Key Points:
- New model simulations have increased supercooled liquid clouds over the Southern Ocean
- A nudged GCM can qualitatively reproduce detailed in situ aircraft observations, including size distributions and water contents
- Detailed comparisons with in situ observations provide better understanding of process biases

Simulating Observations of Southern Ocean Clouds and Implications for Climate

A. Gettelman[1] [iD], C. G. Bardeen[1] [iD], C. S. McCluskey[1] [iD], E. Järvinen[1] [iD], J. Stith[1] [iD], C. Bretherton[2] [iD], G. McFarquhar[3], C. Twohy[4] [iD], J. D'Alessandro[3], and W. Wu[3] [iD]

[1]National Center for Atmospheric Research, Boulder, CO, USA, [2]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [3]School of Meteorology, University of Oklahoma, Norman, OK, USA , [4]NorthWest Research Associates, Redmond, WA, USA

Abstract Southern Ocean (S. Ocean) clouds are important for climate prediction. Yet previous global climate models failed to accurately represent cloud phase distributions in this observation-sparse region. In this study, data from the Southern Ocean Clouds, Radiation, Aerosol, Transport Experimental Study (SOCRATES) experiment is compared to constrained simulations from a global climate model (the Community Atmosphere Model, CAM). Nudged versions of CAM are found to reproduce many of the features of detailed in situ observations, such as cloud location, cloud phase, and boundary layer structure. The simulation in CAM6 has improved its representation of S. Ocean clouds with adjustments to the ice nucleation and cloud microphysics schemes that permit more supercooled liquid. Comparisons between modeled and observed hydrometeor size distributions suggest that the modeled hydrometeor size distributions represent the dual peaked shape and form of observed distributions, which is remarkable given the scale difference between model and observations. Comparison to satellite observations of cloud physics is difficult due to model assumptions that do not match retrieval assumptions. Some biases in the model's representation of S. Ocean clouds and aerosols remain, but the detailed cloud physical parameterization provides a basis for process level improvement and direct comparisons to observations. This is crucial because cloud feedbacks and climate sensitivity are sensitive to the representation of S. Ocean clouds.

Plain Language Summary Clouds over the Southern Ocean are important for climate prediction and may influence the evolution of global temperatures. Thus, these clouds are important to represent properly in models; however, recent studies have revealed models inadequately represent Southern Ocean cloud occurrence and phase, which drive large biases in radiation and subsequent climate sensitivity. Observations from research aircraft over the Southern Ocean south of Australia are compared to simulations with a global climate model which is "nudged" to reproduce the day-to-day cloud systems which are sampled. Despite being a coarse horizontal and vertical resolution, the model is able to reproduce many details of cloud phase and water content during the flights. However, the model has some biases, and these observations have been used to improve the model to better represent cloud phase. These results point to specific observational constraints for improving model simulations.

1. Introduction

Southern Ocean (S. Ocean) clouds are important for climate, regulating both local energy input and interacting with the deep ocean circulation (Trenberth & Fasullo, 2010). Earth System Models (ESMs) have been heavily biased in this region (Trenberth & Fasullo, 2010; Tsushima et al., 2006), with too much absorption of shortwave radiation, a result of too few clouds. Some models have mitigated the biases against observations with clouds that are too bright (Bodas-Salcedo et al., 2012; Lohmann & Neubauer, 2018). It has recently been realized that one major reason for these biases has been the incorrect phase of the clouds in models. S. Ocean clouds are mostly supercooled liquid water (SLW), while many climate models represent them as ice (e.g., Bodas-Salcedo et al., 2012).

The processes that maintain supercooled liquid clouds over the S. Ocean are complex, and not well constrained. Tan et al. (2016) found that S. Ocean low clouds were sensitive to the vapor deposition (Wegener-Bergeron-Findeisen, or WBF) process and ice nucleation. Vergara-Temprado et al. (2018)
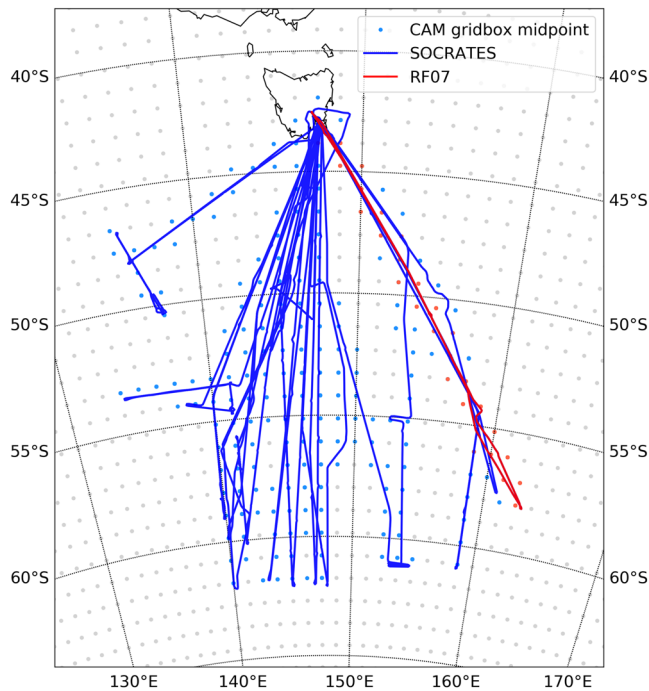
Figure 1. Map of SOCRATES mission flight tracks from the NSF G-V aircraft. Red is Flight RF07 on 31 January 2018 detailed later in the text. Solid dots indicate locations of CAM6 grid point centers used for comparison.

found S. Ocean cold-sector stratocumulous clouds were sensitive to ice nucleation schemes. McCluskey et al. (2018) found that the S. Ocean ice nucleating particle (INP) number concentrations were some of the lowest reported. Mace and Protat (2018) have found large discrepancies between satellite-derived and ship-based remote sensing cloud phase estimates; recent observations from O'Shea et al. (2017) suggest secondary ice production (SIP) may be a contributing processes for ice formation in this region and could contribute to explaining the discrepancies. S. Ocean supercooled liquid clouds have been identified as a significant contributor to cloud feedbacks and climate sensitivity: the response of the Earth system to anthropogenic radiative forcing (Bodas-Salcedo et al., 2019; Gettelman et al., 2019; Tan et al., 2016).

Previous work has highlighted the discrepancies between models (especially the Community Atmosphere Model, CAM) and observations over the S. Ocean. D'Alessandro et al. (2019) compared CAM5 simulations with in situ observations from the NSF ORCAS campaign and found a lack of supercooled liquid below temperatures of -15°C. Wu et al. (2017) compared CAM5 simulations to HIPPO campaign observations and found significant missing clouds.

To help better understand the processes controlling S. Ocean clouds, the Southern Ocean Clouds, Radiation, Aerosol, Transport Experimental Study (SOCRATES) was conducted January–March 2018 in the context of an international series of linked experiments in the Australian region of the S. Ocean. SOCRATES featured a heavily instrumented aircraft (the NSF G-V "HIAPER" aircraft) with a payload of in situ and remote sensing instrumentation (see section 2.4).

Figure 1 illustrates the SOCRATES flight tracks from Hobart, Tasmania, Australia into the S. Ocean. Flights targeted different portions of extratropical cyclones as they tracked across the S. Ocean storm track South of Tasmania in January and February 2018.

As one of the key goals of SOCRATES was to evaluate and improve cloud and aerosol processes in ESMs, detailed simulations of the SOCRATES environment and flight tracks were conducted and compared to observations. This will form the basis for subsequent work comparing specific processes and observations as part of SOCRATES data analysis. In this work we describe constrained model simulations that enable even a coarse-resolution climate model to be compared to detailed in situ and remote sensing observations of multiple cloud microphysical properties. Xie et al. (2008) have shown how forecast experiments can compare climate models to bulk cloud properties. We evaluate model simulations with a state of the art ESM and conduct sensitivity tests of different cloud processes. We then illustrate how the observations can inform and constrain cloud processes, which are important for climate projections. We go from detailed size distributions of hydrometeors all the way to climate simulations and show how observations can help constrain climate models at many levels.

Section 2 contains a description of the model formulation, simulations, and observations. Section 3 presents the core results and evaluation of the model simulations, including campaign averages, selected cases, sensitivity tests, and the global implications. Discussion is in section 4, and conclusions and ideas for future work in section 5.

2. Methods

2.1. Model

The CAM Version 6 (CAM6) is the atmospheric component of the Community Earth System Model Version 2 (Danabasoglu et al., 2020). CAM6 features a two-moment stratiform cloud microphysics scheme, MG2 (Gettelman et al., 2015; Gettelman & Morrison, 2015), with prognostic liquid, ice, rain, and snow hydrometeor classes. MG2 permits ice supersaturation.

MG2 is also coupled to a four-mode aerosol model (Liu et al., 2016) with liquid activation following Abdul-Razzak and Ghan (2002). Marine biogenic aerosol over the S. Ocean in CAM6 is limited to a representation of DMS emissions of sulfur. CAM6 includes a physically based ice mixed phase dust ice nucleation scheme (Hoose et al., 2010) with modifications for a distribution of contact angles Wang et al. (2014) and accounts for preexisting ice in the cirrus ice nucleation of Liu and Penner (2005) as described by Shi et al. (2015).

MG2 is coupled to a unified moist turbulence scheme, Cloud Layers Unified by Binormals (CLUBB), developed by Golaz et al. (2002) and Larson et al. (2002) and implemented in CAM by Bogenschutz et al. (2013). CLUBB handles stratiform clouds, boundary layer moist turbulence, and shallow convective motions. CAM6 also has an ensemble plume mass flux deep convection scheme described by Zhang and McFarlane (1995) and Neale et al. (2008), which has very simple microphysics. The radiation scheme is the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) (Iacono et al., 2000).

CAM6 is the result of a long development process that concluded shortly after the SOCRATES campaign. For comparison (see below) we also include simulations using the older version of the model, CAM5 (Neale et al., 2010). CAM5 had a different treatment of boundary layer and shallow convective turbulence (Bretherton & Park, 2009; Park & Bretherton, 2009) and a simpler treatment of cloud microphysics and supercooled liquid (Gettelman et al., 2010; Morrison & Gettelman, 2008) with ice nucleation in the mixed phase a function of temperature following Meyers et al. (1992).

2.2. Model Configuration

CAM6 is run in a "nudged" (or specified dynamics) configuration with standard 32 vertical levels from the surface to 3hPa, a 30 min time step and horizontal resolution of 0.9° latitude by 1.25° longitude. The resolution of the model is shown by marking the model grid point centers on Figure 1. Nudging means that winds and optionally temperatures are relaxed to an analysis system, in this case the NASA Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) (Molod et al., 2015). Data are read in from files every 3 hr and linearly interpolated to the model time. Sea surface temperatures (SSTs) are also read from the MERRA2 analysis. Two critical elements are worth noting. First, the model uses a 24 hr relaxation time to the MERRA2 winds (U, V) and temperatures (T). Second, the MERRA2 analysis is interpolated in the vertical to the CAM6 vertical level structure. These two adjustments were found to enable a global simulation to reproduce the top-of-atmosphere balance of a free-running CAM6 simulation to within 2 W m^{-2}, so that the "climate" of the free-running simulation is the same with nudged U, V and T.

Simulations were spun up for 1 year using 2017 meteorology. The model was then restarted from 1 January 2018 and run over the SOCRATES flight period for 2 months. Model output is archived along the flight tracks and is sampled at 1 min resolution.

2.3. Sensitivity Tests

We conduct several sensitivity tests with the same configuration described above (Table 1). *CAM6* is the control case. *CAM5* uses physical parameterizations as described by Neale et al. (2010). *Meyers* switches the CAM6 dust-dependent mixed phase ice nucleation (Hoose et al., 2010; Wang et al., 2014) back to the temperature dependence of Meyers et al. (1992). *Berg0.25* reduces the efficiency of the vapor deposition (Wegner-Bergeron-Findeisen, or WBF) process by 75%. *SIP* experiments modify the SIP in the MG2 scheme (Cotton et al., 1986) by either setting it to zero (*SIP0*) or increasing it by a factor of 5 (*SIP5*).

We also perform several different experiments in response to the initial comparisons in section 3. These focus around first altering the representation of rain formation (autoconversion). First, we modify the existing formation by reducing autoconversion by a factor of 10 (*Auto/10*) or by replacing the modified formulation of Khairoutdinov and Kogan (2000) with that of Seifert and Beheng (2001), as discussed by Gettelman (2015) (*SB2001*). Second, the *Eta* experiment reduces the dispersion of the size distribution of cloud drops (η in Morrison & Gettelman, 2008) by switching from the formulation of Rotstayn and Liu (2003) used in CAM6 back to that of Martin et al. (1994) used in CAM5 (Morrison & Gettelman, 2008). Two additional simulations are discussed: increasing INPs for mixed phase clouds with temperatures above -10°C in CAM6 (*In10-10*) and narrowing the CAM6 rain size distribution by setting the shape parameter of the gamma distribution (μ) to a nonzero value (*MuR = 5*).

**Table 1**
Sensitivity Tests With Nudged CAM Simulations

Name	Description
CAM6	Control
CAM5	CAM5 physical parameterizations
Meyers	Meyers et al. (1992) mixed phase ice nuc
Berg0.25	WBF efficiency $1 \rightarrow 0.25$
SIP0	No secondary ice production
SIP5	$5 \times$ secondary ice production
Auto/10	Autoconversion/10.
SB2001	Seifert and Beheng (2001) autoconversion formulation
Eta	Reduced width of size distribution
In10-10	CAM6 with increased ice nucleation (rate)
MuR = 5	Nonzero rain shape parameter ($\mu = 5$)
INx0.1	Scale INP by a factor of 0.1
CCNx2	Scale CCN by a factor of 2
Fix T	MERRA U, V, and T
Free T	No T nudging (U, V only)
Nudge 1 hr	Nudging reduced from 24 to 1 hr

CAM6 does not contain marine biogenic aerosol as an INP or cloud condensation nuclei (CCN) but does contain a representation of marine sulfur emissions as dimethyl sulfide (DMS). Specifically adding marine biogenic aerosol is a subject for future work, and a paper is in process on this. In order to better understand uncertainties in aerosol distribution and the availability of CCN and INP on the simulations, we also scale the activated numbers of droplets (CCN) and crystals (INP). We tried twice and five times CCN and show 2xCCN. Because 5xCCN has very large impacts on drop number, we do not show it. We also looked at 0.1xINP and 10xINP. The latter produces impacts on crystal number that make the climate simulations unrealistic, so we only show 0.1xINP.

We also explore the impact of nudging, by running additional simulations with temperatures and winds fixed to MERRA2 (*Fix T*), only *U* and *V* nudging, and free-running temperatures (*Free T*) and using a relaxation time scale of 1 hr nudging for winds and temperatures (*Nudge 1 hr*). These experiments help elucidate whether any temperature biases are from CAM or from the input (MERRA2) analysis. Verification of temperatures is against SOCRATES in situ data from the aircraft and dropsondes.

2.4. SOCRATES Data

During SOCRATES, the U.S. National Science Foundation (NSF) HIAPER aircraft operated by the National Center for Atmospheric Research (NCAR) was equipped with a suite of in situ and remote sensing instruments. In situ instruments included cloud microphysical probes for measurement of both liquid and ice phase. Cloud droplet spectra was measured with the Cloud Droplet Probe (CDP; Lance et al., 2010) that provides cloud droplet PSDs for particle diameters (D_p) of $2 < D_p < 50$ micron. We use the King probe (King et al., 1978) as an estimate of the liquid water content (LWC). Supercooled liquid and liquid are separated based on $T < 0°$ C and $T > 0°$ C, respectively. Total condensed water content (CWC) is defined using the total mass from the Counterflow Virtual Impactor (CVI) instrument. The uncertainty in the CVI measurement is about 15% (Baumgardner et al., 2017; Twohy et al., 1997). Ice water content (IWC) can be estimated from several imaging probes, but this requires uncertainty in categorizing shapes and has issues with size selection. So we define IWC as the residual of CWC - LWC, constrained to have a minimum of zero. Because IWC is a residual, it is subject then to large uncertainties. Improved methods for determining IWC are currently being developed for SOCRATES. IWC can also be estimated from CVI CWC and phase identification. We analyzed data from several flights and found the residual method yields results very similar to the D'Alessandro et al. (2019) phase identification method. It also has the advantage that it is straightforward to sample the model in the same way.
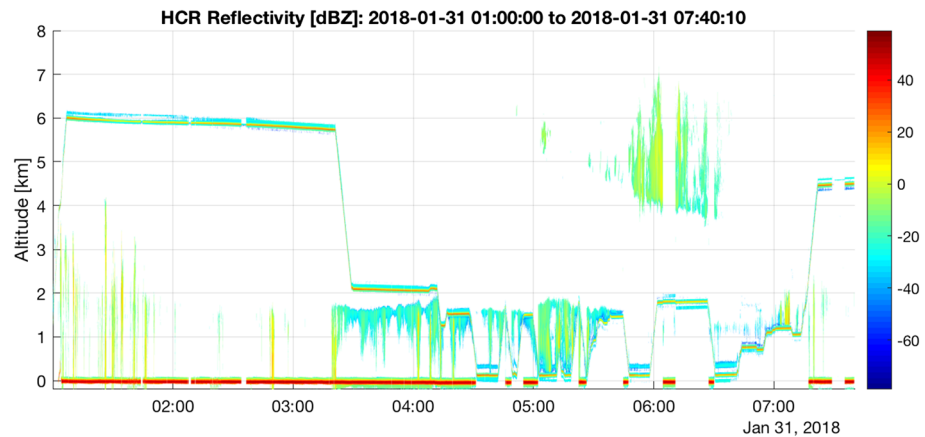
Figure 2. HIAPER Cloud Radar data from SOCRATES Research Flight 7 (RF07) illustrating flight altitude (thin red line) and observed clouds over time. The color bar indicates reflectivity in dBZ. The deep red at the surface is ground clutter when the radar is downward pointing.

CCN are obtained with the Scripps CCN counter (Roberts & Nenes, 2005). CCN are observed at a supersaturation of 0.43%.

A 2-D stereo probe (2DS) was used to determine PSDs from particle shadow graphs for particles in the size range of $0.05 < D_p < 3.2$ mm. The size limit of 2DS is 0.01 mm, but here particles below 0.05 mm are not considered due to uncertainties in the probe's depth of field and sample area (McFarquhar et al., 2017). The 2DS has a set of four arms that deliver shadow graphs both in the horizontal (H) and vertical (V) direction. During SOCRATES, the vertical direction was not working properly, and, therefore, only horizontal data (2DS-H) were used.

Relative humidity measurements use the VCSEL laser hygrometer and a temperature sensor, which combine for an uncertainty of about 7% RH (Zondlo et al., 2010).

Remote sensing probes included Radar, Lidar, and Dropsondes. The HIAPER Cloud Radar (HCR) is a W-band radar that, as operated in SOCRATES, was pointed up or down depending on the mission needs in SOCRATES. A High Spectral Resolution Lidar (HSRL) (EOL, 2018) was also used on the aircraft. A description of the dropsonde data, including data processing and quality assurance methods, is provided in Young and Vömel (2018) and Young (2018).

Additional information on HIAPER airborne data (e.g., temperature, humidity, winds, pressure, and position) and data processing methods is provided by EOL (2018) (and at https://www.eol.ucar.edu/aircraft-instrumentation).

2.5. Research Flight 7

In order to present the results and show impacts, we show campaign averages of all flights but also focus on a particular sample flight that is representative of many flights from SOCRATES. We focus on Research Flight 7 (RF07), which took place on 31 January 2018. This flight (the red line in Figure 1) targeted a region of clouds in the cold sector of an extratropical cyclone South of Macquarie island (54.6°S, 158.9°E). The clouds were of a type that kept "disappearing" in forecast models into a broken cloud deck, while satellite images continued to show solid cloud cover. The models being used in which such clouds disappeared included the European Centre for Medium Range Weather Forecasts (ECMWF) Integrated Forecast System (IFS), the National Oceanic and Atmospheric
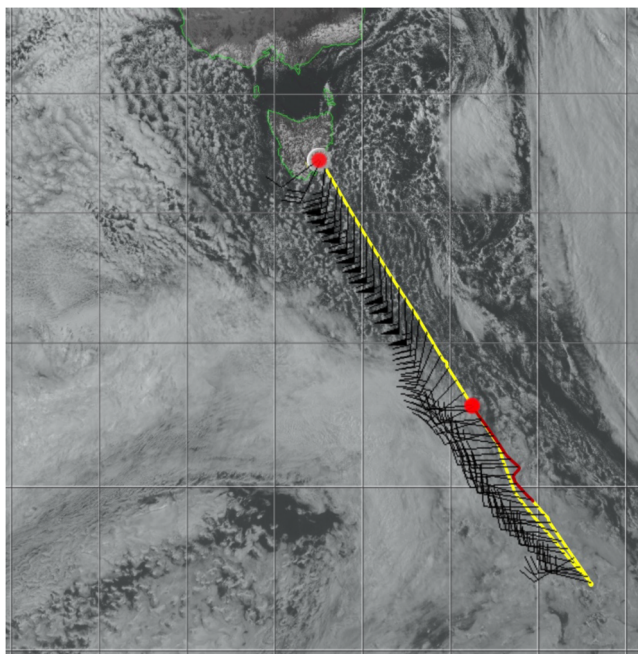
Figure 3. Himawari-8 Visible satellite image at 600 UTC, 31 January 2018 showing the cloud field. Gridlines are every 5° latitude and longitude from 40–60°S and 135–165°E. Also indicated is the aircraft flight track up to 600 UTC with wind vectors from aircraft observations along the flight track. Yellow indicates the flight track and red 500–600 UTC.

Figure 4. Aircraft forward camera image from 410 UTC near turnaround latitude.

Administration (NOAA) Global Forecast System (GFS), and the Australian Community Climate and Earth-System Simulator (ACCESS). The composite radar image from RF07 is illustrated in Figure 2.

As illustrated in Figure 2, RF07 featured broken strato-cumulus cloud between Hobart and Macquarie Island at 56°S (310 UTC). This is also seen in a Himawari-8 visible satellite image from 0600 UTC (Figure 3). After Macquarie island at about 330 UTC, the aircraft descended to above the boundary layer and turned around to head north at 412 UTC. The aircraft then began cloud sampling with an above cloud leg over a supercooled air mass. Cloud top was about 1.5 km for the whole layer, and cloud base at 1 km with some precipitation shafts extending lower. The cloud deck was solid on top, but thin with cellular structure (Wang et al., 2020). Figure 4 is a visible wing camera image of the cloud layer at 410 z just before turning north (58°S), illustrating it was optically thick. There were spots where the ocean was visible through small holes in the cloud. There was some thin cloud at 4.5–6 km in this region, seen in the distance of the image in Figure 4.

The plane then headed north, sampling in and out of the cloud layer. There was pretty significant supercooled liquid and icing on probe tips in the cloud, and the temperatures were just less than freezing (see temperature curtain in Figure 5). Near Macquarie Island (500 UTC on the return), there were multiple cloud layers, with more extensive cloud and drizzle. Mixed phase graupel or snow was seen in the particle instruments and visible in some shafts from the plane. North of Macquarie island, the lower cloud deck was more broken, and a shallow cumulus deck extended from about 1–2 km.

3. Results

In order to better characterize the flights, we show examples of model and observational comparisons from RF07 and then show how this generalizes to averages over the whole campaign and the model climatology. We use observations from the aircraft as well as broader scale satellite observations.

3.1. RF07 Results

Figure 5 illustrates temperatures along the flight track from RF07 and the base *CAM6* nudged simulation. Model temperatures are generally within 1–2° of the aircraft at all times, as the temperatures are nudged to MERRA2 with a 24 hr relaxation time. The top of the boundary layer in the cloud layer from 4 to 6 UTC is just below freezing ($T \sim 268$ K), with the ocean surface temperature just above freezing ($T \sim 280$ K).

The structure of the temperature biases is more easily seen in a comparison to the last dropsonde at 3:44 UTC (Figure 6). At 800–750 hPa, right at the top of the boundary layer, *CAM6* is missing the temperature inversion seen in the observations. The inversion transition region is narrower than the vertical resolution of the model, but even the binned average has a bias of ~2° C in this layer. The lack of resolution of the inversion results in high humidity in the layer above the boundary layer top. There is a -15% bias in the boundary layer up to 800 hPa in *CAM6*. While the zonal wind is well reproduced (perhaps too high right near the surface), the meridional wind has a significant bias.

To check whether this bias is the result of the nudging data, we fixed the temperatures and winds to MERRA2 analysis and reran the simulation (*Fix T*). Figure 7 shows the comparison between MERRA2 winds and temperatures and the dropsonde observations. The temperature bias at 800 hPa is reduced from 2°C to zero, leading to improved humidity above the boundary layer. But the wind biases remain. The zonal wind bias is larger than the base case at the top of the boundary layer. Thus, the wind biases may come from the input reanalysis data, while the temperature bias and inversion bias seem to be a result of CAM simulations pushing the model away from the analysis. Experiments with 1 hr nudging (*Nudge 1hr*), or no temperature nudging (*Free T*), confirm this trend: 1 hr nudging has an intermediate temperature bias between analysis temperatures (Figure 7) and 24 hr nudged temperatures (Figure 6), while no temperature nudging yields a larger bias than 24 hr nudging in Figure 6.
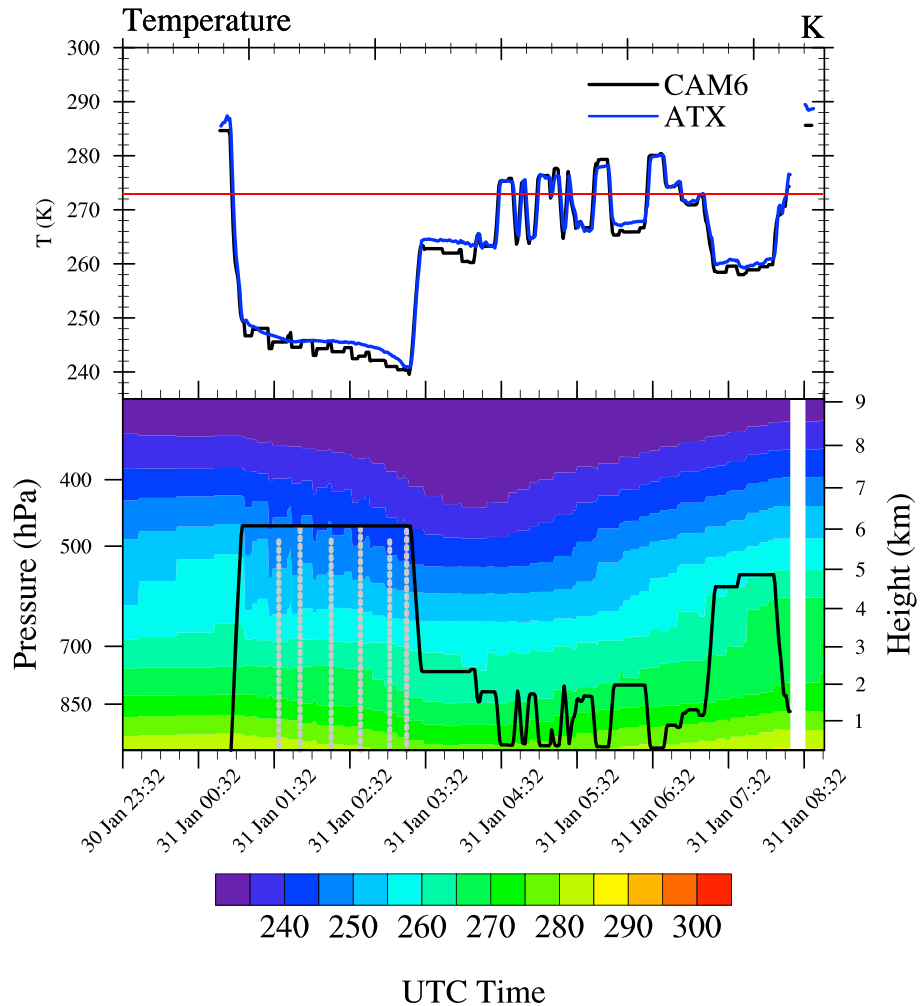
Figure 5. Temperatures along the flight track from RF07, showing the entire flight as a function of time from left to right. Note that latitude decreases (southward flight to 4:11 UTC and then increases again as the plane turned around). Freezing level (273K) is the thin red line. The bottom panel shows aircraft altitude (solid black) and dropsonde locations (dashed gray) on top of the simulated temperature curtain from the CAM6 base case. Top panel illustrates the aircraft temperature at flight level (ATX blue) and model temperature (CAM6) interpolated to the flight level (black).

Figure 8 illustrates that these temperature biases are a general feature of the CAM6 simulations for the whole campaign (119 dropsondes). There are consistent ∼1° (range of -2.5 to 0) temperature biases at the top of the PBL, indicating the lack of an inversion in the base CAM6 24 hr nudged simulation. Associated with this temperature bias is a positive ∼20% relative humidity bias. This is a specific humidity bias, since much of the bias remains even if the temperature bias is reduced (red line in Figure 8c and both red and black in Figure 8d). Figure 8b indicates that this is not due to the input data, as the MERRA2 reanalysis temperatures are on average only 0.2°C colder than the dropsondes (range of -1 to 0). This also significantly reduces the humidity bias (Figure 8d) and reduces the error due to temperature (compare to red line in Figures 8c and 8d). Note that the fixed temperature (MERRA2) simulation does have interactive (not fixed) specific humidity. Also note that the dropsondes are generally dropped from altitude near the beginning of a flight and so are not coincident with many of the low cloud decks. They likely represent a more clear-sky PBL.

Figure 9 extends the analysis of Figure 8 from dropsonde points to the entire campaign of aircraft locations for temperature (Figure 9a) and relative humidity (RH) (Figure 9d). The shaded range represents 2/3 of the data (between the 16th and 83rd percentiles), showing the large variability across the different flights. Only CAM data from the flight level is used for direct comparison to observations. Here the bias in PBL structure
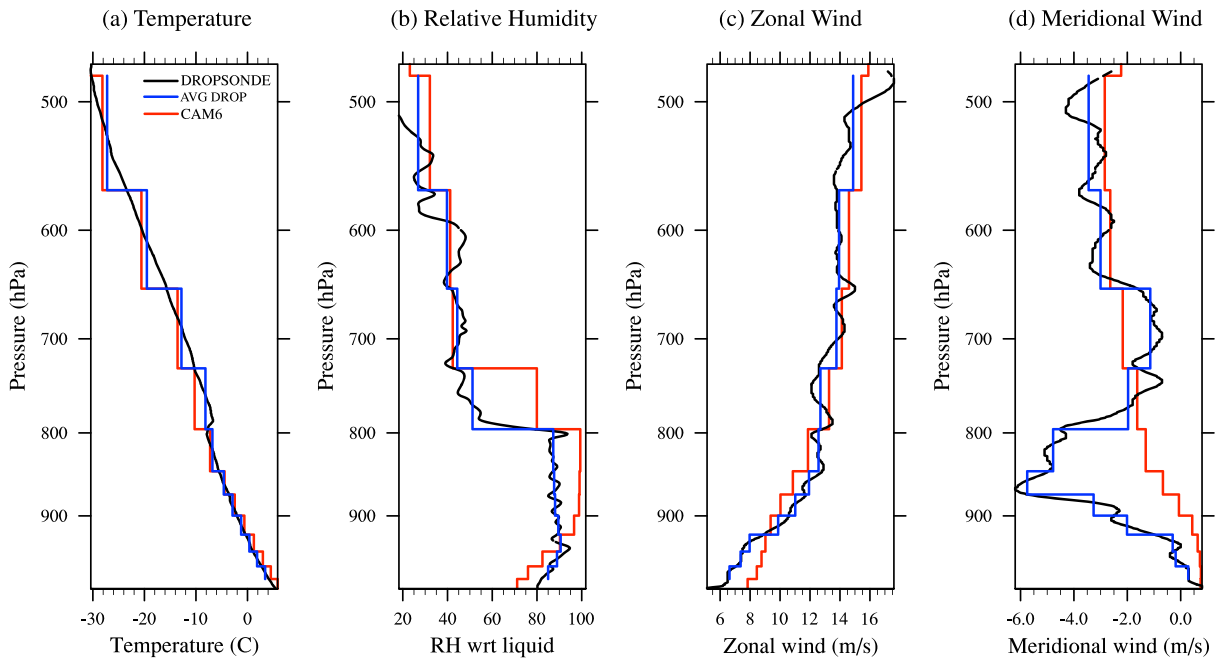
Figure 6. Comparison of dropsonde between CAM6 (red), dropsonde (black), and dropsonde binned to CAM6 levels (blue). (a) Temperature, (b) relative humidity with respect to liquid (RH wrt liquid, %), (c) zonal wind (m/s), and (d) meridional wind (m/s).

is not that evident in temperature (Figure 9a), and the inversion is evident in CAM. These points are more representative of cloudy boundary layers sampled similarly to RF07 (Figures 5 and 10). The positive RH bias remains (Figure 9d) both at lower altitudes and in the free troposphere. Since humidity is not fixed by the analysis, the large-scale advection of free tropospheric RH may have a bias.

Also illustrated in Figure 9 are the water contents in different phases. From observations, the CVI and King probe were used to estimate total CWC and LWC mass concentrations, respectively. Supercooled liquid is LWC for $T < 0°C$. IWC is defined as IWC = CWC - LWC when CWC > 0. Note that because the CWC
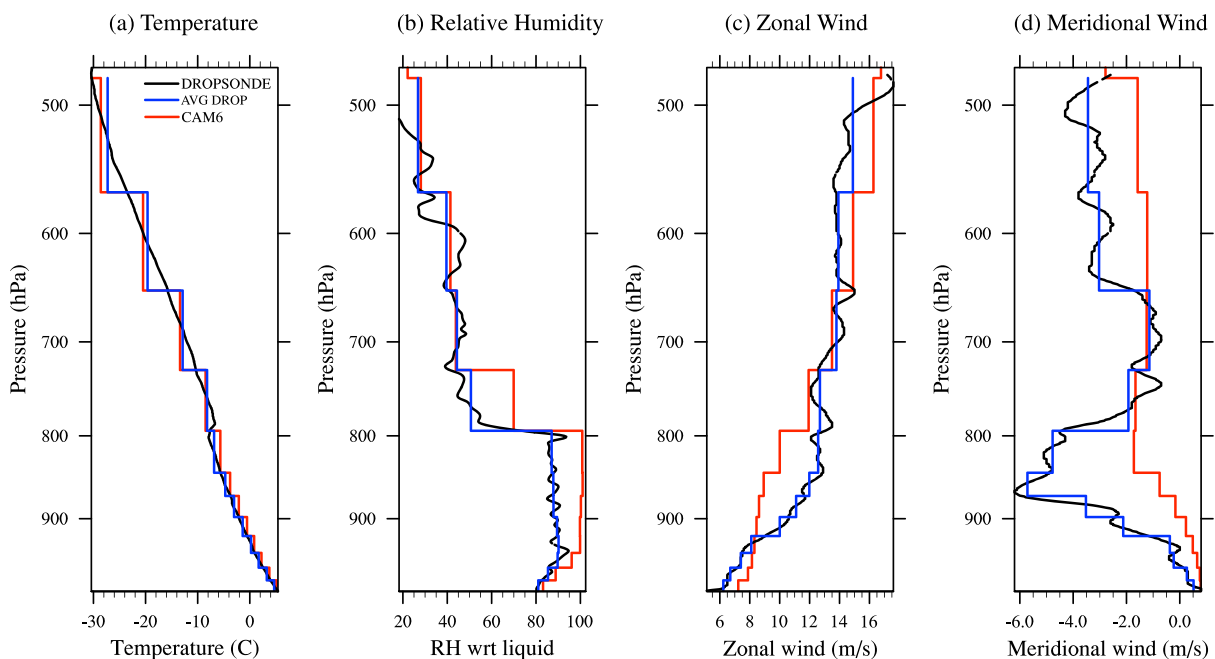
Figure 7. The same as Figure 6 but for a CAM6 simulation with fixed MERRA-2 temperatures and winds.
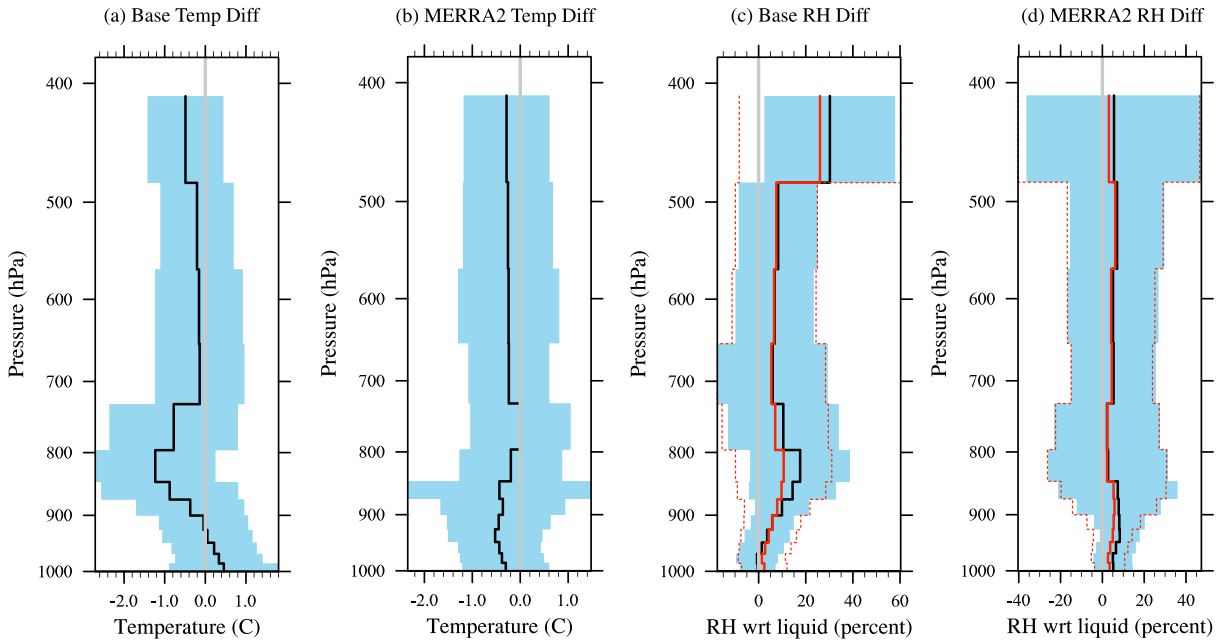
**Figure 8.** Average temperature (a and b) and mean relative humidity (c and d) differences (black line) from SOCRATES dropsondes at the sonde locations and times from CAM6 24 hr nudged simulation (a and c) and using fixed MERRA-2 Temperatures (b and d). Red line in c and d (RH plots) is the simulated RH estimated assuming the radiosonde temperature for saturation (no temperature error). Light blue shading is the range of all radiosonde differences; red dashed lines in c and d are the range of all RH differences where simulated RH is estimated using model specific humidity and saturation humidity is estimated using dropsonde temperature.
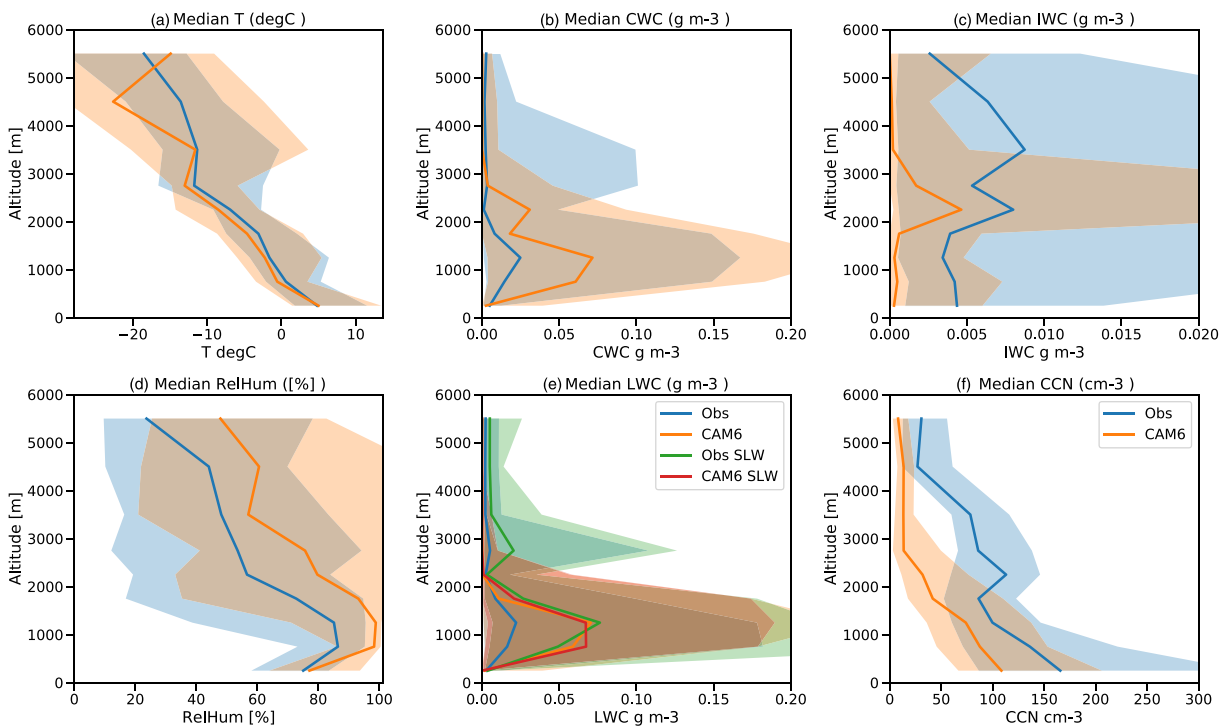
Figure 9. Median binned by altitude for the entire SOCRATES campaign from the aircraft (Obs, blue) and colocated CAM6 simulation (CAM6, orange). In each panel, median value is the solid line. The shaded range is the data between the 16th and 83rd quantiles (67% of the data). (a) Temperature, (b) condensed water content (CWC), (c) ice water content (IWC), (d) relative humidity (RelHum), (e) liquid water content (LWC) for all temperatures, and (f) cloud condensation nuclei (CCN). Supercooled liquid water (SLW) when $T < 0°C$ only is shown in (e) from observations (green) and CAM6 (red).
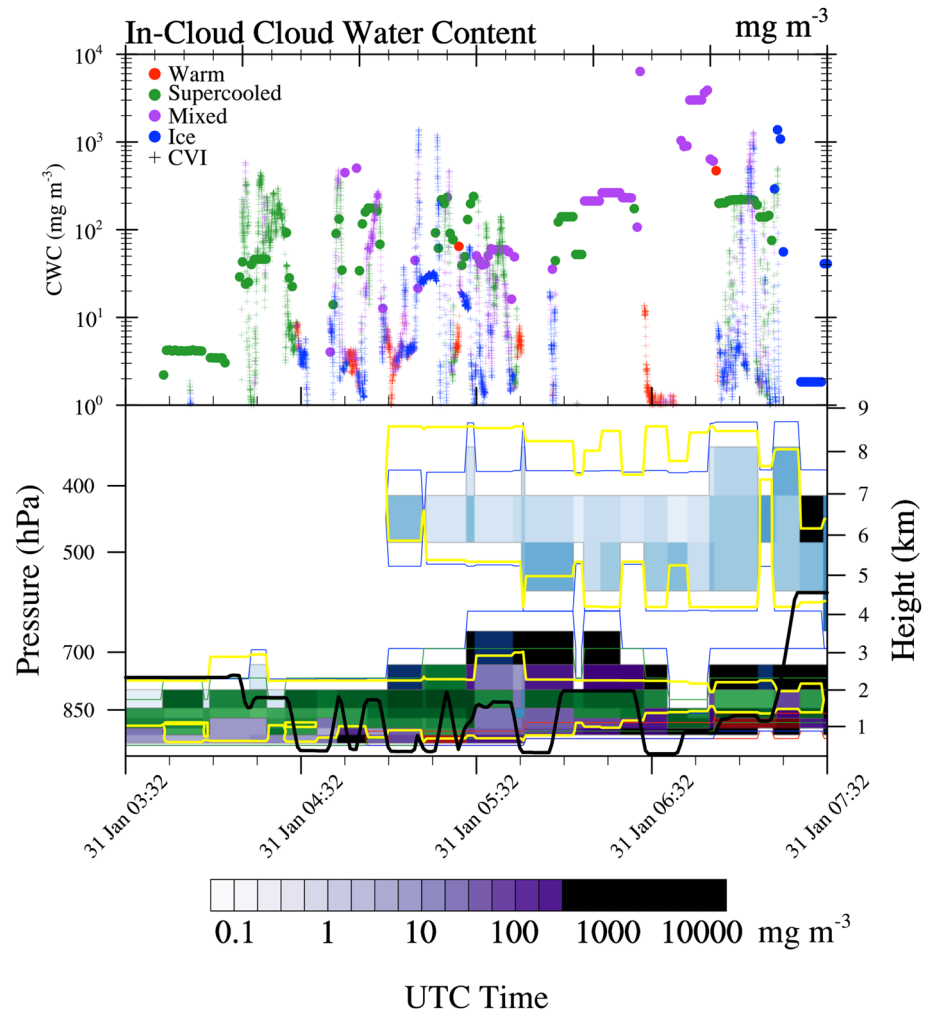
**Figure 10.** Cloud hydrometeors along the fight track from RF07, showing the low cloud section of the fight as a function of time from left (3:32 UTC) to right (7:32 UTC). Southward fight to 4:11 UTC and then northward after that. For both panels, dominant cloud hydrometeor phase (see text for definition) is shown by color: liquid (red), ice (blue) supercooled liquid (green), and mixed phase (purple). Top panel: 4 s averaged aircraft data along the flight track from the CVI (plus symbols) representing total condensed water content (CWC) with colocated model simulation sampled every minute in solid circles. For observations, CWC = CVI (CVCWCC) and liquid water content LWC = King (PLWCC). IWC is the positive difference (IWC = CWC - LWC). Simulation LWC is rain and cloud liquid, IWC is snow and cloud ice, and simulated CWC = LWC + IWC. Bottom panel: aircraft altitude (solid black) on top of the simulated cloud mass from the CAM6 base case. Increased intensity of the color indicates higher water content. The colorbar shows the scale for mixed phase. Yellow contour is cloud fraction greater than 10%. The thin blue, red, green, and purple lines correspond to the extent of the ice, liquid, supercooled liquid, and mixed phase within modeled clouds.

measurement from the CVI and LWC measurement from the King probe have uncertainties on the order of 15%, some low IWC events will be missed. The king probe might also miss some large drizzle and rain drops captured by the CVI. CAM6-simulated LWC is rain and cloud liquid, IWC is snow and cloud ice, and simulated CWC = LWC + IWC.

CAM6 reproduces (to within the variability of the data sets) the mean total CWC Figure 9b across all the flights. CAM6 simulated CWC is lower, a function of reduced IWC (Figure 9c). We expect some bias to the average observed aircraft IWC since it is a residual of CWC - LWC, and small values are harder to measure. The comparisons get better in the region of mixed phase (Alt < 2,000 m) if a threshold of IWC > 15% of CWC is put on CAM simulations. SLW is well reproduced in the boundary layer clouds (Figure 9e). When SLW is found, it has a higher concentration than for $T > 0°$ in the observations. Almost all of the LWC
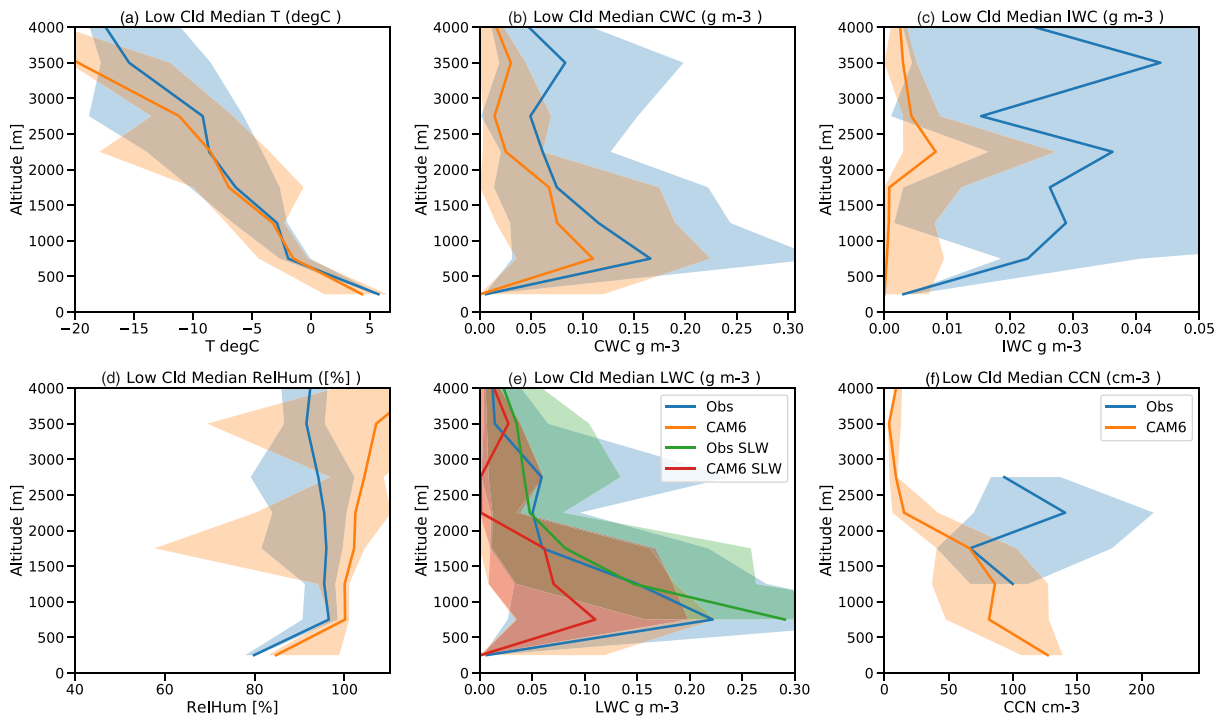
Figure 11. The same as Figure 9 but for only selected low cloud cases as described in the text.

in CAM6 is supercooled, and it has the same concentration. This is consistent with less ice in the PBL in CAM. Finally, CAM reproduces the vertical structure of the average number of CCN observed in the boundary layer (Figure 9f), with a low bias to the CCN in and above the boundary layer. CCN are observed at a supersaturation of 0.43% and compared to simulated CCN at 0.5% supersaturation.

Figure 10 illustrates a curtain of cloud hydrometeors (liquid, ice, and supercooled liquid) observed and simulated for the low cloud portion of RF07. The water contents are estimated as in Figure 9. The CVI and King probe were used to estimate liquid and ice mass concentrations, respectively. The dominant phase is determined by phase ratio (PR = IWC/CWC) where ice is where PR > 0.9, mixed is 0.9 > PR > 0.1, liquid is PR < 0.1, and supercooled liquid is PR < 0.1 and temperature $T < 0°C$. The low cloud layer sampled in the observations and model from 4–6 UTC is a mix of supercooled liquid and ice, of about the same mass concentration. CAM does not capture the peaks in the CWC distribution. Clouds are present at the top of the PBL, with no cloud in the surface layers. The dominant hydrometeor for much of this time in both models and observations is supercooled liquid and mixed phase (where the liquid is also supercooled), which appears to be about the right mass over the flight, with wide variation of the liquid and ice in the model and observations.

Figure 11 extends the low cloud comparison of Figure 10 across the campaign, by taking a subset of the data in similar to Figure 9 for low cloud events. Low cloud events are defined using the observations as times when CWC > 0.001 g m^{-3} for more than 10 s, grouped into events within 2 min of each other that in total span at least 10 min. This represents events sampling multiple low clouds in a similar air mass. Entirely warm clouds ($T > 0°$ C) were excluded. There are between 2 and 6 events per flight, and the events last from 10–60 min transits (altitude > 6,000 m) were excluded. All times were included in the event, and the model is sampled over the same time range and altitude as the in situ observations. For these cold low cloud events, the model reproduces (within the variability) boundary layer temperature (Figure 11a) and humidity (Figure 11d), with an anomaly in the topmost bin (related to the CAM vertical level structure). In the free troposphere above 2,000 m, RH is biased high in the model during these events but is constant with altitude as are the observations. The RH bias may be related to a warm temperature bias from 3,000–4,000 m. Low cloud CWC (Figure 11b) is similar to observed in the boundary layer, while the model has a low IWC bias during low cloud events: They are mostly SLW (Figure 11e) in both the observations and model, giving the same impression of 50% less LWC as Figure 11b for CWC. Limited CCN measurements are available in these
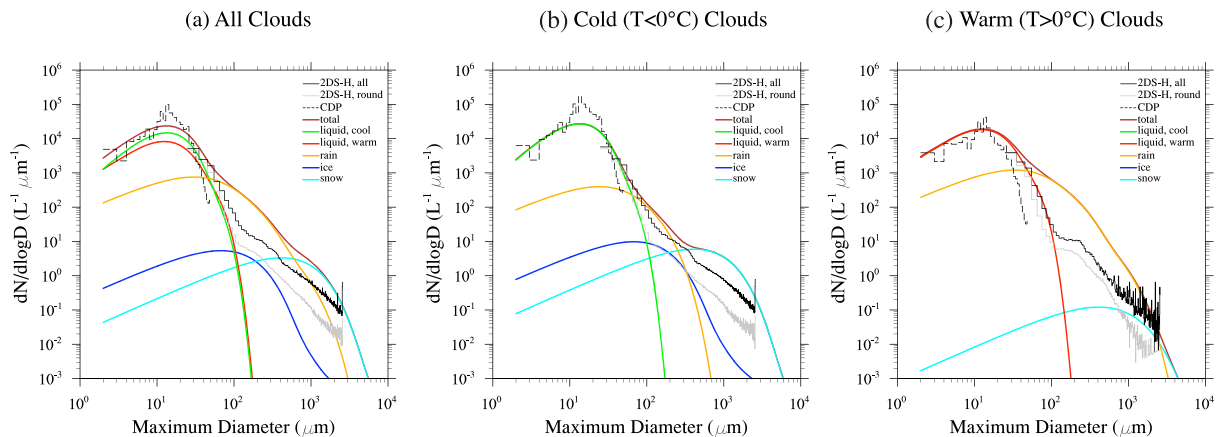
Figure 12. Size distributions from observations (thin lines) and reconstructed model hydrometeor size distributions (thick colored lines) for low level clouds ($P > 750$ mb) as indicated in the legend. Selected cloud probe data shown as 2DS for all particles (thin black) and round particles (thin gray) and CDP (thin black dotted). (a) All clouds, (b) cold clouds, and (c) warm clouds. Model is sampled along the flight track at aircraft altitude.

legs except above low cloud (Figure 11f, where the model is lower than observed, consistent with Figure 9f for all flights).

SOCRATES is a unique campaign for its extensive sampling of cloud drop and crystal size distributions in S. Ocean supercooled liquid clouds. CAM6 is uniquely placed to take advantage of this evaluation opportunity, since the two-moment microphysics scheme (Gettelman & Morrison, 2015; Morrison & Gettelman, 2008) has a prognostic representation of the size distribution and uses information from an aerosol scheme that estimates CCN to create drops. Here we use the moments of the size distribution with the functional form of the gamma distribution assumed in the MG2 scheme, to reconstruct the size distribution for all the hydrometeors (liquid, ice, rain, and snow) in CAM6, and compare this to observations from the suite of instruments on the GV aircraft during SOCRATES. MG2 uses gamma functions for the size distributions of all hydrometeors, with variable shape (dispersion) for liquid, and fixed shape ($\mu = 0$, exponential) for ice, snow, and rain (Morrison & Gettelman, 2008). Figure 12 illustrates the reconstructed distributions for (a) all, (b) cold ($T < 0°$C), and (c) warm ($T > 0°$C) clouds at pressures greater than 750 hPa, to isolate shallow clouds near the surface. The model is sampled along the flight track at aircraft altitude and instantaneous model output for mass, and number is used with the assumed functional form of the gamma distributions for each class of hydrometeor to reconstruct the size distribution. These are then averaged and compared to the accumulated sizes from the measurements.

Note the extreme scale separation for this comparison. Observed size distributions for in situ instruments are constructed from 1 Hz data, representing a sample volume of few cm^2 cross section and 150 m of flight distance (in 1 s). However, with about 5,000 samples total, this yields 750 km of cloud sampling. Simulated size distributions are assumed functional averages of a single "in-cloud" number per grid volume, typically 100 km × 100 km horizontal by 200 m vertical. Given the limitations of a functional size distribution (e.g., fixed width), CAM6 does a remarkably good job at reproducing size distributions observed from the aircraft. Individual flights have similar characteristics.

Several aspects are notable. First, the size distribution for warm liquid clouds looks reasonable (Figure 12c) with a peak between 10 and 20 μm. However, for cold clouds, in general, there does not seem to be enough supercooled liquid (see below for a discussion of sensitivity tests); this is consistent with the mean mass profiles from Figures 9e and 11e. The size distribution appears to be broader than observed from the aircraft cloud probes, with not enough peak number concentration. This may be consistent with lower CCN in the simulations for low cloud cases (Figure 11f). The snow size distribution seems well reproduced (Figure 12b), but there appears to be too much warm rain (Figure 12c), leading to too many cloud drops between 50 and 300 μm, though the 50–200 μm range is a difficult area for instruments to observe, and there are discrepancies between the instrumentation. A similar plot for only flight RF07 indicates 10% less warm rain, and 10% more liquid in the shallow clouds for this flight, but the amount of liquid is still underrepresented relative to the measurements. It may be that the rain formation process (autoconversion and accretion) is too active in the model, leading to less water remaining in clouds.

Figure 13. January–February 2018 2 month mean over 65–45°S and 135–160°E of (a) liquid water path (TGCLDLWP), (b) ice water path (TGCLDIWP), (c) total cloud cover (CLDTOT), (d) cloud top effective radius (ACTREL), (e) cloud top drop number concentration (ACTNL), (f) top-of-atmosphere shortwave cloud radiative effect (SW CRE), and (g) total cloud optical depth. Sensitivity tests from CAM as described in Table 1 colors, and CERES observations in red where available.

We conducted an experiment to reduce the width of the size distribution for liquid (*Eta*) by changing the shape parameter (η) of the gamma function in the cloud microphysics parameterization (Gettelman & Morrison, 2015). This did indeed reduce the width to look more like the observations in Figure 12, as illustrated in Figure S1 in the supporting information. However, decreased width does not significantly increase the number of supercooled liquid drops, which actually drops at cloud top (Figure 13e). Increasing CCN by a factor of 2 increased drop number by 20% (Figure 13e) but is not sufficient to reproduce the peak in the size distribution (Figure S1). Reduced autoconversion (*Auto/10*) also results in small increases in the total liquid number (Figure 13e) associated with increases in liquid water path (Figures 13a and S1). Conversely in the *SB2001* autoconversion experiment, there are decreases in total number (Figures 13e and S1) and liquid water path (Figure 13a) from the base case. This implies accretion may be more important than autoconversion.

Narrowing the size distribution for rain from an exponential (shape parameter $\mu = 0$) to $\mu = 5$ (*MuR = 5*) reduced the larger rain sizes as expected (Figure S1) but significantly increased rain mass (due to reduced fall speed for smaller raindrops), not improving the comparison to observations, and not changing total number of liquid.

3.2. Sensitivity Tests

We now turn to sensitivity tests where we vary the model formulation to test how it impacts the cloud and radiation simulation in the SOCRATES region and how it compares to observations. For a broader perspective more related to weather and especially climate, we look at regional averages from satellite data for January and February 2018. These are taken from the Clouds and the Earth's Radiant Energy System (CERES) retrievals (Loeb et al., 2018; Wielicki et al., 1996). Specifically, we use Version 4.1 of the Energy Balance Adjusted Flux (EBAF) product (https://doi.org/10.5067/TERRA-AQUA/CERES/EBAF-TOA_L3B004.1) and of the Synoptic product (SYN) Version 4.1 (https://doi.org/10.5067/Terra+Aqua/CERES/SYN1degMonth_L3.004A). We look at monthly averages for January and February 2018, as well as daily averages over this period, and long-term 15 year climatologies to try to understand the model solutions and comparisons in a broader context (see section 3.3 below). We do not use a satellite simulator to compare to model output; however, we do sample the model cloud properties at cloud top as observed to ensure that the properties are more comparable and minimize uncertainties due to sampling. Radiation fields will not be affected that much. Liquid and ice water paths do have large remaining uncertainties due to retrievals however as noted below.

Figure 13 illustrates regional (45–65°S, 135–160°E) 2 month means from the simulations and CERES data for large-scale quantities that are important for cloud physics and for driving radiative fluxes. Higher water amounts (LWP, Figure 13a) are found with *Fixed T* or 1 hr nudging (*Nudge 1hr*), and lower LWP with free-running temps (*Free T*) or for the *CAM5* simulations. The revised Seifert and Beheng (2001) autoconversion scheme (*SB2001*) results in lower LWP, similar to *CAM5*. Ice water path (Figure 13b) is higher for *CAM5* and the Meyers et al. (1992) empirical ice nucleation as a function of temperature (*Meyers*). *Meyers* is an element of CAM5 physics. Less liquid and more ice is expected from this changes to phase partitioning. Interestingly, reduced Bergeron (*Berg0.25*, vapor deposition) results in an increase in IWP, but also an increase in LWP. The increase in liquid condensate is likely resulting in still more production of ice.

The CERES SYN LWP product mean for these 2 months in the SOCRATES region is lower than most CAM simulations except the *CAM5* and *SB2001* simulations, though it is not that well correlated with CAM simulations on a day-to-day basis. This is a similar result to that implied by the SOCRATES in situ data in Figure 9b across all flights, which will be analyzed further below. CERES LWP and IWP are estimated from an assumed particle size (10 μm for liquid and 30 μm for ice) and a retrieved optical depth from infrared reflectance (CERES SYN Edition 4 Data Quality Summary). As such, particularly for ice water path, CERES may not match the observed SOCRATES ice and snow sizes (Figure 12). Accordingly, we do not show the CERES IWP (0.2 kg m^{-2}), which is much larger than LWP in this region. In addition, CERES and modeled LWP include the entire atmospheric column, whereas Figure 12 includes only pressures >750 hPa, thus a subset of clouds. Even the 10μm liquid radius is significantly smaller than simulated (Figure 13d).

Thus, while the observational comparisons with CERES in Figure 13 are generally good and consistent with some of the in situ data, caution is warranted. The LWP (Figure 13a) is a heavily derived product from CERES and is subject to large retrieval uncertainties (e.g., Mace & Protat, 2018).

Figure 13c indicates less total cloudiness for *CAM5* than the other simulations. CERES EBAF 4.1 total cloud amounts for the same region and a 2 month average of January and February 2018 are shown on the figure and fall between *CAM5* and *CAM6* simulations. Total cloud area on a day-to-day basis is fairly well correlated between the CAM simulations and CERES (coefficient of 0.3 to 0.4). *CAM5* is slightly better correlated than *CAM6*. Cloud fraction is low in *CAM5*, while *CAM6* has too many, but the differences are small: ±5% at around 89% cloud cover.

CAM5 simulated cloud top drop size (Figure 13d) is notably smaller than *CAM6* and its variants, and corresponding to larger cloud drop number (Figure 13e). The CAM6 size distribution dual peaked shape is the same as SOCRATES data (Figure 12), but with reduced smaller particles, at the peak, resulting in a smaller effective size. The result of the smaller sizes, with less liquid and more ice, is reduced (less negative) cloud forcing over this 2 month period (Figure 13f). The CAM5 (Meyers et al., 1992) ice nucleation parameterization (*Meyers*) seems to be responsible for this, as it has results closer to *CAM5*. We also explored increasing
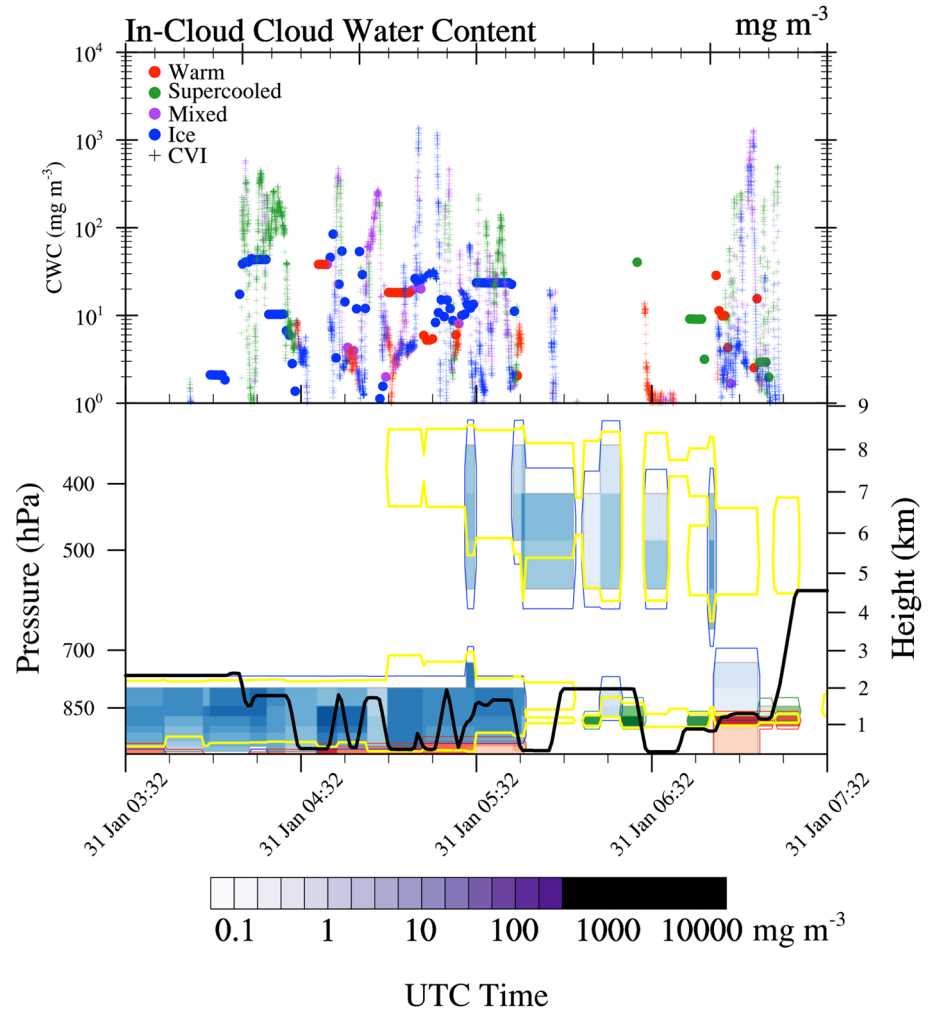
Figure 14. The same as Figure 10 but for a simulation using "CAM5" physical parameterizations.

ice nuclei for temperatures $T > -10°C$ (*In10-10*), and this increased IWP to even larger values than CAM5 (off scale on Figure 13b). CAM5 has the often seen model bias of too few and too bright S. Ocean clouds.

Interestingly, the less negative SW cloud forcing (radiative effect) is associated with lower cloud optical thickness, with CERES having a lower mean optical thickness than most of the CAM simulations. Note that CERES optical thickness is derived from infrared radiances on geostationary satellites and MODIS and also has assumptions in it. The *SB2001* simulation, with lower LWP and cloud optical depth, but higher cloud fraction and larger effective radius, as well as less ice water path (and significant supercooled liquid), seems to best reproduce the CERES observations during the SOCRATES period. The SW Cloud Radiative Effect (Figure 13) is similar to CERES with similar optical thickness but too much cloud cover.

With respect to some of the other sensitivity tests, it is notable that adjusting the SIP parameterization does not do much to the water path or number concentration, whether it is turned off (*SIP0*) or increased (*SIP5*). Note that the SIP parameterization in CAM represents the Hallett-Mossep process of rime splintering, which is only one of at least four possible mechanisms (Cotton et al., 1986). As noted, *Meyers* makes ice and liquid partitioning (and radiative effects) look more like *CAM5* and is a big reason for the difference between model versions over the S. Ocean. These results demonstrate that the radiative properties of S. Ocean clouds in CAM are sensitive to the ice nucleation scheme, similar to findings by Tan et al. (2016). This is discussed further in section 3.3 below. Changing autoconversion (*SB2001* and *Auto/10*) has large impacts on LWP and cloud radiative properties.
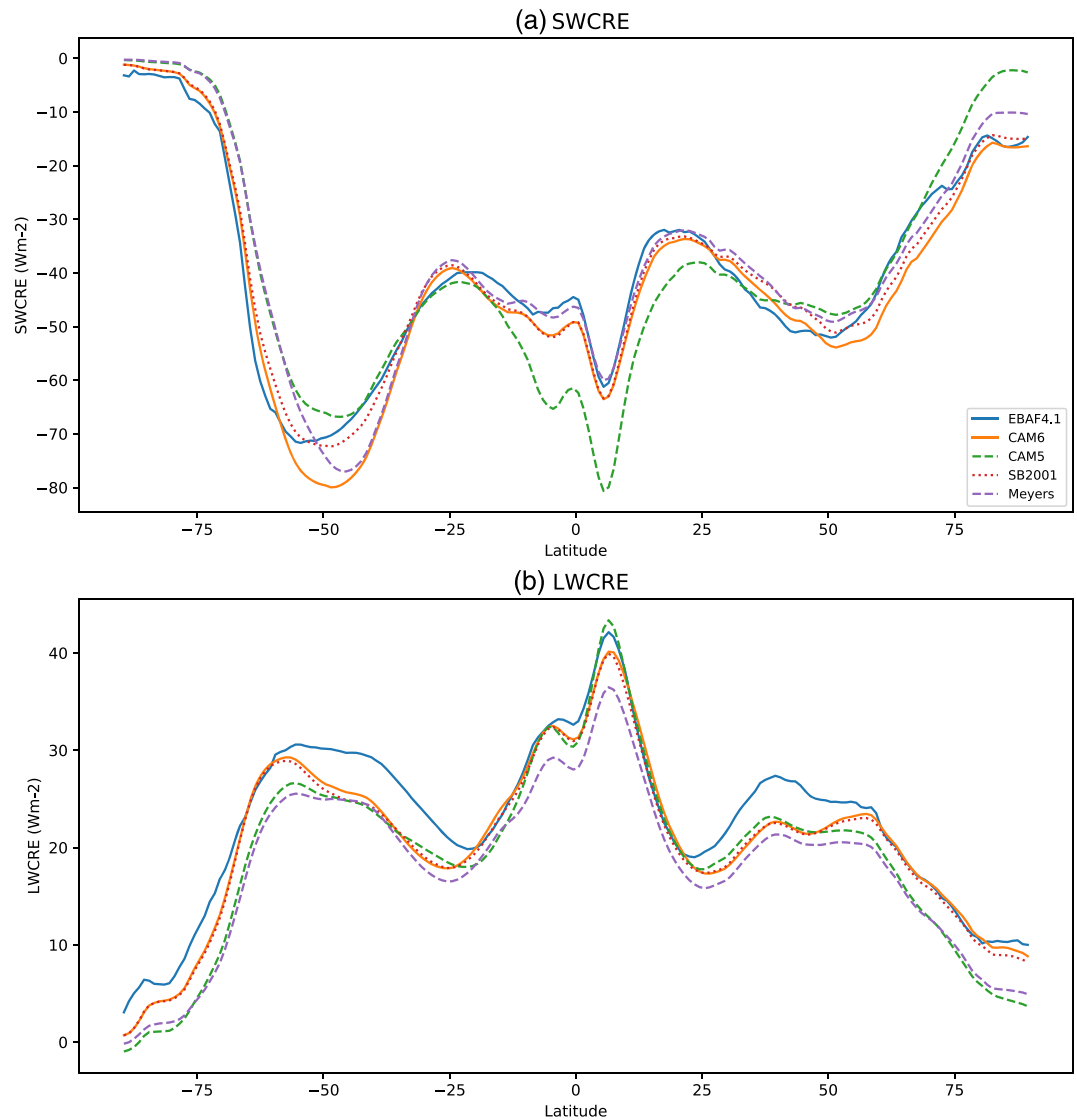
Figure 15. Zonal annual mean climatology of (a) SW and (b) LW Cloud radiative effects from CAM simulations and CERES observations (EBAF4.1).

Nudging has a nonnegligible impact on water and ice partitioning. *FixT* and *Nudge 1 hr* have less *T* bias but higher cloud water (Figure 13a) and stronger cloud forcing (Figure 13g). The free-running temperature simulation (*Free T*) has less cloudiness (Figure 13c), smaller sizes (Figure 13d), and reduced magnitude of SW Cloud Radiative Effect (Figure 13f). But the PBL structure has a larger bias in *Free T* (Figure 8).

As a more detailed illustration and comparison to SOCRATES observations, Figure 14 illustrates a simulation of flight RF07 with *CAM5* cloud microphysics, for comparison to Figure 10. *CAM5* features the Meyers et al. (1992) representation of ice nucleation as a function of temperature and diagnostic precipitation. Both these were found to reduce SLW (Gettelman et al., 2015). This does not match observations in the top panel of Figure 14, where the *CAM5* simulation has ice (blue) and some warm liquid (red), but almost none of the SLW (green) or mixed phase (purple) seen in the observations. This is clear indication that the revisions to cloud phase representation and partitioning in *CAM6* are an improvement over *CAM5* when compared to SOCRATES observations, even if the overall radiative effects in *CAM5* are closer to CERES (Figure 13f). The *SB2001* simulation has improved SW CRE (Figure 13f) but maintains supercooled liquid similar to CAM6 in Figure 10.

One additional note is that in CAM5 clouds are present all the way down to the lowest model layer ("strato-fogulus"), which was not observed during RF07 or other flights. The improvement in CAM6 is likely related to the new unified moist turbulence scheme (CLUBB), in which turbulence is driving cloud formation, in better agreement with observations.

3.3. Global Implications

Finally, we look at the longer term and global implications of these results. The different model formulations do not just have different results in the S. Ocean, but their climate is different globally. We have tested *CAM5*, *CAM6 Meyers*, and *SB2001* formulations. These simulations are detailed in Gettelman et al. (2019). Simulations are similar to the nudged runs (same code basis and same resolution) but run with climatological SSTs and no nudging. Simulations are 10 years long.

Figure 15 illustrates four different configurations of 10 yearlong free-running CAM simulations compared to a long-term annual climatology from 15 years of CERES EBAF 4.1 data. Here some of the results of Figure 13 can be put into context. In the S. Ocean, over all longitudes between 65°S and 45°S, *CAM5* has too weak SWCRE and LWCRE relative to CERES. CAM5 SW biases in the S. Ocean are due to low LWP. The SWCRE is too strong in *CAM6*, while the *SB2001* formulation is much closer to CERES observations. The Northern Hemisphere has lower magnitude SWCRE and biases in all formulations, likely due more land and less ocean, and a more constrained oceanic storm track. This is consistent with Figure 13 in the smaller SOCRATES region. However, the LWCRE has less bias in *CAM6* and *SB2001*, and the tropics are significantly better. CAM5 tropical SW biases are due to tuning of the deep convective parameterization over land.

LW biases actually are a consequence of different analysis methods. LWCRE from CERES is defined differently to CAM6: It is clear-sky-only points minus all sky. Whereas CAM6 is clear sky calculated everywhere (with more water vapor at points with high cloudiness), so there is less LW emission, a lower clear-sky value, and reduced difference between clear sky and all sky.

The SOCRATES region seasonal (DJF) SW root mean square error (RMSE) for all latitude-longitude grid points 65–45°S between the CAM simulations and CERES EBAF4.1 is larger for *CAM6* (24 W m^{-2}) than *CAM5* (9.7 W m^{-2}), but the Global Annual RMSE is smaller for *CAM6* (9.1 W m^{-2}) than *CAM5* (12.4 W m^{-2}) while CAM6 with Meyers et al. (1992) ice nucleation (*Meyers*) is intermediate between them. The use of Seifert and Beheng (2001) autoconversion (*SB2001*) yields lower RMSE versus CERES than CAM6 for the SOCRATES region seasonal DJF RMSE (16 W m^{-2}) and the lowest global RMSE (8.2 W m^{-2}).

The difference in mean state yields a different climate response. As noted by Gettelman et al. (2019), CAM5 and CAM6 have different climate sensitivity (the surface temperature response to an imposed forcing), which was found to be a result of different cloud feedbacks (the radiative response of clouds to surface warming). Gettelman et al. (2019) found that this difference was partially due to high-latitude cloud processes and the different distribution of SLW. As noted by Tan et al. (2016) and others, without supercooled liquid (CAM5), there is a negative cloud phase feedback when ice clouds become liquid in a warmer world. But if these clouds are supercooled liquid (CAM6), this negative feedback is not present.

4. Discussion

CAM6 nudged simulations are able to represent the locations and times SOCRATES observations of clouds and cloud microphysics. The nudging technique reproduces cloud regimes in similar locations to the aircraft, particularly with respect to supercooled liquid clouds. There are biases in the structure of the inversion at the top of boundary layer in the simulations in some situations, while an inversion is seen at aircraft altitudes in the simulations. Errors in the temperature inversion can be partially mitigated by fixing temperatures to the input data. Setting nudging timescales and parameters (whether to nudge temperature or not) will affect the cloud simulation, and while temperatures may move closer to observations (Figure 7), cloud simulation (cloud fraction, cloud phase, water content, and radiative effects) may change significantly (more than internal variability) and be further from CERES observations (Figure 13). Thus, there is an interplay between clouds and the dynamics of boundary layers in the S. Ocean that make it difficult for global models (either forecast models or climate models) to match the boundary layer structure and clouds simultaneously. CAM6 generates clouds from CCN with total water contents (CWC) not significantly different from observations, perhaps biased slightly high. Supercooled liquid in low clouds appears slightly low, with lower drop numbers, and fewer CCN than observed. CAM6 does this even with PBL biases, but eliminating

those biases and preventing feedbacks between clouds and temperatures (FixT experiment) degrades the cloud simulation.

Given these caveats about the method, the resulting cloud properties (total water and size distributions) agree with SOCRATES observations on individual flights and campaign averages given the uncertainty and variability of observations (which is large). Supercooled liquid clouds are produced extensively in cold sectors of cyclones in the S. Ocean targeted by SOCRATES. Supercooled liquid is better than in previous versions (CAM5) where large discrepancies have been found between CAM5 and observations (D'Alessandro et al., 2019; Wu et al., 2017; Zhang et al., 2019), and this is largely due to the new mixed phase ice nucleation, which is now dependent on available ice nuclei rather than an empirical function of temperature.

Previous work has touched on part of these issues in other regions. Xie et al. (2013) noted that the Meyers et al. (1992) empirical ice nucleation parameterization created biases. Xie et al. (2008) noted using Arctic observations that the mixed phase is important. In addition, models have errors in the boundary layer top and inversion strength, which will create errors in cloud position and bulk properties. But much of this work is macrophysical in nature focusing on bulk ice and radiation. The focus of analysis here is on detailed microphysical observations of cloud hydrometeors enabled by SOCRATES observations. Here we show this is also important in the S. Ocean, and a coherent picture of mass, number, and phase linked to radiative properties developed.

Cloud hydrometeor size distributions are also broadly reproduced across both ice and liquid from small cloud drops to large rain and snow particles. Note that because of size sorting across instruments, it is difficult to do direct number comparisons between particle counts and simulations. The model has some systematic deficiencies however. For warm clouds, there may be too much mass of rain, particularly around 100 micron diameter. For cold clouds, snow is well reproduced, but supercooled droplet size distributions tend to have too few numbers and an insufficient peak in the size distribution at 10–20 microns. Modification of the dispersion of the size distribution or increases in CCN improves these results slightly but does not increase overall drop numbers to match observation. Overall number is increased by reducing autoconversion or by increasing CCN. But decreasing LWP yields better comparisons to observed regional LWP, cloud optical depth, and SW cloud radiative effect. It seems as if higher CCN with reduced autoconversion would match observations better from in situ to regional and climate scale. This analysis thus provides a framework for how observations can inform model results and suggest changes: for example, to the warm rain formation process and perhaps to CCN and/or droplet activation.

Achieving radiative closure for cloud microphysics and radiation is thus difficult, since it requires accurately representing these moments (e.g., phase, mass, and hydrometeor number), even with observations. The *CAM5* simulated LWP over the entire region and period is 50% lower than *CAM6* (Figure 13a), and the IWP is 50% higher (Figure 13b). This likely leads to the lower cloud fraction (Figure 13c) and ultimately weaker SW CRE (Figure 13f) and lower optical depth (Figure 13g) simulated by *CAM5* compared to *CAM6*. Small (Figure 13d) and numerous (Figure 13e) drops compensate for low LWP, but since drop numbers are low, perhaps more CCN provide a better fit to observations. Meanwhile, in situ observations from SOCRATES suggest that the dominant cloud phase simulated in *CAM5* (ice) is far different from observed (supercooled liquid) and the cloud location (boundary layer top) also differs in *CAM5*. CERES also retrieves more ice than liquid, which does not match SOCRATES in situ observations. These comparisons call into question cloud products from the broader CERES observations in Figure 13. However, the *SB2001* experiment looks much closer to the CERES observations for LWP and SW CRE with less water, while maintaining significant SLW (similar to *CAM6*), demonstrating that multiple physical processes (activation, ice nucleation, and autoconversion) likely play an important role in how S. Ocean clouds are represented in CAM6.

The size distribution biases may contribute to the inability to reproduce the zonal mean structure of overall climatological cloud radiative effect, and having too few cloud drops may imply a larger mean size. However, the experiments with adjusted autoconversion indicate that lower water path (found in SB2001) can also improve the comparisons with observations. SB2001 is designed to have a hysteresis effect of being harder to initiate without rain and follows observations of rain formation v. LWP more closely in CAM (Gettelman, 2015). The mass seems to be the first-order effect, with size distributions a second-order effect. However, with larger drops it may be possible to maintain a larger liquid water path. Note that the CERES LWP product assumes a 10 micron size, so comparisons with in situ observations are perhaps more

relevant, which would indicate that increasing CCN in the region would improve the reproduction of observed microphysics from the in situ to regional and climate scale.

This analysis with observations provides a process (and observationally constrained) pathway to improve simulations further. Better constraints on condensate mass (and number) from the observations are still being developed for SOCRATES, and these will be valuable in adding an additional constraint on the simulations and resulting radiative properties. Because cloud feedbacks and climate sensitivity are dependent on the microphysics (phase and water content) of S. Ocean clouds, this is important for constraining climate projections. SOCRATES observations confirm that S. Ocean clouds are mostly supercooled liquid, similar to CAM6.

5. Conclusions

Nudged simulations with a global climate model (CAM6) even at coarse horizontal and vertical resolution are able to capture many of the important features of specific cloud systems observed by SOCRATES. Successful simulations have some biases in the boundary layer structure related to vertical resolution and to nudging itself, and some care must be taken in understanding the purpose of nudging as changing the temperature structure changes the overall cloud simulation. The fact that improving the temperatures relative to analysis temperatures may degrade the overall cloud simulation indicates problems fitting one model (CAM) to another model (MERRA2) state and/or compensating biases in CAM. It also indicates that there remain biases in the links between microphysics and turbulent structure. This is not surprising but perhaps provides a pathway for improving these links.

Comparisons between model and observations for flights into supercooled liquid clouds during SOCRATES show that improvements to the ice nucleation scheme in CAM6 result in significant improvements in the representation of SLW. CAM is not sensitive to Hallet-Mossep SIP process (the type represented in CAM) in the S. Ocean region. However, this is only one type of SIP and further work is being conducted on specific SIP cases to investigate other mechanisms more completely from observations and model extensions. CAM6 is sensitive to ice nucleation and changes in warm rain formation (autoconversion). Simulated CCN concentrations in shallow clouds may be low, contributing to low drop numbers.

One of the most unique features of this study is the ability to compare detailed cloud microphysics (phase and size distributions of different hydrometeors) across scales between large-scale models and in situ observations. This works particularly well in the relatively uniform cloud regimes observed during SOCRATES RF07 and other SOCRATES flights. The method and comparisons also allow identification of what specific processes lead to biases in microphysics and even the large-scale radiative effect of clouds over the S. Ocean.

Closure between microphysics and radiation is difficult. While the overall microphysics and phase of clouds in CAM6 look quite good for SOCRATES clouds, when a broader climatological picture is explored over the SOCRATES region, there are significant biases in radiative fluxes. The details of the cloud physics might be creating biases such that the right radiative response is occurring for the wrong reasons in either the model or satellite retrievals. Biases in PBL structure hint that the macrophysical environment and humidity structure feeding the microphysics may also be biased. The radiative response can be improved with less water path through the use of a revised autoconversion scheme (*SB2001*) but still does not match droplet numbers seen in the aircraft observations, implying higher CCN and drop number may be necessary. The biases in rain size distributions also point at a failure to simulate the warm rain process effectively.

Observations are also not consistent between satellites and in situ data. It is likely that the CERES retrievals of microphysics (LWP and IWP) from radiative fluxes and radiances have significant biases due to fixed specification of particle size. This makes comparisons with satellite retrievals from the top of the atmosphere difficult to compare, not least because of uncertainty in the satellite retrievals themselves, which is a useful subject for further study against SOCRATES data. Note that the satellite comparisons of mean radiative effects involves all clouds, not just low clouds. It argues for more detailed analysis of observations themselves to better extrapolate from the aircraft data to the regional and long-term coverage of satellite data. It also confirms the focus here not on a total number concentration or an effective radius but on a size distribution as well as CCN, which is more appropriate for in situ comparison as well as providing more insight into the processes as noted above.

Because model formulations with different cloud microphysics (i.e., CAM6 and CAM5) have different high latitude cloud feedbacks, it is critical to understand and constrain the phase partitioning and cloud microphysics of S. Ocean clouds. In this case, CAM6 with more supercooled liquid and more positive cloud feedbacks (and higher climate sensitivity) looks more physically plausible in the S. Ocean due to better cloud phase simulation.

These results should be tested against different scales of cloud models for the SOCRATES regime, and against different global simulations. In addition, better constraints on in situ observed condensate mass and phase would be useful for better constraining the observations. There are still large uncertainties in the retrieval of condensate mass and discrimination into liquid and ice from the in situ cloud probes and is thus the focus of a separate manuscript. Detailed analysis of activation, ice nucleation, and CCN is also the focus of future work.

In particular, advanced two-moment cloud physics schemes such as Gettelman and Morrison (2015) provide more detail about potential causes for discrepancies against observations, and a multiscale observational approach from in situ microphysics to satellite data provides unprecedented detail that has and can continue to help guide model improvements in this important region for climate projections.

Data Availability Statement

All observation data are available from the NCAR Earth Observation Lab (EOL) Field Catalog (http://catalog.eol.ucar.edu/socrates) (and specifically from https://data.eol.ucar.edu/master_lists/generated/socrates/). Model simulations are available online (at https://doi.org/10.5065/qzke-cp31).

References

Abdul-Razzak, H., & Ghan, S. J. (2002). A parameterization of aerosol activation 3. Sectional representation. *Journal of Geophysical Research*, *107*(D3), 4026. https://doi.org/10.1029/2001JD000483

Baumgardner, D., Abel, S. J., Axisa, D., Cotton, R., Crosier, J., Field, P., et al. (2017). Cloud ice properties: In situ measurement challenges. *Meteorological Monographs*, *58*, 9.1–9.23. https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0011.1

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., & Elsaesser, G. S. (2019). Strong dependence of atmospheric feedbacks on mixed-phase microphysics and aerosol-cloud interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*, 1735–1758. https://doi.org/10.1029/2019MS001688

Bodas-Salcedo, A., Williams, K., Field, P., & Lock, A. (2012). The surface downwelling solar radiation surplus over the Southern Ocean in the Met Office model: The role of midlatitude cyclone clouds. *Journal of Climate*, *25*(21), 7467–7486.

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., & Schanen, D. P. (2013). Higher-order turbulence closure and its impact on climate simulation in the Community Atmosphere Model. *Journal of Climate*, *26*(23), 9655–9676. https://doi.org/10.1175/JCLI-D-13-00075.1

Bretherton, C. S., & Park, S. (2009). A new moist turbulence parameterization in the Community Atmosphere Model. *Journal of Climate*, *22*, 3422–3448.

Cotton, W. R., Tripoli, G. J., Rauber, R. M., & Mulvihill, E. A. (1986). Numerical simulation of the effects of varying ice crystal nucleation rates and aggregation processes on orographic snowfall. *Journal of Applied Meteorology and Climatology*, *25*, 1658–1680.

D'Alessandro, J. J., Diao, M., Wu, C., Liu, X., Jensen, J. B., & Stephens, B. B. (2019). Cloud phase and relative humidity distributions over the Southern Ocean in austral summer based on in situ observations and CAM5 simulations. *Journal of Climate*, *32*(10), 2781–2805. https://doi.org/10.1175/JCLI-D-18-0232.1

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The Community Earth System Model Version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001916. https://doi.org/10.1029/2019MS001916

EOL (2018). SOCRATES 2018 Project Manager Report. NCAR - Earth Observing Laboratory.

Gettelman, A. (2015). Putting the clouds back in aerosol-cloud interactions. *Atmospheric Chemistry and Physics*, *15*(21), 12,397–12,411. https://doi.org/10.5194/acp-15-12397-2015

Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A. G., Danabasoglu, G., et al. (2019). High climate sensitivity in the Community Earth System Model Version 2 (CESM2). *Geophysical Research Letters*, *46*, 8329–8337. https://doi.org/10.1029/2019GL083978

Gettelman, A., Liu, X., Ghan, S. J., Morrison, H., Park, S., Conley, A. J., et al. (2010). Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the Community Atmosphere Model. *Journal of Geophysical Research*, *115*, D18216. https://doi.org/10.1029/2009JD013797

Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes. *Journal of Climate*, *28*(3), 1268–1287. https://doi.org/10.1175/JCLI-D-14-00102.1

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. M. (2015). Advanced two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol-cloud interactions. *Journal of Climate*, *28*(3), 1288–1307. https://doi.org/10.1175/JCLI-D-14-00103.1

Golaz, J. C., Larson, V. E., & Cotton, W. R. (2002). A PDF-based model for boundary layer clouds. Part II: Model results. *Journal of the Atmospheric Sciences*, *59*, 3552–3571.

Hoose, C., Kristjánsson, J. E., Chen, J. P., & Hazra, A. (2010). A classical-theory-based parameterization of heterogeneous ice nucleation by mineral dust, soot, and biological particles in a Global Climate Model. *Journal of the Atmospheric Sciences*, *67*(8), 2483–2503. https://doi.org/10.1175/2010JAS3425.1

Iacono, M. J., Mlawer, E. J., Clough, S. A., & Morcrette, J. J. (2000). Impact of an improved longwave radiation model, RRTM, on the energy budget and thermodynamic properties of the NCAR Community Climate Model, CCM3. *Journal of Geophysical Research*, *105*(D11), 14,873–14,890.

Khairoutdinov, M. F., & Kogan, Y. (2000). A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Monthly Weather Review*, *128*, 229–243.

King, W. D., Parkin, D. A., & Handsworth, R. J. (1978). A hot-wire liquid water device having fully calculable response characteristics. *Journal of Applied Meteorology and Climatology*, *17*(12), 1809–1813. https://doi.org/10.1175/1520-0450(1978)017<1809:AHWLWD>2.0.CO;2

Lance, S., Brock, C. A., Rogers, D., & Gordon, J. A. (2010). Water droplet calibration of the Cloud Droplet Probe (CDP) and in-flight performance in liquid, ice and mixed-phase clouds during ARCPAC. *Atmospheric Measurement Techniques*, *3*(6), 1683–1706. https://doi.org/10.5194/amt-3-1683-2010

Larson, V. E., Golaz, J. C., & Cotton, W. R. (2002). Small-scale and mesoscale variability in cloudy boundary layers: Joint probability density functions. *Journal of the Atmospheric Sciences*, *59*(24), 3519–3539. https://doi.org/10.1175/1520-0469(2002)059<3519:SSAMVI>2.0.CO;2

Liu, X., Ma, P. L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., et al. (2016). Description and evaluation of a new four-mode version of the Modal Aerosol Module (MAM4) within Version 5.3 of the Community Atmosphere Model. *Geoscientific Model Development*, *9*(2), 505–522. https://doi.org/10.5194/gmd-9-505-2016

Liu, X., & Penner, J. E. (2005). Ice nucleation parameterization for global models. *Meteor Z*, *14*, 499–514.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018). Clouds and the Earth's Radiant Energy System (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) Edition-4.0 data product. *Journal of Climate*, *31*(2), 895–918. https://doi.org/10.1175/JCLI-D-17-0208.1

Lohmann, U., & Neubauer, D. (2018). The importance of mixed-phase and ice clouds for climate sensitivity in the global aerosol-climate model ECHAM6-HAM2. *Atmospheric Chemistry and Physics*, *18*(12), 8807–8828. https://doi.org/10.5194/acp-18-8807-2018

Mace, G. G., & Protat, A. (2018). Clouds over the Southern Ocean as observed from the R/V investigator during CAPRICORN. Part I: Cloud occurrence and phase partitioning. *Journal of Applied Meteorology and Climatology*, *57*(8), 1783–1803. https://doi.org/10.1175/JAMC-D-17-0194.1

Martin, G. M., Johnson, D. W., & Spice, A. (1994). The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *Journal of the Atmospheric Sciences*, *51*, 1823–1842.

McCluskey, C. S., Hill, T. C. J., Humphries, R. S., Rauker, A. M., Moreau, S., Strutton, P. G., et al. (2018). Observations of ice nucleating particles over Southern Ocean waters. *Geophysical Research Letters*, *45*, 11,989–11,997. https://doi.org/10.1029/2018GL079981

McFarquhar, G. M., Baumgardner, D., Bansemer, A., Abel, S. J., Crosier, J., French, J., et al. (2017). Processing of ice cloud in situ data collected by bulk water, scattering, and imaging probes: Fundamentals, uncertainties, and efforts toward consistency. *Meteorological Monographs*, *58*, 11.1–11.33. https://doi.org/10.1175/AMSMONOGRAPHS-D-16-0007.1

Meyers, M. P., DeMott, P. J., & Cotton, W. R. (1992). New primary ice-nucleation parameterizations in an explicit cloud model. *Journal of Applied Meteorology*, *31*, 708–721.

Molod, A., Takacs, L., Suarez, M., & Bacmeister, J. (2015). Development of the GEOS-5 atmospheric general circulation model: Evolution from MERRA to MERRA2. *Geoscientific Model Development*, *8*(5), 1339–1356. https://doi.org/10.5194/gmd-8-1339-2015

Morrison, H., & Gettelman, A. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the NCAR Community Atmosphere Model (CAM3), Part I: Description and numerical tests. *Journal of Climate*, *21*(15), 3642–3659.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., et al. (2010). Description of the NCAR Community Atmosphere Model (CAM5.0) (NCAR/TN-486+STR). Boulder, CO, USA: National Center for Atmospheric Research.

Neale, R. B., Richter, J. H., & Jochum, M. (2008). The impact of convection on ENSO: From a delayed oscillator to a series of events. *Journal of Climate*, *21*, 5904–5924. https://doi.org/10.1175/2008JCLI2244.1

O'Shea, S. J., Choularton, T. W., Flynn, M., Bower, K. N., Gallagher, M., Crosier, J., et al. (2017). In situ measurements of cloud microphysics and aerosol over coastal Antarctica during the MAC campaign. *Atmospheric Chemistry and Physics*, *17*(21), 13,049–13,070. https://doi.org/10.5194/acp-17-13049-2017

Park, S., & Bretherton, C. S. (2009). The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the Community Atmosphere Model. *Journal of Climate*, *22*, 3449–3469.

Roberts, G. C., & Nenes, A. (2005). A continuous-flow streamwise thermal-gradient CCN chamber for atmospheric measurements. *Aerosol Science and Technology*, *39*(3), 206–221. https://doi.org/10.1080/027868290913988

Rotstayn, L. D., & Liu, Y. (2003). Sensitivity of the first indirect aerosol effect to an increase of cloud droplet spectral dispersion with droplet number concentration. *Journal of Climate*, *16*(21), 3476–3481. https://doi.org/10.1175/1520-0442(2003)016<3476:SOTFIA>2.0.CO;2

Seifert, A., & Beheng, K. D. (2001). A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmospheric Research*, *59–60*, 265–281.

Shi, X., Liu, X., & Zhang, K. (2015). Effects of pre-existing ice crystals on cirrus clouds and comparison between different ice nucleation parameterizations with the Community Atmosphere Model (CAM5). *Atmospheric Chemistry and Physics*, *15*(3), 1503–1520. https://doi.org/10.5194/acp-15-1503-2015

Tan, I., Storelvmo, T., & Zelinka, M. D. (2016). Observational constraints on mixed-phase clouds imply higher climate sensitivity. *Science*, *352*(6282), 224–227. https://doi.org/10.1126/science.aad5300

Trenberth, K. E., & Fasullo, J. T. (2010). Simulation of present-day and twenty-first-century energy budgets of the Southern Oceans. *Journal of Climate*, *23*, 440–454. https://doi.org/10.1175/2009JCLI3152.1

Tsushima, Y., Emori, S., Ogura, T., Kimoto, M., Webb, M. J., Williams, K. D., et al. (2006). Importance of the mixed-phase cloud distribution in the control climate for assessing the response of clouds to carbon dioxide increase: A multi-model study. *Climate Dynamics*, *27*, 113–126. https://doi.org/10.1007/s00382-006-0127-7

Twohy, C. H., Schanot, A. J., & Cooper, W. A. (1997). Measurement of condensed water content in liquid and ice clouds using an airborne counterflow virtual impactor. *Journal of Atmospheric and Oceanic Technology*, *14*(1), 197–202. https://doi.org/10.1175/1520-0426(1997)014<0197:MOCWCI>2.0.CO;2

Vergara-Temprado, J., Miltenberger, A. K., Furtado, K., Grosvenor, D. P., Shipway, B. J., Hill, A. A., et al. (2018). Strong control of Southern Ocean cloud reflectivity by ice-nucleating particles. *Proceedings of the National Academy of Sciences of the United States of America*, *115*, 201721627. https://doi.org/10.1073/pnas.1721627115

Wang, Y., Liu, X., Hoose, C., & Wang, B. (2014). Different contact angle distributions for heterogeneous ice nucleation in the Community Atmospheric Model version 5. *Atmospheric Chemistry and Physics*, *14*(19), 10,411–10,430. https://doi.org/10.5194/acp-14-10411-2014

Wang, Y., McFarquhar, G. M., Rauber, R. M., Zhao, C., Wu, W., Finlon, J. A., et al. (2020). Microphysical properties of generating cells over the Southern Ocean: Results from SOCRATES. *Journal of Geophysical Research: Atmospheres*, *125*, e2019JD032237. https://doi.org/10.1029/2019JD032237

Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Smith, G. L., & Cooper, J. E. (1996). Clouds and the Earth's Radiant Energy System (CERES): An Earth observing system experiment. *Bulletin of the American Meteorological Society*, *77*(5), 853–868.

Wu, C., Liu, X., Diao, M., Zhang, K., Gettelman, A., Lu, Z., et al. (2017). Direct comparisons of ice cloud macro- and microphysical properties simulated by the Community Atmosphere Model Version 5 with HIPPO aircraft observations. *Atmospheric Chemistry and Physics*, *2017*, 1–57. https://doi.org/10.5194/acp-2016-1106

Xie, S., Boyle, J., Klein, S. A., Liu, X., & Ghan, S. (2008). Simulations of Arctic mixed-phase clouds in forecasts with CAM3 and AM2 for M-PACE. *Journal of Geophysical Research*, *113*, D04211. https://doi.org/10.1029/2007JD009225

Xie, S., Liu, X., Zhao, C., & Zhang, Y. (2013). Sensitivity of CAM5-simulated arctic clouds and radiation to ice nucleation parameterization. *Journal of Climate*, *26*(16), 5981–5999. https://doi.org/10.1175/JCLI-D-12-00517.1

Young, K. (2018). NCAR/EOL quality controlled dropsonde data. Version 1.0: NCAR - Earth Observing Laboratory. https://doi.org/10.5065/D6QZ28SG

Young, K., & Vömel, H. (2018). NCAR/EOL Quality Controlled Dropsonde Data. Version 1.0. UCAR/NCAR - Earth Observing Laboratory. https://doi.org/10.5065/D6QZ28SG

Zhang, M., Liu, X., Diao, M., D'Alessandro, J. J., Wang, Y., Wu, C., et al. (2019). Impacts of representing heterogeneous distribution of cloud liquid and ice on phase partitioning of Arctic mixed-phase clouds with NCAR CAM5. *Journal of Geophysical Research: Atmospheres*, *125*, 13,071–13,090. https://doi.org/10.1029/2019JD030502

Zhang, G. J., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Center General Circulation Model. *Atmosphere Ocean*, *33*, 407–446.

Zondlo, M. A., Paige, M. E., Massick, S. M., & Silver, J. A. (2010). Vertical cavity laser hygrometer for the National Science Foundation Gulfstream-V Aircraft. *Journal of Geophysical Research*, *115*, D20309. https://doi.org/10.1029/2010JD014445