Nielsen, S. H. H., Fleisher, M. Q., Anderson, B. E., & Burckle, L. H. (2009). Wind-driven upwelling in the southern ocean and the deglacial rise in atmospheric CO_2. *Science*, *323*, 1443–1448. https://doi.org/10.1126/science.1167441

Arakawa, A., & Lamb, V. R. (1977). Computational design of the basic dynamical process of the UCLA general circulation model. *Methods in Computational Physics*, *17*, 173–265. https://doi.org/10.1016/B978-0-12-460817-7.50009-4

Arrigo, K. R., van Dijken, G., & Long, M. (2008). Coastal Southern Ocean: A strong anthropogenic CO_2 sink. *Geophysical Research Letters*, *35*, L21602. https://doi.org/10.1029/2008GL035624

Bernardello, R., Marinov, I., Palter, J. B., Galbraith, E. D., & Sarmiento, J. L. (2014). Impact of Weddell Sea deep convection on natural and anthropogenic carbon in a climate model. *Geophysical Research Letters*, *41*, 7262–7269. https://doi.org/10.1002/2014GL061313

Bitz, C. M., Gent, P. R., Woodgate, R. A., Holland, M. M., & Lindsay, R. (2006). The influence of sea ice on ocean heat uptake in response to increasing CO_2. *Journal of Climate*, *19*, 2437–2450. https://doi.org/10.1175/JCLI3756.1

Böning, C., Dispert, A., Visbeck, M., Rintoul, S. R., & Schwarzkopf, F. U. (2008). The response of the Antarctic Circumpolar Current to recent climate change. *Nature Geoscience*, *1*, 864–869. https://doi.org/10.1038/ngeo362

Bourassa, M. A., Gille, S. T., Bitz, C., Carlson, D., Cerovecki, I., Clayson, C. A., et al. (2013). High-latitude ocean and sea ice surface fluxes: Challenges for climate research. *Bulletin of the American Meteorological Society*, *94*, 403–423. https://doi.org/10.1175/BAMS-D-11-00244.1

Bracegirdle, T. J., Bertler, N., Carleton, A. M., Ding, Q., Fogwill, C. J., Fyfe, J. C., et al. (2016). A multi-disciplinary perspective on climate model evaluation for Antarctica. *Bulletin of the American Meteorological Society*, *97*, ES23–ES26. https://doi.org/10.1175/BAMS-D-15-00108.1

Bracegirdle, T. J., Shuckburgh, E., Sallee, J.-B., Wang, Z., Meijers, A. J. S., Bruneau, N., et al. (2013). Assessment of surface winds over the Atlantic, Indian, and Pacific Ocean sectors of the Southern Ocean in CMIP5 models: Historical bias, forcing response, and state dependence. *Journal of Geophysical Research: Atmospheres*, *118*, 547–562. https://doi.org/10.1002/jgrd.50153

Cheon, W. G., Park, Y., Toggweiler, J. R., & Lee, S. (2014). The relationship of Weddell Polynya and open-ocean deep convection to the southern hemisphere westerlies. *Journal of Physical Oceanography*, *44*, 694–713. https://doi.org/10.1175/JPO-D-13-0112.1

Chereskin, T. K., Donohue, K. A., & Watts, D. R. (2012). cDrake: Dynamics and transport of the Antarctic circumpolar current in drake passage. *Oceanography*, *25*, 134–135. https://doi.org/10.5670/oceanog.2012.86

Chidichimo, M. P., Donohue, K. A., Watts, D. R., & Tracey, K. L. (2014). Baroclinic transport time series of the Antarctic circumpolar current measured in drake passage. *Journal of Physical Oceanography*, *44*, 1829–1853. https://doi.org/10.1175/JPO-D-13-071.1

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., et al. (2013). Carbon and other biogeochemical cycles. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press.

Collins, W. J., Bellouin, N., Doutriaux-Boucher, M., Gedney, N., Hinton, T., Jones, C. D., et al. (2008). *Evaluation of the HadGEM2 model* (Tech. Note HCTN 74). Exeter, UK: Met Office Hadley Centre. Available at http://www.metoffice.gov.uk/publications/HCTN/index.html

Cunningham, S. A., Alderson, S. G., King, B. A., & Brandon, M. A. (2003). Transport and variability of the Antarctic circumpolar current in drake passage. *Journal of Geophysical Research*, *108*(C5), 8084. https://doi.org/10.1029/2001JC001147

Dalan, F., Stone, P. H., Kamenkovich, I. V., & Scott, J. R. (2005). Sensitivity of the Ocean's climate to diapycnal diffusivity in an EMIC. Part I: Equilibrium state. *Journal of Climate*, *18*, 2460–2481. https://doi.org/10.1175/JCLI3411.1

Dalmonech, D., Foley, A. M., Anav, A., Friedlingstein, P., Friend, A. D., Kidston, M., et al. (2014). Challenges and opportunities to reduce uncertainty in projections of future atmospheric CO_2: A combined marine and terrestrial biosphere perspective. *Biogeosciences Discussions*, *11*, 2083–2153. https://doi.org/10.5194/bgd-11-2083-2014

Dong, S., Sprintall, J., & Gille, S. T. (2006). Location of the polar front from AMSR-E satellite sea surface temperature measurements. *Journal of Physical Oceanography*, *36*, 2075–2089. https://doi.org/10.1175/JPO2973.1

Donohue, K. A., Tracey, K. L., Watts, D. R., Chidichimo, M. P., & Chereskin, T. K. (2016). Mean Antarctic circumpolar current transport measured in drake passage. *Geophysical Research Letters*, *43*, 11760–11767. https://doi.org/10.1002/2016GL070319

Dufour, C. O., Le Sommer, J., Gehlen, M., Orr, J. C., Molines, J.-M., Simeon, J., et al. (2013). Eddy compensation and controls of the enhanced sea-to-air CO_2 flux during positive phases of the Southern Annular Mode. *Global Biogeochemical Cycles*, *27*, 950–961. https://doi.org/10.1002/gbc.20090

Dunne, J. P., John, J. G., Adcroft, A. J., Griffies, S. M., Hallberg, R. W., Shevliakova, E., et al. (2012). GFDL's ESM2 global coupled climate–carbon earth system models. Part I: Physical formulation and baseline simulation characteristics. *Journal of Climate*, *25*, 6646–6665. https://doi.org/10.1175/JCLI-D-11–00560.1

Dunne, J. P., John, J. G., Shevliakova, E., Stouffer, R. J., Krasting, J. P., Malyshev, S. L., et al. (2013). GFDL's ESM2 global coupled climate–carbon earth system models. Part II: Carbon system formulation and baseline simulation characteristics. *Journal of Climate*, *26*, 2247–2267. https://doi.org/10.1175/JCLI-D-12–00150.1

Exarchou, E., Kuhlbrodt, T., Gregory, J. M., & Smith, R. S. (2015). Ocean heat uptake processes: A model intercomparison. *Journal of Climate*, *28*, 887–908. https://doi.org/10.1175/JCLI-D-14-00235.1

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., et al. (2016). ESMValTool (v1.0): A community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geoscientific Model Development*, *9*, 1747–1802. https://doi.org/10.5194/gmd-9-1747-2016

Fasullo, J. T., & Trenberth, K. E. (2012). A less cloudy future: The role of subtropical subsidence in climate sensitivity. *Science*, *338*, 792–794. https://doi.org/10.1126/science.1227465

Ferrari, R., & Wunsch, C. (2009). Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annual Review of Fluid Mechanics*, *41*, 253–282. https://doi.org/10.1146/annurev.fluid.40.111406.102139

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., et al. (2013). Evaluation of climate models. In T. F. Stocker et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press.

Flato, G. M., Boer, G. J., Lee, W. G., McFarlane, N. A., Ramsden, D., Reader, M. C., et al. (2000). The Canadian Centre for climate modelling and analysis global coupled model and its climate. *Climate Dynamics*, *16*, 451–467. https://doi.org/10.1007/s003820050339

Freeman, N. M., & Lovenduski, N. S. (2016). Mapping the Antarctic polar front: Weekly realizations from 2002 to 2014. *Earth System Science Data*, *8*, 191–198. https://doi.org/10.5194/essd-8-191-2016

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*, 862–886. https://doi.org/10.1175/JCLI-D-14–00117.1

Galbraith, E. D., Dunne, J. P., Gnanadesikan, A., Slater, R. D., Sarmiento, J. L., Dufour, C. O., et al. (2015). Complex functionality with minimal computation: Promise and pitfalls of reduced-tracer ocean biogeochemistry model. *Journal of Advances Modeling Earth Systems*, *7*, 2012–2028. https://doi.org/10.1002/2015MS000463

Ganachaud, A., & Wunsch, C. (2000). Improved estimates of global ocean circulation, heat transport and mixing from hydrographic data. *Nature*, *408*, 453–456. https://doi.org/10.1038/35044048

Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., et al. (2014). World Ocean Atlas 2013. In S. Levitus and A. Mishonov (Eds.), *Volume 4: Dissolved inorganic nutrients (phosphate, nitrate, silicate)*. (NOAA Atlas NESDIS 76, 25 pp). Boulder, CA: NOAA.

Gent, P. R. (2016). Effects of Southern Hemisphere wind changes on the meridional overturning circulation in ocean models. *Annual Review Marine Science*, *8*, 79–94. https://doi.org/https://doi.org/10.1146/annurev-marine-122414-033929

Gille, S. T. (2014). Meridional displacement of the Antarctic circumpolar current. *Philosophical Transactions of the Royal Society A*, *372*, 20130273. https://doi.org/10.1098/rsta.2013.0273

Gregory, J. M. (2000). Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, *16*, 501–515. https://doi.org/10.1007/s003820000059

Gruber, N., Sarmiento, J. L., & Stocker, T. (1996). An improved method for detecting anthropogenic CO_2 in the oceans. *Global Biogeochemical Cycles*, *10*, 809–837. https://doi.org/10.1029/96GB01608

Hallberg, R. W., & Gnanadesikan, A. (2006). The role of eddies in determining the structure and response of the wind-driven Southern Hemisphere overturning: Results from the modeling eddies in the Southern Ocean project. *Journal of Physical Oceanography*, *36*, 2232–2252. https://doi.org/10.1175/JPO2980.1

Heuzé, C., Heywood, K. J., Stevens, D. P., & Ridley, J. K. (2013). Southern Ocean bottom water characteristics in CMIP5 models. *Geophysical Research Letters*, *40*, 1409–1414. https://doi.org/10.1002/grl.50287

Heuzé, C., Heywood, K. J., Stevens, D. P., & Ridley, J. K. (2015). Changes in global ocean bottom properties and volume transports in CMIP5 models under climate change scenarios. *Journal of Climate*, *28*, 2917–2944. https://doi.org/10.1175/JCLI-D-14-00381.1

Hogg, A. M. C. (2010). An Antarctic Circumpolar Current driven by surface buoyancy forcing. *Geophysical Research Letters*, *37*, L23601. https://doi.org/10.1029/2010GL044777

Holland, P. R., & Kwok, R. (2012). Wind-driven trends in Antarctic sea-ice drift. *Nature Geoscience*, *5*, 872–875. https://doi.org/10.1038/NGEO1627

Huang, B., Stone, P. H., & Hill, C. (2003). Sensitivities of deep-ocean heat uptake and heat content to surface fluxes and subgrid-scale parameters in an ocean general circulation model with idealized geometry. *Journal of Geophysical Research*, *108*(C1), 3015. https://doi.org/10.1029/2001JC001218

Jeffrey, S. J., Rotstayn, L. D., Collier, M. A., Dravitzki, S. M., Hamalainen, C., Moeseneder, C., et al. (2013). Australia's CMIP5 submission using the CSIRO Mk3.6 model. *Australian Meteorological and Oceanographic Journal*, *63*, 1–13. http://www.bom.gov.au/amoj/docs/2013/jeffrey_hres.pdf

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., et al. (2017). Biogeochemical sensor performance in the SOCCOM profiling float array. *Journal of Geophysical Research: Oceans*, *122*, 6414–6436. https://doi.org/10.1002/2017JC012838

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*, 437–470. https://doi.org/10.1175/1520-0477(1996)077 < 0437:TNYRP>2.0.CO;2

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., et al. (2002). NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the American Meteorological Society*, *83*, 1631–1643. https://doi.org/10.1175/BAMS-83-11-1631

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., et al. (2004). A global ocean carbon climatology: Results from Global Data Analysis Project (GLODAP). *Global Biogeochemical Cycles*, *18*, GB4031. https://doi.org/10.1029/2004GB002247

Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2015). *Global Ocean Data Analysis Project, Version 2 (GLODAPv2)* (ORNL/CDIAC-162, NDP-093). Oak Ridge, TN: Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy. https://doi.org/10.3334/CDIAC/OTG.NDP093_GLODAPv2

Khatiwala, S., Tanhua, T., Fletcher, S. M., Gerber, M., Doney, S. C., Graven, H. D., et al. (2013). Global ocean storage of anthropogenic carbon. *Biogeosciences*, *10*, 2169–2191. https://doi.org/10.5194/bg-10-2169-2013.

Kidston, J., Taschetto, A. S., Thompson, D. W. J., & England, M. H. (2011). The influence of Southern Hemisphere sea-ice extent on the latitude of the mid-latitude jet stream. *Geophysical Research Letters*, *38*, L15804. https://doi.org/10.1029/2011GL048056

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., & Meehl, G. A. (2010). Challenges in combining projections from multiple climate models. *Journal of Climate*, *23*, 2739–2758. https://doi.org/10.1175/2009JCLI3361.1

Kostianoy, A., Ginzburg, A. I., Frankignoulle, M., & Delille, B. (2004). Fronts in the Southern Indian Ocean as inferred from satellite sea surface temperature data. *Journal of Marine Systems*, *45*, 55–73. https://doi.org/10.1016/j.jmarsys.2003.09.004

Kuhlbrodt, T., & Gregory, J. M. (2012). Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change. *Geophysical Research Letters*, *39*, L18608. https://doi.org/10.1029/2012GL052952

Landschützer, P., Gruber, N., Haumann, F. A., Rödenbeck, C., Bakker, D. C. E., van Heuven, S., et al. (2015). The reinvigoration of the Southern Ocean carbon sink. *Science*, *349*, 1221–1224. https://doi.org/10.1126/science.aab2620

Lauvset, S. K., Key, R. M., Olsen, A., van Heuven, S., Velo, A., Lin, X., et al. (2016). A new global interior ocean mapped climatology: The 1°×1° GLODAP version 2. *Earth System Science Data*, *8*, 325–340. https://doi.org/10.5194/essd-8-325-2016

Lenton, A., & Matear, R. J. (2007). Role of the Southern Annular Mode (SAM) in Southern Ocean CO_2 uptake. *Global Biogeochemical Cycles*, *21*, GB2016. https://doi.org/10.1029/2006GB002714

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, E. T., Langenfelds, R., Gomez, A., et al. (2007). Saturation of the Southern Ocean CO_2 sink due to recent climate change. *Science*, *316*, 1735–1738. https://doi.org/10.1126/science.1136188

Levitus, S. (1982). *Climatological Atlas of the World Ocean* (NOAA Prof. Pap. 13, pp. 191). Rockville, MD: U.S. Government Printing Office.
Locarnini, R. A., Mishonov, A. V., Antonov, J. I., Boyer, T. P., Garcia, H. E., Baranova, O. K., et al. (2013). World Ocean Atlas 2013. In S. Levitus and A. Mishonov (Eds.), *Volume 1: Temperature* (NOAA Atlas NESDIS 73, 40 pp.). Boulder, CA: NOAA.
Lovenduski, N. S., Gruber, N., & Doney, S. C. (2008). Toward a mechanistic understanding of the decadal trends in the Southern Ocean carbon sink. *Global Biogeochemical Cycles*, *22*, GB3016. https://doi.org/10.1029/2007GB003139
Lumpkin, R., & Speer, K. (2007). Global ocean meridional overturning. *Journal of Physical Oceanography*, *37*, 2550–2562. https://doi.org/10.1175/JPO3130.1
Macdonald, A. M., & Wunsch, C. (1996). An estimate of global ocean circulation and heat fluxes. *Nature*, *382*, 436–439. https://doi.org/10.1038/382436a0.s
Manabe, S., Bryan, K., & Spelman, M. J. (1990). Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide. *Journal of Physical Oceanography*, *20*, 722–749. https://doi.org/10.1175/1520-0485(1990)020<0722:TROAGO>2.0.CO;2
Marinov, I., Gnanadesikan, A., Sarmiento, J. L., Toggweiler, J. R., Follows, M., & Mignone, B. K. (2008). Impact of oceanic circulation on biological carbon storage in the ocean and atmospheric pCO_2. *Global Biogeochemical Cycles*, *22*, GB3007. https://doi.org/10.1029/2007GB002958
Marinov, I., Gnanadesikan, A., Toggweiler, J. R., & Sarmiento, J. R. (2006). The Southern Ocean biogeochemical divide. *Nature*, *441*, 964–967. https://doi.org/10.1038/nature04883
Marshall, J., & Radko, T. (2003). Residual-mean solutions for the Antarctic circumpolar current and its associated overturning circulation. *Journal of Physical Oceanography*, *33*, 2341–2354. https://doi.org/10.1175/1520-0485(2003)033 < 2341:RSFTAC>2.0.CO;2
Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, *5*, 171–180. https://doi.org/10.1038/ngeo1391
Mazloff, M., Heimbach, P., & Wunsch, C. (2010). An eddy-permitting Southern Ocean state estimate. *Journal of Physical Oceanography*, *40*, 880–899. https://doi.org/10.1175/2009JPO4236.1
McNeil, B. I., & Matear, R. J. (2008). Southern Ocean acidification: A tipping point at 450-ppm atmospheric CO2. *Proceedings of the National Academy of Sciences United States of America*, *105*, 18860–18864. https://doi.org/10.1073/pnas.0806318105
Meijers, A. J. S., Bindoff, N. L., & Rintoul, S. R. (2011). Frontal movements and property fluxes: Contributions to heat and freshwater trends in the Southern Ocean. *Journal of Geophysical Research*, *116*, C08024. https://doi.org/10.1029/2010JC006832
Meijers, A. J. S., Shuckburgh, E., Bruneau, N., Sallee, J.-B., Bracegirdle, T. J., & Wang, Z. (2012). Representation of the Antarctic Circumpolar Current in the CMIP5 climate models and future changes under warming scenarios. *Journal of Geophysical Research*, *117*, C12008. https://doi.org/10.1029/2012JC008412
Meredith, M., Watkins, J. L., Murphy, E. J., & Ward, P. (2003). The southern ACC front to the northeast of South Georgia: Pathways, characteristics and fluxes. *Journal of Geophysical Research*, *108(C5)*, 3162. https://doi.org/10.1029/2001JC001227
Meredith, M. P., Woodworth, P. L., Chereskin, T. K., Marshall, D. P., Allison, L. C., Bigg, G. R., et al. (2011). Sustained monitoring of the Southern Ocean at Drake Passage: Past achievements and future priorities. *Reviews of Geophysics*, *49*, RG4005. https://doi.org/10.1029/2010RG000348
Moore, J. K., Abbott, M. R., & Richman, J. G. (1999). Location and dynamics of the Antarctic polar front from satellite sea surface temperature data. *Journal of Geophysical Research*, *104*, 3059–3073. https://doi.org/https://doi.org/10.1029/1998JC900032
Munro, D. R., Lovenduski, N. S., Takahashi, T., Stephens, B., Newberger, T., & Sweeney, C. (2015). Recent evidence for a strengthening CO_2 sink in the Southern Ocean from carbonate system measurements in the Drake Passage (2002–2015). *Geophysical Research Letters*, *42*, 7623–7630. https://doi.org/10.1002/2015GL065194
Olsen, A., Key, R. M., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2016). The Global Ocean Data Analysis Project version 2 (GLODAPv2): An internally consistent data product for the world ocean. *Earth Systems Science Data*, *8*, 297–323. https://doi.org/10.5194/essd-8-297-2016
Orsi, A. H., Whitworth, T., & Nowlin, W. D. (1995). On the meridional extent and fronts of the Antarctic Circumpolar Current. *Deep Sea Research Part I: Oceanographic Research Papers*, *42*, 641–673. https://doi.org/10.1016/0967-0637(95)00021-W
Palmer, M., Haines, K., Antonov, J., Barker, P., Bindoff, N., Boyer, T., et al. (2010). Future observations for monitoring global ocean heat content. In J. Hall, D. E. Harrison, & D. Stammer (Eds.), *Proceedings of the "OceanObs'09: Sustained Ocean Observations and Information for Society" Conference* (Vol. 2), Venice, Italy, 21–25 September 2009, *ESA Publication* WPP-306. https;//doi.org/10.5270/OceanObs09
Purkey, S. G., & Johnson, G. C. (2010). Warming of global abyssal and deep Southern Ocean waters between the 1990s and 2000s: Contributions to global heat and sea level rise budgets. *Journal of Climate*, *23*, 6336–6351. https://doi.org/10.1175/2010JCLI3682.1
Purkey, S. G., & Johnson, G. C. (2012). Global contraction of Antarctic Bottom Water between the 1980s and 2000s. *Journal of Climate*, *25*, 5830–5844. https://doi.org/10.1175/JCLI-D-11-00612.1
Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., et al. (2007). Climate models and their evaluation. In S. Solomon et al. (Eds.), *Climate change 2007: The physical science basis. Contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press.
Rhein, M., Rintoul, S. R., Aoki, S., Campos, E., Chambers, D., Feely, R. A., et al. (2013). Observations: Ocean. In T. F. Stocker et al (Eds.), *Climate change 2013: The physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press.
Rintoul, S. R., Sparrow, M., Meredith, M. P., Wadley, V., Speer, K., Hofmann, E., et al. (2012). *The Southern Ocean observing system: Initial science and implementation strategy* (74 pp.). Cambridge, UK: SCAR. Available at https://www.scar.org/library/scar-publications/occasional-publications/3502-soos-strategy/
Roemmich, D., Church, J., Gilson, J., Monselesan, D., Sutton, P., & Wijffels, S. (2015). Unabated planetary warming and its ocean structure since 2006. *Nature Climate Change*, *5*, 240–245. https://doi.org/10.1038/nclimate2513
Rotstayn, L. D., Jeffrey, S. J., Collier, M. A., Dravitzki, S. M., Hirst, A. C., Syktus, J. I., et al. (2012). Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmospheric Chemistry and Physics*, *12*, 6377–6404. https://doi.org/10.5194/acp-12-6377-2012 (http://www.atmos-chem-phys.net/12/6377/2012/acp-12-6377-2012.html).
Russell, J. L., Benway, H., Bracco, A., Deutsch, C., Ito, T., Kamenkovich, I., et al. (2015). *Ocean's carbon and heat uptake: Uncertainties and metrics* (U.S. CLIVAR Rep. 2015-3, 33 pp.). Available at https://usclivar.org/sites/default/files/documents/2015/SO-OCU-Workshop-Report-final_0.pdf
Russell, J. L., Dixon, K. W., Gnanadesikan, A., Stouffer, R. J., & Toggweiler, J. R. (2006b). The Southern Hemisphere westerlies in a warming world: Propping open the door to the deep ocean. *Journal of Climate*, *19*, 6382–6390. https://doi.org/10.1175/JCLI3984.1
Russell, J. L., Stouffer, R. J., & Dixon, K. W. (2006a). Intercomparison of the Southern Ocean circulations in the IPCC coupled model control simulations. *Journal of Climate*, *19*, 4560–4575. https://doi.org/10.1175/JCLI3869.1
Sabine, C., Feely, R., Gruber, N., Key, R., Lee, K., Bullister, J., et al. (2004). The oceanic sink for anthropogenic CO_2. *Science*, *305*, 367–371. https://doi.org/10.1126/science.1097403

Saha, S., Moorthi, S., Pan, H., Wu, X., Wang, J., Nadiga, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*, 1015–1057. https://doi.org/10.1175/2010BAMS3001.1

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., Wang, Z., et al. (2013a). Assessment of Southern Ocean water mass circulation and characteristics in CMIP5 models: Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, *118*, 1830–1844. https://doi.org/10.1002/jgrc.20135

Sallée, J.-B., Shuckburgh, E., Bruneau, N., Meijers, A. J. S., Bracegirdle, T. J., & Wang, Z. (2013b). Assessment of Southern Ocean mixed layer depths in CMIP5 models: Historical bias and forcing response. *Journal of Geophysical Research: Oceans*, *118*, 1845–1862. https://doi.org/10.1002/jgrc.20157

Sarmiento, J. L., Gruber, N., Brzezinski, M. A., & Dunne, J. P. (2004). High-latitude controls of thermocline nutrients and low latitude biological productivity. *Nature*, *427*, 56–60. https://doi.org/10.1038/nature02127

Sen Gupta, A., Santoso, A., Taschetto, A. S., Ummenhofer, C. C., Trevena, J., & England, M. H. (2009). Projected changes to the Southern Hemisphere Ocean and sea ice in the IPCC AR4 climate models. *Journal of Climate*, *22*, 3047–3078. https://doi.org/10.1175/2008JCLI2827.1

Shao, A. E., Gille, S. T., Mecking, S., & Thompson, L. (2015). Properties of the Subantarctic Front and Polar Front from the skewness of sea level anomaly. *Journal of Geophysical Research: Oceans*, *120*, 5179–5193. https://doi.org/10.1002/2015JC010723

Shiogama, H., Stone, D., Emori, S., Takahashi, K., Mori, S., Maeda, A., et al. (2016). Predicting future uncertainty constraints on global warming projections. *Scientific Reports*, *6*, 18903. https://doi.org/10.1038/srep18903

Sloyan, B. M., & Kamenkovich, I. V. (2007). Simulation of subantarctic mode and Antarctic intermediate waters in climate models. *Journal of Climate*, *20*, 5061–5080. https://doi.org/10.1175/JCLI4295.1

Sokolov, S., & Rintoul, S. R. (2007). Multiple jets of the Antarctic Circumpolar Current South of Australia. *Journal of Physical Oceanography*, *37*, 1394–1412. https://doi.org/10.1175/JPO3111.1

Sokolov, S., & Rintoul, S. R. (2009). Circumpolar structure and distribution of the Antarctic Circumpolar Current fronts: 1. Mean circumpolar paths. *Journal of Geophysical Research*, *114*, C11018. https://doi.org/10.1029/2008JC005108

Stouffer, R. J., Russell, J. L., & Spelman, M. J. (2006). Importance of oceanic heat uptake in transient climate change. *Geophysical Research Letters*, *33*, L17704. https://doi.org/10.1029/2006GL027242

Swart, N. C., & Fyfe, J. C. (2012). Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress. *Geophysical Research Letters*, *39*, L16711. https://doi.org/10.1029/2012GL052810

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

Talley, L. D. (2008). Freshwater transport estimates and the global overturning circulation: Shallow, deep and throughflow components. *Progress in Oceanography*, *78*, 257–303. https://doi.org/10.1016/j.pocean.2008.05.001

Talley, L. D. (2013). Closure of the global overturning circulation through the Indian, Pacific, and southern oceans. *Oceanography*, *26*, 80–97. https://doi.org/10.5670/oceanog.2013.07

Tamsitt, V., Talley, L. D., Mazloff, M. R., & Cerovečki, I. (2016). Zonal variations in the Southern Ocean heat budget. *Journal of Climate*, *29*, 6563–6579. https://doi.org/10.1175/JCLI-D-15-0630.1

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, *93*, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

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

Toggweiler, J. R., & Samuels, B. (1998). On the Ocean's large-scale circulation near the limit of no vertical mixing. *Journal of Physical Oceanography*, *28*, 1832–1852. https://doi.org/10.1175/1520-0485(1998)028 < 1832:OTOSLS>2.0.CO;2

Tulloch, R., Ferrari, R., Jahn, O., Klocker, A., LaCasce, J., Ledwell, J. R., et al. (2014). Direct estimate of lateral eddy diffusivity upstream of drake passage. *Journal of Physical Oceanography*, *44*, 2593–2616. https://doi.org/10.1175/JPO-D-13-0120.1

Turner, J., Bracegirdle, T. J., Phillips, T., Marshall, G. J., & Hosking, J. S. (2013). An initial assessment of Antarctic sea ice extent in the CMIP5 models. *Journal of Climate*, *26*, 1473–1484. https://doi.org/10.1175/JCLI-D-12-00068.1

Uppala, S. M., KÅllberg, P. W., Simmons, A. J., Andrae, U., Da Costa Bechtold, V., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal Royal Meteorological Society*, *131*, 2961–3012. https://doi.org/10.1256/qj.04.176

Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research: Oceans*, *122*, 6968–6988. https://doi.org/10.1002/2016JC012650

Watanabe, M., Chikira, M., Imada, Y., & Kimoto, M. (2011). Convective control of ENSO simulated in MIROC. *Journal of Climate*, *24*, 543–562. https://doi.org/10.1175/2010JCLI3878.1

Watson, A. J., Ledwell, J. R., Messias, M.-J., King, B. A., Mackay, N., Meredith, M. P., et al. (2013). Rapid cross-density ocean mixing at mid-depths in the Drake Passage measured by tracer release. *Nature*, *501*, 408–411. https://doi.org/10.1038/nature12432

Whitworth, T., & Peterson, R. G. (1985). Volume transport of the Antarctic Circumpolar Current from bottom pressure measurements. *Journal of Physical Oceanography*, *15*, 810–816. https://doi.org/10.1175/1520-0485(1985)015<0810:VTOTAC>2.0.CO;2

Yukimoto, S., Yoshimura, H., Hosaka, M., Sakami, T., Tsujino, H., Hirabara, M., et al. (2011). *Meteorological Research Institute Earth System Model Version 1 (MRI-ESM1): Model description* (Tech. Rep. of MRI, 64, 83 pp.). Available at http://www.mri-jma.go.jp/Publish/Technical/index_en.html

Zunz, V., Goosse, H., & Massonet, F. (2013). How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent. *Cryosphere*, *7*, 451–468. https://doi.org/10.5194/tc-7-451-2013

Zweng, M. M., Reagan, J. R., Antonov, J. I., Locarnini, R. A., Mishonov, A. V., Boyer, T. P., et al. (2013). World Ocean Atlas 2013. In S. Levitus and A. Mishonov (Eds.), *Volume 2: Salinity*. (NOAA Atlas NESDIS 74, 39 pp.). Boulder, CA: NOAA.