

*JGR-Oceans SOCCOM Special Issue*

Supporting Information for

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

**Joellen L. Russell1, Igor Kamenkovich2, Cecilia Bitz3, Raffaele Ferrari4, Sarah T. Gille5, Paul J. Goodman1, Robert Hallberg6, Kenneth Johnson7, Karina Khazmutdinova10, Irina Marinov8, Matthew Mazloff5, Stephen Riser3, Jorge L. Sarmiento9, Kevin Speer10, Lynne D. Talley5 and Rik Wanninkhof11**

1University of Arizona, Tucson, Arizona, 2Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, 3University of Washington, Seattle, Washington, 4Massachusetts Institute of Technology, Cambridge, Massachusetts, 5Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, 6National Oceanic and Atmospheric Administration, Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, 7Monterey Bay Aquarium Research Institute, Moss Landing, California, 8University of Pennsylvania, Philadelphia, Pennsylvania, 9Princeton University, Princeton, New Jersey, 10Florida State University, Tallahassee, Florida, 11National Oceanic and Atmospheric Administration, Atlantic Oceanographic and Meteorological Laboratory, Miami, Florida.

Corresponding author: Joellen Russell (jrussell@email.arizona.edu)

**Contents of this file**

Figures S1 to S15

**Additional Supporting Information (Files uploaded separately)**

Captions for Datasets S1 to Sx

Captions for Tables S1 to Sx (if larger than 1 page, upload as separate file)

Captions for Movies S1 to Sx

Captions for Audio S1 to Sx

**Introduction**

As part of our work, we created/designed many other metrics that have proven useful to us and other researchers, but were not essential for inclusion in the current study. We present several of these here as a record of our thoughts and deliberations and experience. The following figures and captions are “free standing” and are presented with minimal discussion.

**Figure S1:** Annual-mean cloud fraction (percent) from the reanalysis and from 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. Panel a) CFSR dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. Panel H) shows the zonal mean for each of the seven maps.

Figure S1 shows the annual-mean cloud coverage (as a percentage) among all cloud types (high, medium and low cloud amount), averaged over all months in the period of record. Cloud cover affects heat and fresh water fluxes over the Southern Ocean. Except for CSIRO, the models generally have less cloud cover than is seen in the reanalysis.

**Figure S2:** Annual cycle (latitude vs time) of the zonal-mean cloud fraction (percent) from the reanalysis and from several of the coupled climate models submitted as part of the CMIP5/IPCC-AR5 process. Note that 2 repeated calendar years are shown. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) CFSR dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1.

Figure S2 shows the zonally-averaged, annual climatology of the cloud coverage (as a percentage) over the Southern Ocean. The same year is repeated twice in each panel. Climatologies are created by averaging all similar months in the period of record (e.g. all Januarys were averaged to create the January climatology, etc.), and the months were compiled to create the annual climatology. From this dataset, we then zonally-average along each latitude to create the zonal average climatologies above. The reanalysis indicates the peak cloud coverage occurs in winter (JJA) at all latitudes, but several of the models have their peak cloud cover in summer (DJF) over the ACC.

**Figure S3:** Annual cycle (latitude vs time) of the total heat flux (zonally-averaged) over the Southern Ocean from the reanalysis and from several of the coupled climate models submitted as part of the CMIP5/IPCC-AR5 process. Positive contours (red shading) indicate a net input of heat into the ocean and negative contours (blue shading) indicate a net loss of heat by the ocean. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) NCEP2 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. The total heat flux is obtained by summing the net shortwave radiative flux at the surface, the net longwave radiative flux at the surface, the latent heat and the sensible heat terms. Panel H shows the area-weighted average heat flux over the entire Southern Ocean (80°S-30°S).

Figure S3 shows the zonally-averaged, annual climatology of the atmosphere-to-ocean heat flux (in W/m2, positive is down, warming the ocean) over the Southern Ocean. The same year is repeated twice in each panel. Climatologies are created by averaging all similar months in the period of record (e.g. all Januarys were averaged to create the January climatology, etc.), and the months were compiled to create the annual climatology. From this dataset, we then zonally-average along each latitude to create the zonal average climatologies above. The reanalysis indicates the peak heat input occurs in summer (DJF) at all latitudes, as expected. The northward slant of the -20 W/m2 contour from 70°S in March to ~65°S in September is associated with the reduction in heat loss from the ocean due to the growth of sea ice – some of the models have different slants and some have displaced ice edges. All of the models overestimate the net amount of heat lost by the ocean in winter (MJJA).

**Figure S4:** Annual mean difference between evaporation and precipitation (E-P, m/yr, negative values or blue shading indicate addition of freshwater through the surface) from the reanalysis and from 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. Panel a) CFSR dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. Panel H shows the zonally-integrated annual mean, net evaporation minus precipitation (in 106 m2/yr) for each of the other seven panels.

Figure S4 shows the annual mean water flux at the surface due to evaporation and precipitation in m/yr. The monthly precipitation flux was subtracted from the monthly evaporation flux, and then all months in the period of record were averaged together. The flux in kg/m2/s was then converted to m/yr by dividing by the density of water (1000 kg/m3). Positive values (E – P > 0, yellow/brown) indicate a net flux of water from the ocean to the atmosphere and an increase in local salinity, while negative values (E – P < 0, blue) indicate a net addition of water to the ocean and a reduction in local salinity. Panel H show the zonal integral of the other seven panels (over the ocean only). None of the models capture the intense band of precipitation seen in the reanalysis at 65°S in Panel H. The models show reasonably good agreement with each other, although the MIROC model has much more precipitation between 50°S and 45°S that the other models do.

**Figure S5:** Annual-mean sea surface height (cm) from the observations and from 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. Panel a) AVISO dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. Panel H) shows the zonal mean for each of the seven maps. The area-mean sea surface height in the Southern Ocean (90°S-30°S) was subtracted from each model.

Figure S5 shows the annual mean sea surface height (in cm), averaged over all the months in the period of record. In order to make the maps directly comparable, the area-weighted average sea surface height over the entire Southern Ocean (south of 30°S) was linearly removed from each point. The slope of the sea surface across the Antarctic Circumpolar Current is directly related to the strength of the barotropic transport in the ACC. The models generally agree with each other, and the models with smallest difference in height across the current (MRI and CSIRO) have the weakest transports through Drake Passage.

**Figure S6:** The winter (JJAS) maximum mixed layer depth (MLD) over the Southern Ocean from the observations and the model simulations. MLD was defined similarly to the original Levitus 1982 definition: the depth at which the density has increased from its surface value by 0.125 kg/m3 (Levitus, 1982, Monterey and Levitus 1997). All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA2013 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. In panel G) for the MRI simulation, the white area around 60°S in the Indian Ocean sector has persistent convection down to the bottom, so the density at every depth are less than 0.125 kg/m3 different from the surface value, making the MLD undefined in this area.

Figure S6 shows the maximum winter (JJAS) mixed layer depth (MLD) in meters. The MLD is defined as the depth at which the potential density (referenced to the surface, ) is 0.125 kg/m3 larger than the surface value at that location, based the definition given by Levitus (1982) and Monterey and Levitus (1997). Much finer-scale definitions of the MLD are used for measurements from individual profiles and higher resolution (in space and time) data, but for monthly averages on a coarse-to-moderate grid, this definition is appropriate. The potential density was determined from the climatological temperature and salinity each month and then the depth of the deepest MLD in any of the winter months is shown. The simulations capture the basic pattern of maximum MLD along the Subantarctic Front (not shown), although the MLD in the HadGEM simulation is significantly shallower than observed. Two of the models, ESM2M in the Weddell Sea and MRI south of the Kerguelen Plateau have persistent convection down to the bottom, so the density in those locations is always close to the surface value and the MLD is undefined (at the bottom).

**Figure S7:** Upper ocean heat content in the top 400m (units of 109 J/m2) from the observations (WOA13) and the simulated differences for each of the models. Heat content is calculated by integrating the temperature over the top 400m and then multiplying by the specific heat (3895 J/kg-K) and a reference density (1030 kg/m3, as in Stephenson et al., 2012). All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA2013 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1.

Figure S7 shows the annual-mean upper ocean heat content (in 109 J/m2) from the WOA13 dataset, and the difference (anomaly/bias/error) in each of the model simulations calculated by averaging all months in the period of record. Heat content is calculated by integrating the temperature (in K) from 0-400m and then multiplying by the specific heat and a reference density. In order to determine the difference from observed for the models, each simulation was regridded onto the standard 1°-by-1° grid of WOA13: values at the standard points on the new grid are determined by linearly interpolating from the nearest 4 points on the old grid. Common features in the models include warm biases downstream from Africa in the South Indian and downstream from New Zealand in the South Pacific, generally accompanied by cool biases to the south over the ACC. These indicate too little mixing of subtropical gyre water into the northern flank of the ACC. The South Atlantic around 30°S is generally cooler than observed.

**Figure S8:** Meridional slope of the annual-mean depth of the σθ=27.6 isopycnal surface from the World Ocean Atlas 2013 in meters per degree of latitude. Red shading indicates a deepening of the isopycnal toward the north (away from Antarctica). All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA2013 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1.

Figure S8 shows the meridional slope of the σθ = 27.6 isopycnal surface in meters of depth per degree of latitude. Steeper isopycnals can be expected to lead to more energetic eddy fields through enhanced baroclinic instability and higher levels of the available potential energy. Intense upwelling around Antarctica, driven by the surface Ekman divergence leads to sharply sloping isopycnals and presumably high levels of eddy activity. The isopycnal slopes are therefore an important proxy for internal ACC dynamics and eddy-induced transports. Meridional eddy-induced heat transports play a key role in the ventilation along isopycnals that outcrop in the ACC. The meridional slopes also control the baroclinic part of the ACC, through the thermal-wind balance. In order to create the figure, we first determined the depth of the isopycnal surface and then took the meridional derivative of this field. The steepest sloping isopycnals generally occur over bottom topography (Drake Passage and the Scotia Arc, the Eltanin and Udintsev Fracture Zones, south of the Campbell Plateau, near Kerguelen, etc.)

**Figure S9:** Simulated annual density (σθ) averaged from 32°-30°S in each of the IPCC models and the observations. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA2013 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. Panels H, I and J show the zonal-mean simulated annual potential density, in each of the Indian, Pacific and Atlantic basins, respectively. Profiles include: WOA2013 (black); CanESM2 (pink); CSIRO-Mk3.6 (red); GFDL-ESM2M (blue); HadGEM2-ES (green); MIROC-ESM (magenta); and MRI-ESM1 (cyan).

Figure S9 shows the annual mean density (again σθ, referenced to the surface) averaged between 32°S and 30°S at the northern edge of the Southern Ocean. Isopycnal depths here are an important indicator of the stratification and water mass properties. Isopycnals that outcrop to the surface effectively link the Southern Ocean mixed layer and the deeper layers at mid- and low-latitudes. Along-isopycnal transports of heat and biogeochemical tracers can therefore supply these quantities to the intermediate and deep layers in the Pacific, Atlantic and Indian basins. Stratification at 30°S is an important metric that determines the penetration depth of heat and carbon anomalies and helps to interpret biases in the lateral fluxes across this section. It is clear that the model simulations are consistently too weakly stratified at 30°S near the surface, and all of the models have consistently too-light water between 100-700m in the Indian and between 300-800m in the Pacific. Weak stratification can also be caused by overestimated vertical diffusion, including spurious numerical diapycnal mixing which is difficult to control in these z-coordinate models. The depth 27.5 isopycnals, which in the real ocean corresponds to the upper part of CDW, is simulated accurately in all models. Simulations in the deep ocean tend to straddle the observations, primarily due to different bottom water temperatures.

**Figure S10:** Surface pH (at the insitu temperature and total scale) calculated from the dissolved inorganic carbon (DIC) and potential alkalinity observations and the model simulations. On the left are the wintertime values (JAS), and the right panel shows the summertime values (JFM). The calculated pH are taken from the recent Takahashi et al. (2014) climatology for reference year 2005, available through CDIAC, which has a fairly coarse 5°-by-5° resolution. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) CDIAC; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1. Note that the CSIRO model is a coupled climate model, not an Earth System model and therefore does not simulate the pH. The MIROC model only has annual mean data available through PCMDI. The pH observations are based on surface water DIC, nitrate, and alkalinity data primarily from the GLODAP, CARINA, and LDEO databases taken from the top 50m of the water column.

Figure S10 shows the surface (0-50m average) pH in the winter (JAS, left) and summer (JFM) Climatologies are created by averaging all similar months in the period of record (e.g. all Januarys were averaged to create the January climatology, etc.), and the months were compiled to create the annual climatology. The Ekman-driven surface divergence brings old, carbon-rich, low pH water to the surface: this is generally captured by the models with some differences between the specific pH values present. Several of the simulations have excessively alkaline waters north of the ACC and several have too acidic water in the upwelling region. The seasonal differences seen in the observations are also clearly seen in the HadGEM simulation and possibly in ESM2M. The MRI simulation clearly has more acidic (lower pH) water at the surface during the summer.

**Figure S11:** (TOP PANELS) Surface (0-100m), (LOWER LEFT) zonal mean Atlantic, and (LOWER RIGHT) zonal mean Pacific concentrations of dissolved inorganic carbon (DIC) from the observations and the model simulations. The observations are from the GLODAPv2 dataset. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) GLODAP; b) CanESM2; c) GFDL-ESM2M; d) HadGEM2-ES; e) MIROC-ESM; and f) MRI-ESM1. Note that the color scale is different for the surface (upper panels) than for the zonal averages in each basin (lower panels).

Figure S11 shows the annual mean concentration of dissolved inorganic carbon (DIC, in mol/kg) at the surface (0-100m average, top panels), in the Atlantic (zonal average, lower left panels), and the Pacific (zonal average, lower right panels). This figure was created by averaging all of the months in the period of record; the Atlantic is the zonal average from 60°W-20°E, and the Pacific is the zonal average from 140°E-60°W. The Southern Ocean is responsible for about 50% of the global uptake of anthropogenic CO2 by the ocean, despite being only ~30% of the total ocean area (Gruber et al. 2009, Takahashi et al. 2009, 2012). The total amount of carbon in the surface ocean, along with the pH, determines the surface pCO2 and therefore greatly affects the air/sea exchange of carbon. Simulations of the surface carbon can be expected to affect the simulated uptake of CO2 in transient forcing scenarios, and therefore the global atmospheric temperature response to these scenarios. There are significantly different amounts of total carbon, globally and in each of the different reservoirs (atmosphere, ocean, vegetation and soil), in the various Earth System Model simulations. The amount of carbon in each reservoir can potentially affect the transient response (uptake or degassing) based on potentially unrealistic initial conditions.

**Figure S12:** (TOP PANELS) Surface (0-100m), (LOWER LEFT) zonal mean Atlantic, and (LOWER RIGHT) zonal mean Pacific concentrations of nitrate (NO3-) from the observations and the model simulations. The observations are from the 2009 World Ocean Atlas. All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA13; b) CanESM2; c) GFDL-ESM2M; d) HadGEM2-ES; e) MIROC-ESM; and f) MRI-ESM1.

Figure S12 shows the annual mean concentration of nitrate (NO3-, in mol/kg) at the surface (0-100m average, top panels), in the Atlantic (zonal average, lower left panels), and the Pacific (zonal average, lower right panels). This figure was created by averaging all of the months in the period of record; the Atlantic is the zonal average from 60°W-20°E, and the Pacific is the zonal average from 140°E-60°W. Old nutrient-rich Circumpolar Deep Water is upwelled in the Southern Ocean, but due to the lack of sunlight for part of the year, significant quantities of unused nutrients are subducted in newly formed Mode and Intermediate waters. The Southern Ocean is responsible for about 70% of the nutrients exported to the rest of the world ocean. The upwelling of nitrate and its subsequent export to the rest of the world ocean can have significant impacts ocean biogeochemical simulations. Overall, the models generally get the annual mean surface concentration of nitrate correct, although the CanESM simulation has values about 15-20% too high in the ACC, and virtually no nitrate north of the Subantarctic front. The zonal-mean sections in the Atlantic and the Pacific also show a general agreement with the observations, but here too some of the models have issues. In the Atlantic, HadGEM, and to a lesser extent MIROC have too little nitrate in the intermediate water at ~1000m, while in the Pacific the CanESM simulation again has most of its nitrate too high.

**Figure S13:** Annual-mean zonal velocity at the surface (cm/s, 0-50m average) from the observations and from 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. Panel a) SOSE; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1.

Figure S13 shows the annual mean surface (0-50m) zonal velocity (cm/s). This figure was created by averaging all of the months in the period of record. Flow at the surface is dominated by the eastward flowing Antarctic Circumpolar Current. Most of the transport is carried in the Subantarctic Front and the Subpolar front which are too narrow to be fully resolved by the medium resolution (~1°) employed in the ocean components of the CMIP5 coupled climate models. The SOSE data shown here is from the 2006-2010 assimilation using a higher resolution (1/6°) model; the braided, narrow streamlines, as are inferred from altimetry, are visible. Even at the coarser resolution, the models accurately simulate the increased current speeds south of Africa in the Agulhas Current (westward) and the Agulhas Retroflection (eastward), and the topographically-influenced increased speeds through the Drake Passage, south of the Campbell Plateau and near the Kerguelen Plateau.

**Figure S14:** A) Bathymetry of the Southern Ocean (m) with locations cited in the text indicated. Bathymetry data from ETOPO5 (5’ resolution, 1/12 degree). Panels B-G show the difference between the model bathymetry and the observed bathymetry, regridded onto a standard 1° grid. Positive values (reds) indicate where the model bathymetry is too shallow, while negative values (blues) indicate where the model bathymetry is too deep. Light gray indicates and error of less than 100m either way.

Figure 14 shows the bottom topography from the ETOPO5 (5’ arc resolution or 1/12°) dataset, as well as the difference between the model bottom topographies and the observations. In order to determine the difference from observed for the models, each simulation was regridded onto the standard 1°-by-1° grid of WOA13: values at the standard points on the new grid are determined by linearly interpolating from the nearest 4 points on the old grid. The ACC extends all the way to the bottom of the ocean and the locations of major surface features, including the Subantarctic and Polar Fronts, are tied to topographic features Geostrophic flow in the meridional direction depends on zonal pressure gradients so the meridional overturning of the Southern Ocean and the flow across 30°S is closely tied to the topography of the sea floor. Not only the average water depth matters, but also the “roughness” of the bottom, as motion over bumps leads to increased mixing and potentially export across 30°S. Bottom water formation and mode water formation in the Southern Ocean are responsible for most of the uptake of anthropogenic heat and a significant fraction of the ocean’s uptake of anthropogenic carbon, so the details of the topography will affect a simulation’s total heat and carbon storage and export.

**Figure S15:** Thickness of the very cold layer: water less than 2°C; from the observations and the model simulations. The very cold layer is an approximation of the AABW simulated by the model (in the Southern Hemisphere). All model figures cover the simulated years 1986-2005 from the HISTORICAL forcing scenario. Panel a) WOA2013 dataset; b) CanESM2; c) CSIRO-Mk3.6; d) GFDL-ESM2M; e) HadGEM2-ES; f) MIROC-ESM; and g) MRI-ESM1.

Figure 15 shows the thickness of the layer with a potential temperature under 2°C. This figure was created by averaging all of the months in the period of record. The thickness and volume of the nominal Antarctic Bottom Water layer is a good indicator of potential heat and carbon uptake in warming scenarios. Although true AABW is much colder than 2°C, there are very few sources of deep water this cold, so a temperature below 2°C indicates water that has been directly influenced by AABW. Simulations with a thick cold-water layer in the Southern Hemisphere have a robust connection between the surface and the deep. Observations from repeat hydrographic sections indicate that deep water (below 4000m) are gaining heat; model simulations that simulate this deep uptake of heat have been shown to have relatively reduced atmospheric warming under increasing greenhouse gas scenarios.