Southern Ocean Solar Reflection Biases in CMIP6 Models Linked to Cloud Phase and Vertical Structure Representations

Grégory V. Cesana, Théodore Khadir, Hélène Chepfer, Marjolaine Chiriaco

Geophysical Research Letters

RESEARCH LETTER
10.1029/2022GL099777

Key Points:

- The lack of shortwave reflection is somewhat corrected in Coupled Model Intercomparison Project Phase 6 (CMIP6) models, attributable to a better representation of cloud fraction and cloud phase
- The cloud fraction and radiative effect behavior is different north and south of 55°S, which is where the 0°C isotherm meets the surface
- Contrary to observations, CMIP6 clouds are less reflective south of 55°S, where boundary layer clouds are often topped by mid-level clouds

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

G. V. Cesana,
gregory.cesana@columbia.edu

Citation:

Cesana, G. V., Khadir, T., Chepfer, H., & Chiriaco, M. (2022). Southern ocean solar reflection biases in CMIP6 models linked to cloud phase and vertical structure representations. *Geophysical Research Letters*, *49*, e2022GL099777. https://doi.org/10.1029/2022GL099777

Received 2 JUN 2022
Accepted 1 NOV 2022

Author Contributions:

Conceptualization: Grégory V. Cesana, Théodore Khadir, Hélène Chepfer, Marjolaine Chiriaco
Formal analysis: Grégory V. Cesana, Théodore Khadir
Methodology: Grégory V. Cesana, Théodore Khadir, Hélène Chepfer, Marjolaine Chiriaco
Writing – original draft: Grégory V. Cesana
Writing – review & editing: Théodore Khadir, Hélène Chepfer, Marjolaine Chiriaco

Southern Ocean Solar Reflection Biases in CMIP6 Models Linked to Cloud Phase and Vertical Structure Representations

Grégory V. Cesana[1,2], Théodore Khadir[3], Hélène Chepfer[4], and Marjolaine Chiriaco[5]

[1]Center for Climate Systems Research, Columbia University, New York, NY, USA, [2]NASA Goddard Institute for Space Studies, New York, NY, USA, [3]Department of Environmental Science and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden, [4]LMD/IPSL, CNRS, Sorbonne Université, École Polytechnique, Paris, France, [5]LATMOS/IPSL, University Versailles Saint-Quentin en Yvelines, Guyancourt, France

Abstract Over the Southern Ocean (SO, 40°S–70°S), climate models have consistently underestimated solar reflection. Here we evaluate the relationship between cloud profiles, cloud phase and radiation over the SO in Coupled Model Intercomparison Project Phase 6 (CMIP6) models against Clouds and the Earth's Radiant Energy System and Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations. We find that the lack of solar reflection is slightly improved in CMIP6 models compared to CMIP5's, attributable to a better representation of cloud fraction and phase. We show that clouds have a different vertical structure and radiative effect south and north of where the 0°C isotherm meets the surface (~55°S). Although the models capture the greater vertical extent of clouds south of 55°S, they fail to reproduce the observed increase in solar reflection, which we pinpoint to cloud phase biases. Increasing CMIP6 supercooled liquid cloud opacity should help reduce their persistent shortwave biases.

Plain Language Summary Over the Southern Ocean, defined as the latitudinal band between 40° and 70°S, climate models have consistently overestimated the amount of absorbed solar radiation, mostly because of biases in the representation of clouds. Such biases in climate models are particularly problematic because they may affect the radiative response of clouds to climate warming. We find that the newest generation of climate models better represents clouds than the previous one, compared to satellite observations. We show that clouds have a different vertical structure and radiative properties south and north of the 0°C surface isotherm, around 55°S. Although the models capture the greater vertical extent of clouds south of 55°S, they fail to reproduce the observed increase in solar reflection by clouds there. Finally, we report that increasing the opacity of liquid clouds at subzero temperatures (between 0° and −40°C) should help reduce persistent solar radiation biases attributable to clouds, south of 55°S, in the newest generation of climate models.

1. Introduction

Earth's climate sensitivity has increased in the most recent Coupled Model Intercomparison Project (CMIP) models, largely attributed to the response of clouds to climate change, referred to as cloud feedback, in the extratropics (Zelinka et al., 2020). In particular, over the Southern Ocean (SO; defined as being between 40° and 70°S), the negative shortwave (SW) cloud feedback southward of about 55°S, has weakened compared to the previous CMIP exercise, while the positive SW cloud feedback northward of 55°S has strengthened, making the overall feedback more positive in models (Zelinka et al., 2020). Around this latitude, the boundary layer (BL) typically transitions from positive temperatures on the northern side (between 40° and 55°S) to negative temperatures on the southern side (between 55° and 70°S), whereby climate model radiation schemes are the most sensitive to cloud phase transitions with substantial consequences for cloud feedbacks (Bodas-Salcedo et al., 2019; Frey & Kay, 2018; Gettelman et al., 2019; Tan et al., 2016).

For example, increasing the amount of supercooled liquid clouds in a model (Kay et al., 2016), which are typically more opaque than their frozen counterparts (Cesana & Storelvmo, 2017; Tsushima et al., 2006), can help fix the long-standing lack of solar radiation reflection over the SO (Trenberth & Fasullo, 2010). To address this problem and more generally the lack of supercooled liquid clouds reported in many Earth System Models (ESMs) (Cesana et al., 2012, 2015; Choi et al., 2010; Komurcu et al., 2014; McCoy et al., 2015), some modeling centers have increased the relative part of liquid clouds in their CMIP6 models (Bodas-Salcedo et al., 2019; Gettelman et al., 2019; Kelley et al., 2020; Madeleine et al., 2020). It is therefore useful to determine whether ESMs properly

simulate cloud phase transitions in the present-day climate—the focus of this analysis—besides reproducing the response of cloud properties to climate change.

Here we leverage the recent submission of simulated cloud phase profiles using a lidar simulator to CMIP6 to analyze and evaluate the relationship between cloud phase and radiation over the SO. The use of a lidar simulator (Cesana & Chepfer, 2013), which mimics what a satellite-borne lidar would observe over an ESM atmosphere, offers the possibility to collectively assess CMIP6 models in a consistent manner against observations. We first evaluate how CMIP cloud fraction profiles compare with Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) observations on a global scale. We then study the temperature dependence of the cloud phase in ESMs over the SO. And finally, we analyze how biases in liquid cloud properties may affect top-of-the-atmosphere (TOA) SW fluxes over the SO.

2. Data and Methods

2.1. Observations

We use the General Circulation Model-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP) observations (v2.9, Cesana & Chepfer, 2013; Chepfer et al., 2010), which document cloud properties every 333 m along-track-resolution near-nadir lidar profiles for 480 m height intervals. Liquid- and ice-dominated cloud are distinguished using the state of lidar beam polarization, which changes when backscattered by a non-spherical crystal contrary to spherical droplets (Text S1 in Supporting Information S1). We use the Earth's Radiant Energy System-Energy Balanced and Filled (CERES-EBAF Ed4.1; Loeb et al., 2018) to quantify the SW cloud radiative effect at TOA (referred to as $SWCRE_{TOA}$ in the remainder of the manuscript), defined as the difference between clear-sky and all-sky monthly upward fluxes at TOA. The more negative the $SWCRE_{TOA}$ is, the brighter the cloud is. We also use monthly mean temperature profiles from a mean of three reanalyzes listed in the open research statement (Hersbach et al., 2020; Kanamitsu et al., 2002; Slivinski et al., 2019).

2.2. Model Simulations

We analyze monthly outputs from global simulations with prescribed sea surface temperatures (SST; following the Atmospheric Model Intercomparison Project (AMIP)) from nine CMIP6 models (Table S1 in Supporting Information S1). They use various methods to predict liquid and ice partitioning in clouds from a simple temperature-based dependence (Madeleine et al., 2020; Roehrig et al., 2020) to more sophisticated microphysical schemes (Danabasoglu et al., 2020; Held et al., 2019). In addition, we include the latest version of the Goddard Institute for Space Studies (GISS) ESM (Cesana et al., 2021), referred to as GISS-ModelE3. To ensure a fair evaluation that accounts for the CALIPSO lidar limitations and uses similar cloud and cloud phase definitions, we use the CALIPSO-like outputs from the models, obtained with the CALIPSO lidar simulator (Cesana & Chepfer, 2013; Chepfer et al., 2008), and interpolated on CALIPSO-GOCCP's grid.

3. Results

3.1. Vertical Structure of Clouds

We first analyze zonal profiles of cloud fraction and $SWCRE_{TOA}$ to roughly evaluate the impact of cloud biases on TOA SW radiation in CMIP6 models, during the Austral summer (December January February, Figure 1). On a global scale, the CMIP6 models exhibit qualitatively similar biases than the previous CMIP generation (Figure 1; Cesana & Chepfer, 2012; Cesana & Waliser, 2016), that is, too many clouds in the high levels ($z > 6.5$ km) and too few in the low levels ($z < 3.5$ km) except below 1 km. However, a closer inspection reveals that CMIP6 models simulate more high clouds in the tropics, and substantially more low- and mid-level clouds at midlatitudes and in the polar regions than CMIP5 models (Figure 1d). This increase in cloud fraction can arguably be attributed to modeling center efforts to address long-standing and systematic biases pointed out in the literature, particularly over the SO, where the lack of low-level clouds in CMIP5 models have been linked with TOA SW radiation biases (Wang & Su, 2013; Zhang et al., 2005).

In the southern hemisphere, as the cloud regime evolves from subtropical boundary-layer clouds (shallow cumuli and stratocumuli) to more vertically extended midlatitude-storm clouds (frontal clouds), the observations show

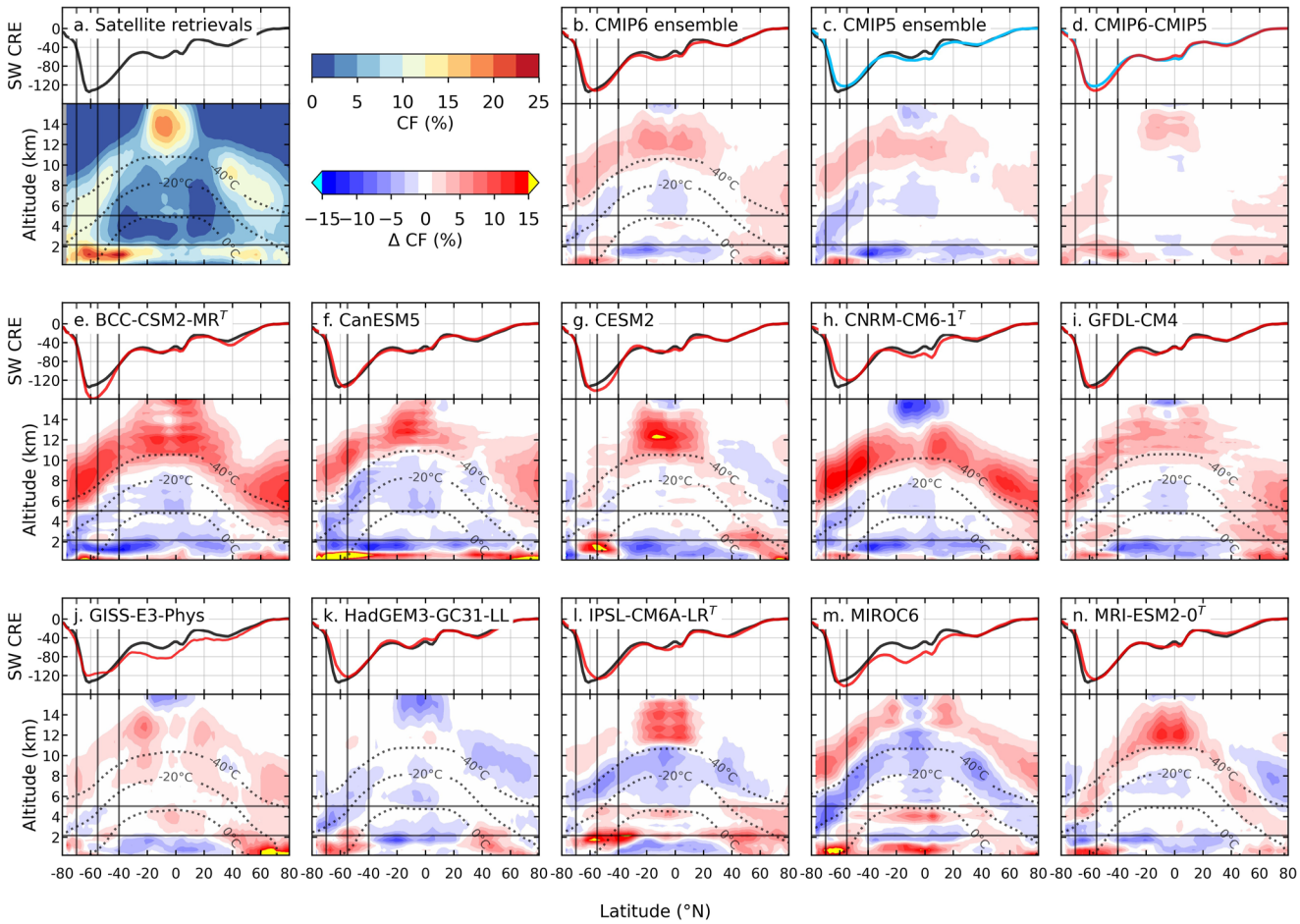

Figure 1. Relationship between cloud fraction profiles and $SWCRE_{TOA}$. During the austral summer zonal profiles of cloud fraction (%) for General Circulation Model-Oriented CALIPSO Cloud Product (2008–2015) and cloud fraction biases (%) for the Coupled Model Intercomparison Project (CMIP) lidar simulator outputs, and zonal means of $SWCRE_{TOA}$ for Clouds and the Earth's Radiant Energy System (2008–2015, black line) and for the CMIP outputs (red line, W/m^2) in the upper part of each subplot. The 0, −20, and −40°C isotherms (dotted black lines) help locate mixed-phase temperatures. The Southern Ocean and low- and mid-levels are emphasized by horizontal black lines and the vertical black lines mark the 40°S, 70°S, and 55°S where the 0°C isotherm meets the surface.

that the $SWCRE_{TOA}$ nearly doubles and peaks near 60°S, where the low- and the mid-level cloud fractions are both large. Although the CMIP6 models reproduce that transition from low-level to mid-level topped clouds rather well, substantial systematic biases persist in both their cloud fraction and $SWCRE_{TOA}$. In light of this clear distinction between these cloud regimes, we separate the SO into two regions: between 40° and 55°S, where boundary-layer temperatures are mostly positive and low-level clouds prevail, and between 55° and 70°S, where boundary-layer temperatures drop below freezing level and more mid-level clouds form.

Between 40°S and 55°S, the CMIP5 models underestimate the cloud fraction between 1 and 3 km, which helps explain their long-standing lack of cloud reflection compared to observations. By contrast, most CMIP6 models, as exemplified by the multimodel mean, overestimate the cloud fraction below 3 km. In this region, BL temperatures are mostly positive and therefore clouds are composed of liquid droplets that strongly reflect SW radiation. Consequently, the addition of BL clouds from CMIP5 to CMIP6 models has increased their amount of solar reflection, which is now slightly overestimated compared to observations. A closer look at the frequency of occurrence low-cloud cover (i.e., as seen from space) also shows that CMIP6 models are overall in better agreement with the observations than CMIP5 models, which produce too many occurrences of small low-cloud cover, except for low-cloud covers greater than 60%, where CMIP6 models simulate slightly too reflective clouds (Figure S1 in Supporting Information S1; Nam et al., 2012; Zhang et al., 2005).

Poleward of 55°S, where BL temperatures drop below 0°C and more clouds vertically extend into the free troposphere, the long-standing lack of solar reflection identified in CMIP5 models has been somewhat corrected in

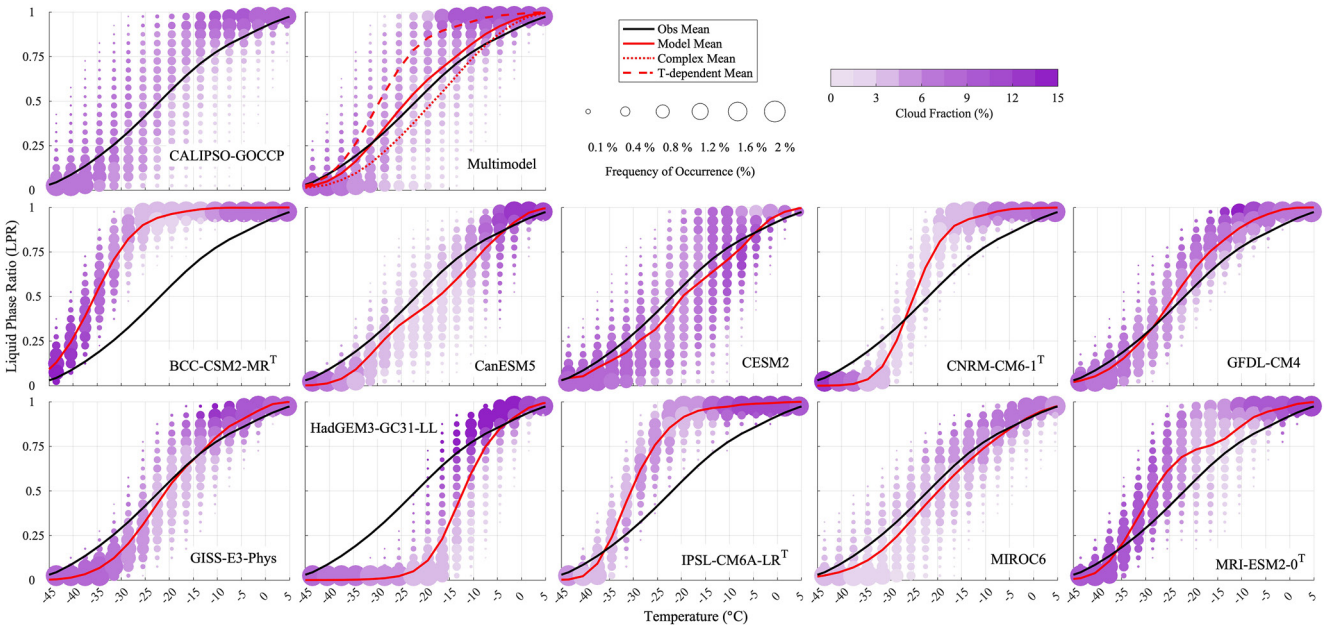

Figure 2. Relation between liquid phase ratio (LPR) and temperature over the Southern Ocean (70°S–40°S) during the austral summer. Frequency of occurrence (circles) of the LPR as a function of the temperature for the Coupled Model Intercomparison Project Phase 6 lidar simulator outputs (red lines) and for General Circulation Model-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP) (black line; 2007–2016). The size of the circles represents the frequency of occurrence of a specific LPR between −91°C and 24°C—which corresponds to the range of temperatures used in CALIPSO-GOCCP temperature profiles of cloud fraction—while its shading corresponds to the mean cloud fraction in each 3°C and 10% bin. The models marked with a T rely mainly on temperature-dependent relationships to partition liquid and ice hydrometeors (Table S1 in Supporting Information S1).

the CMIP6 multimodel mean (Figures 1b and 1c). This improvement is attributable to larger multimodel mean cloud fractions in both low and mid-levels (Figure 1d), in better agreement with observations, yet all CMIP6 models but Beijing Climate Centre (BCC) underestimate the magnitude of the $SWCRE_{TOA}$ (Figures 1e–1n). Both CMIP5 and CMIP6 clouds are not reflective enough for a given low-cloud cover, although CMIP6 models have substantially improved their representation of low-cloud cover occurrences, that is, more frequent large low-cloud cover (Figure S1 in Supporting Information S1). We note that although all models but HadGEM3-GC31-LL overestimate the cloud fraction above 6 km over the SO (Figure 1), we do not expect this to have a substantial impact on TOA SW radiation since these clouds are typically optically thin ice clouds (optical thickness smaller than 3) and have a relatively small $SWCRE_{TOA}$ (Matus & L'Ecuyer, 2017). The cloud fraction underestimate between 1 and 5 km might appear to be the primary reason for the lack of SW reflection. Yet, these altitudes typically correspond to temperature between 0° and −20°C where unrealistic or overly simplistic representation of cloud phase can also cause cloud and radiation biases, which is why we further explore how the cloud phase is simulated in CMIP6 models in the next section.

3.2. Relationship Between Cloud Phase and Temperature

Historically, many GCMs have been pointed out to underestimate the amount of supercooled clouds, causing non-negligible radiative biases both at TOA (Bodas-Salcedo et al., 2016; McCoy et al., 2015) and at the surface (Cesana et al., 2012; Stramler et al., 2011). We therefore analyze the cloud phase representation of CMIP6 models against CALIPSO-GOCCP to help us better understand the origins of the systematic SW radiation bias. We use the so-called liquid phase ratio (LPR), which describes the ratio of liquid cloud frequency with respect to all-cloud frequency, ranging from 0 (pure ice) to 1 (pure liquid). It is important to note that this ratio is a frequency ratio, computed through the lidar simulator and therefore consistent with the observations, rather than a mass ratio, computed from liquid and ice water content as often used in climate model studies (Choi et al., 2010; McCoy et al., 2014; Tsushima et al., 2006). Figure 2 shows the relationship between LPR and temperature between 70°S and 40°S (vertical black lines in Figure 1) during the austral summer in CALIPSO-GOCCP and the CMIP6 models.

Our results show that most models simulate cloud fractions that are in good agreement with the observations in the mixed-phase temperature range (0–−40°C), as evidenced by the multimodel mean. We also find that all models that partition liquid and ice using complex microphysics schemes (Table S1 in Supporting Information S1) reproduce the temperature-LPR mean relationship fairly well (Figure 2, their average is shown in dotted red), except HadGEM3-GC31-LL, which substantially underestimates it. By contrast, the four models that mostly rely on temperature to determine LPR, referred to as the T-dependent models, substantially overestimate LPR for all temperatures (Figure 2, dashed red), meaning that they simulate too many liquid clouds with respect to all clouds compared to the observations. Complex-microphysics models also better represent the variability of temperature-LPR relationship, but most models, regardless of the complexity of their microphysics, simulate too many occurrences of clouds at mixed-phase temperatures (between 0° and −40°C; Figure S2 in Supporting Information S1). However, 3 out of 4 T-dependent models (BCC-CSM2-MR, CNRM-CM6-1 and IPSL-CM6A-LR) and HadGEM3-GC31-LL simulate mixed-phase clouds (LPR between 0.1 and 0.9) within a range of temperature that is too narrow—and too cold for the T-dependent models—compared to observations (Figure S2 in Supporting Information S1). Additionally, T-dependent models tend to overestimate supercooled liquid-dominated cloud frequency (LPR > 0.9) at relatively cold temperatures (up to −25°C). At mixed-phase temperatures, T-dependent models—which constantly simulate both liquid and ice in every gridbox—produce more diagnosed liquid-dominated occurrences than other models would do, because the presence of liquid droplets dominates the lidar diagnostic (observed and consistently simulated by the simulator; Cesana & Chepfer, 2013).

In summary, we report that the complex-microphysics models better reproduce the LPR mean and variability than the T-dependent models, which clearly overestimate the LPR mean. Considering these results, one would expect T-dependent models, which simulate more occurrences of bright liquid clouds, to reflect a greater amount of SW radiation at TOA than their counterparts, yet the evidence shows otherwise (Figure 1 and Figure S1 in Supporting Information S1).

3.3. Impact of Cloud Phase on Radiation

Building on our above findings, we now aim at better understanding why the simulated clouds do not reflect enough solar radiation back to space poleward of 55°S, which does not seem to be solely the result of cloud occurrence or cloud phase biases, but rather a combination of both. To this end, we examine how profiles of cloud fraction north and south of 55°S compare with observations and how they affect their corresponding $SWCRE_{TOA}$ biases. When considering liquid and ice clouds together, the CMIP6 multimodel mean shows substantial improvements compared to CMIP5 models (Figure S3 in Supporting Information S1): more clouds below 3 km south of 55°S and between 3 and 6 km south of 55°S, suggesting that CMIP6 models simulate geometrically thicker clouds, in better agreement with observations. Should the radiative properties of CMIP6 SO clouds be perfectly simulated, their multimodel mean $SWCRE_{TOA}$ would match the observations. Focusing on liquid clouds, Figure 3 shows that CMIP6 multimodel mean captures the observed zonal mean profile of liquid cloud fraction (LCF) very well in both regions, although we note non-negligible inter-model variability. However, as noted in Section 3.1, the CMIP6 model low clouds are not bright enough south of 55°S, for a given cloud fraction (Figure S1 in Supporting Information S1), which is why their SW cloud reflection is underestimated on average (Figure 3) as well as their absorbed solar radiation (Figure S4 in Supporting Information S1, Trenberth & Fasullo, 2010). Such a bias could originate from either the BL clouds, whose optical thickness could be too small when transitioning to subfreezing temperatures (i.e., south of 55°S), or from mid-level topped clouds, which are now better represented compared to CMIP5 models, but still slightly underestimated compared to observations and whose optical thickness could be too small. Mixed-phase temperatures are ubiquitous south of 55°S, therefore the model with the smallest LPR (HadGEM3-GC31-LL) shows the largest lack of SW reflection compared to the observations, as expected, although other errors in microphysical properties of clouds may also contribute to radiative biases. The three other models with the largest lack of SW reflection are T-dependent models (CNRM-CM6-1, IPSL-CM6A-LR and MRI-ESM2-0), which produce more frequent supercooled liquid clouds at cold temperatures than the observations. Their lack of SW reflection is likely due to the fact that at mixed-phase temperatures, these models constantly simulate mixed-phase clouds—in contrast with observations—that are less optically thick than pure liquid clouds. We confirm that both the ability of the model to simulate the correct cloud regime, characterized here by the vertical structure of clouds, and the cloud phase representation are primary contributors to $SWCRE_{TOA}$, by using a perturbed parameter ensemble of six radiatively balanced configurations

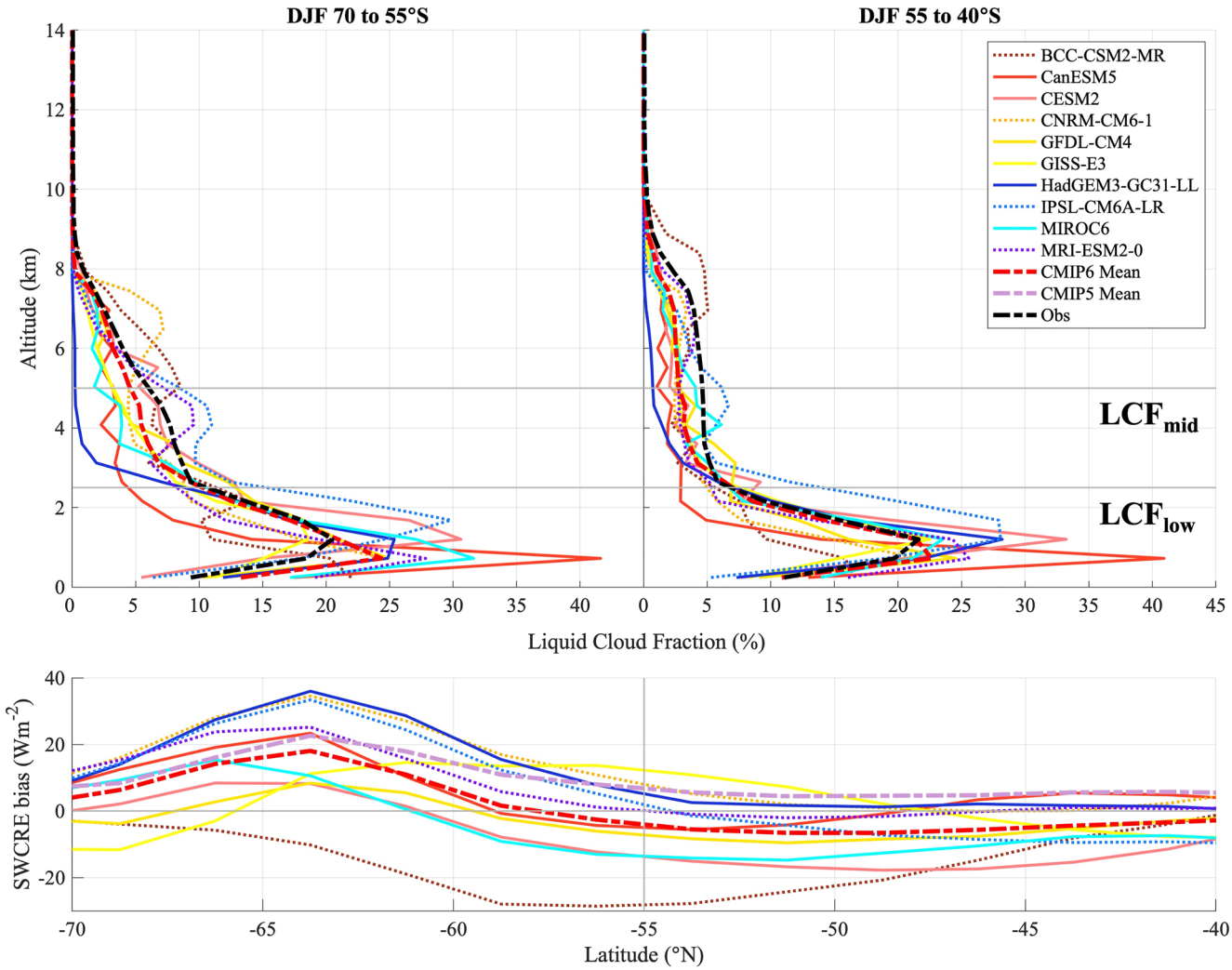

Figure 3. Relationship between liquid cloud fraction (LCF) profiles and $SWCRE_{TOA}$ above the Southern Ocean. (top) During the austral summer (December January February (DJF)) zonal profiles of LCF (%) for the Coupled Model Intercomparison Project Phase 6 (CMIP6) lidar simulator outputs (colored lines) and their mean (red dashed-dotted line) and for CALIPSO-GOCCP (black dashed-dotted line, 2008–2015 DJF) south (left) and north (right) of 55°S. (bottom) Corresponding $SWCRE_{TOA}$ CMIP6 individual and multimodel mean bias (Wm^{-2}), defined as model minus observations. The purple dashed-dotted line shows the corresponding 10 CMIP5 multimodel mean. A positive and negative bias means an underestimate and overestimate of SW reflection from clouds, respectively. The dotted lines correspond to T-dependent models.

of GISS-ModelE3 (Text S2 in Supporting Information S1). These results are consistent with what has been hinted by previous single-model studies (Bodas-Salcedo et al., 2016; Kay et al., 2016).

Figure 3 only assesses the mean state over each latitudinal band and provides little insight about the low and mid-level cloud overlap. To address this question, we further examine the $SWCRE_{TOA}$ as a function of the mean liquid cloud fractions between 0 and 2.5 km and 2.5 and 5 km, referred to as LCF_{low} and LCF_{mid}, within each $2 \times 2°$ gridbox and for each individual monthly mean profile (Figure 4). The observations indicate that nearly half of the cloud occurrences north of 55°S (43%, the largest circles, Figure 4) are BL clouds (moderate to large LCF_{low} on *x*-axis) with little to no overlying clouds ($LCF_{mid} < 2\%$ on *y*-axis). Yet, low- and mid-level clouds also often occur simultaneously, which we interpret as the signature of vertically extended clouds, such as frontal clouds, and more generally correspond to clouds gradually evolving into one regime or the other. The $SWCRE_{TOA}$ varies between about −60 and −130 Wm^{-2} as LCF_{low} and/or LCF_{mid} increase, consistent with an increased cloud optical thickness. On the one hand, most models simulate brighter clouds (more negative $SWCRE_{TOA}$) with increasing LCF_{low} and LCF_{mid}, as in the observations. On the other hand, none is able to reproduce the specific regime that is the most common regime in the observations, which corresponds to a low to large $LCF_{liq,low}$ with

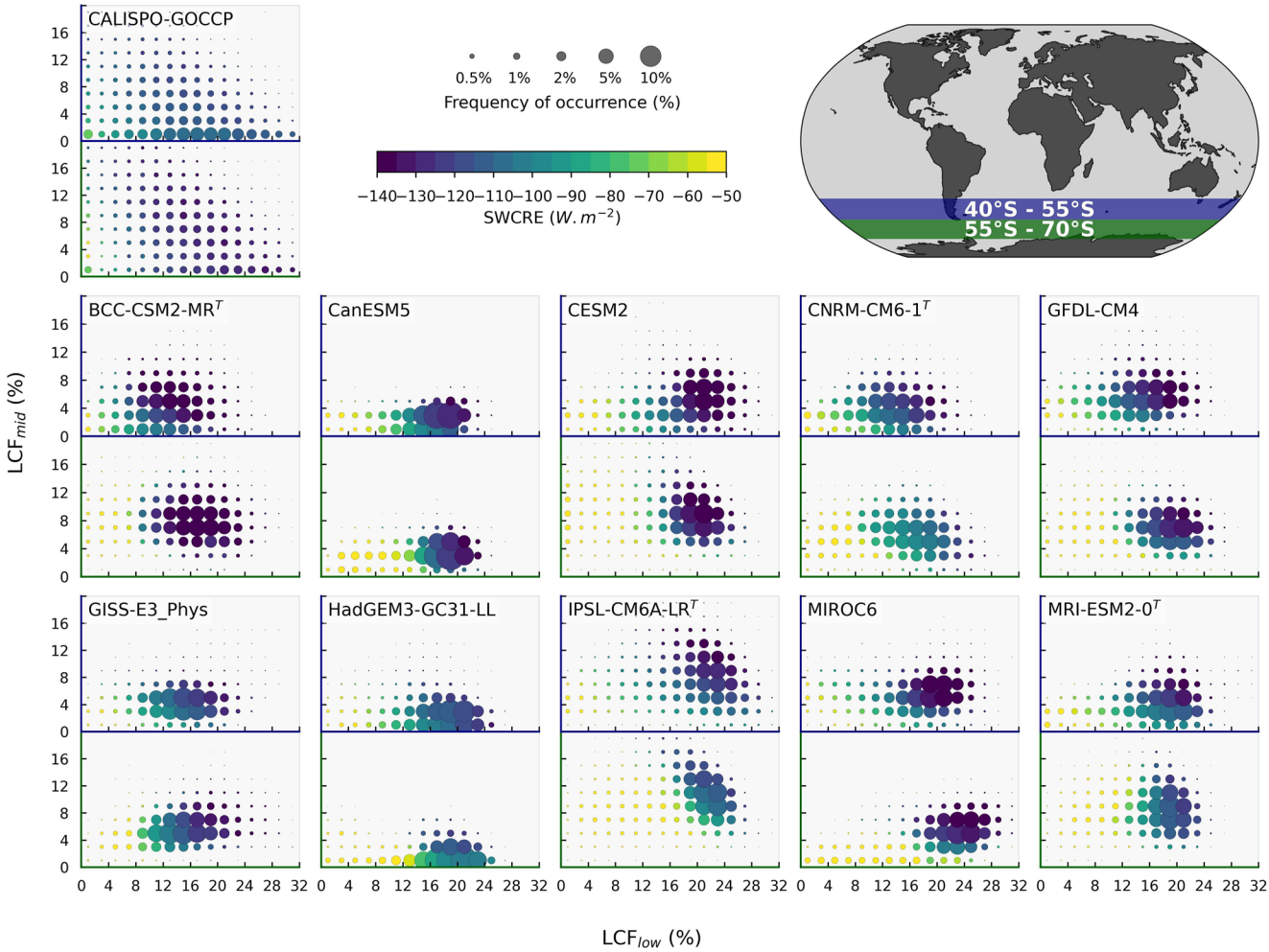

Figure 4. Relationship between cloud phase and $SWCRE_{TOA}$. 2D-histogram of $SWCRE_{TOA}$ as a function of the mean liquid cloud fraction (LCF) between 2.5 and 5 km (*y*-axis, LCF_{mid}) and the mean LCF between 0 and 2.5 km (*x*-axis, LCF_{low}) from Coupled Model Intercomparison Project Phase 6 lidar simulator outputs and General Circulation Model-Oriented CALIPSO Cloud Product (2008–2015 December January February (DJF)) and Clouds and the Earth's Radiant Energy System (2008–2015 DJF) observations. The size of the circles represents the proportion of profiles with respect to all profiles within each latitudinal band (note that the size of the circles does not use a linear scale), while the shading corresponds to the $SWCRE_{TOA}$ in each $2 \times 2°$ bin. In each row, the upper and bottom plots correspond to 40°–55°S and 55°–70°S latitudinal bands, respectively.

no or little overlying LCF_{mid}. We note however that two models, namely CanESM5 and HadGEM3-GC31-LL, come close to the observations. Moreover, the models struggle to reproduce the observed variability, and most of them tend to strongly favor a limited range of LCF_{low} always topped with mid-level clouds. Poleward of 55°S, the observations indicate that the pure BL cloud regime (moderate to large LCF_{low} with no LCF_{mid}) largely decreases (from 43% to 16% of the occurrences) and more clouds form in the mid-levels (see also Figure S6 in Supporting Information S1), resulting in brighter clouds (more negative $SWCRE_{TOA}$; Figure S7 in Supporting Information S1). By contrast, CMIP6 $SWCRE_{TOA}$ systematically produces less bright clouds (Figure 4 and Figure S7 in Supporting Information S1)—except for GISS-ModelE3—although they generate more overlying clouds than north of 55°S (the large circles move higher up in the bottom panels compared to the upper panels), in agreement with the observations (Figure 4 and Figure S6 in Supporting Information S1). Such a result suggests that the inability of the models to produce the correct cloud regimes coupled to their incorrect cloud phase are at the heart of the long-standing lack of outgoing SW radiation at TOA. More specifically, most models tend to generate mid-level clouds too frequently, which potentially include less reflective ice crystals that may grow at the expense of liquid droplets through the Wegener-Bergeron-Findeisen process, accretion, or both (except for IPSL-CM6A model).

4. Conclusions

In this study, we take advantage of CMIP6 cloud phase profiles using a lidar simulator to evaluate the relationship between cloud distribution, cloud phase and radiation over the SO (40°S–70°S), which has remained a long-standing problem in modern ESMs. We find that the newest generation of CMIP models simulate more low- and mid-level clouds in the extratropics, in better agreement with the observations. Concurrently, the lack of SW reflection from clouds is also somewhat reduced but not enough south of 55°S, where mixed-phase clouds are ubiquitous due to constant subfreezing temperatures during the austral summer. We then show that models with complex microphysics schemes better reproduce the observed LPR variability and mean, while the T-dependent models simulate a narrower range of temperature for a given LPR—in the mixed-phase regimes—with too many liquid-dominated clouds. However, the T-dependent models do not produce brighter clouds than their complex counterparts between 55° and 70°S. We speculate that this is related to the fundamental nature of T-dependent schemes that always produce liquid and ice in clouds at mixed-phase temperatures, thereby reducing their reflectivity compared to liquid-only clouds, which are frequent in complex microphysics models and observations. Finally, we show evidence that CMIP6 models still struggle to generate the correct SW reflection poleward of 55°S likely because of a combination of incorrect optical properties and regime variability of mid-level topped clouds forming in the BL, consistent with previous literature (Bodas-Salcedo et al., 2016; Kay et al., 2016). A shortcoming of this study is that we do not evaluate the condensed water path of clouds, which may be wrong—even if the cloud fraction and phase are correctly simulated by a model—and may cause additional radiative bias. Unfortunately, retrievals of liquid and ice water path (IWP) (or content) on a global scale carry large uncertainties especially in regions of substantial precipitation (e.g., over the SO; Elsaesser et al., 2017; Lebsock & Su, 2014) and cannot be used to quantitatively evaluate climate models. We note however, that the cloud fraction and cloud phase are two of the first order variables in determining $SWCRE_{TOA}$ biases over the SO because the IWP of clouds is typically far smaller than their liquid water path (e.g., McFarquhar et al., 2021) and because liquid clouds are often more reflective than ice clouds for a fixed amount of condensed water (e.g., Rogers & Yau, 1989).

A specificity of this study stems from the fact that we separate SO clouds into two different latitudinal bands based on the location of the 0°C isotherm: the first is dominated by warm low-level liquid cloud regime (north of 55°S) and the other comprises of low- and mid-level topped supercooled cloud regimes (south of 55°S), which causes the lack of SW reflection. We find that CMIP6 clouds are systematically less reflective south of 55°S regardless of the cloud regime, which is the opposite of what is observed. We speculate that when cold cloud processes are involved, the models fail to sustain the same or even larger liquid water content as in warm clouds, which can happen in GISS-ModelE3 when enhancing the Wegener-Bergeron-Findeisen process for example, (not shown).

Finally, in the context of SO cloud feedbacks being particularly relevant for models' climate sensitivity (Zelinka et al., 2020), one may wonder if the increase in vertically extended clouds from CMIP5 to CMIP6 models has played a role in the decrease of the negative cloud feedback over the SO; and how biases in the representation of different cloud regimes, and by extension cloud types, affect the overall SO cloud feedback, given that different cloud types have distinct cloud feedbacks (Cesana & del Genio, 2021; Myers et al., 2021).

Data Availability Statement

The CALIPSO-GOCCP observations can be downloaded from the CFMIP-Obs website (https://climserv.ipsl.polytechnique.fr/cfmip-obs/data/GOCCP/). The CMIP ESM outputs used in this study—listed in Table S1 in Supporting Information S1—were downloaded from the ESGF (https://esgf-node.llnl.gov/). The GISS-ModelE3 data are available at https://doi.org/10.5281/zenodo.6564079. CERES-EBAF 4.1 TOA fluxes were downloaded on the CERES website (https://ceres.larc.nasa.gov/data/#ebaftoa-level-3). Three reanalysis data sets were used in the study for temperatures: the fifth generation European Centre for Medium-Range Weather Forecasts atmospheric reanalysis (ERA5, downloaded from https://doi.org/10.24381/cds.6860a573 and by selecting monthly averaged reanalysis and the temperature variable, the National Oceanic and Atmospheric Administration (NOAA)/Cooperative Institute for Research in Environmental Sciences (CIRES)/DOE 20th Century Reanalysis v3, downloaded from https://psl.noaa.gov/data/gridded/data.20thC_ReanV3.html, and National Centers for Environmental Prediction/Department of Energy (NCEP/DOE) reanalysis 2, downloaded from http://www.cpc.ncep.noaa.gov/products/wesley/reanalysis2/.

Acknowledgments
GC was partly supported by a CloudSat-CALIPSO RTOP (Grant 80NSSC18M0133) and NASA Modeling, Analysis, and Prediction Program. We thank NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center for providing computing resources to run GISS ESM, NASA and CNES for giving access to CALIPSO observations, and Climserv for giving access to CALIPSO-GOCCP observations and for providing computing resources. We thank EECLAT and INSU for providing resources to organize a workshop that nurtures discussions between the authors. We also acknowledge the World Climate Research Programme's Working Group on Coupled Modeling, which is responsible for CMIP, and thank the climate modeling groups for producing and making available their model output. We also thank Jennifer Kay for providing additional CESM2 AMIP simulations and Tomoo Ogura and Takuro Michibata for MIROC6 simulations. Finally, we thank the GISS development team (Andrew Ackerman, Gregory Elsaesser, Ann Fridlind and Maxwell Kelley) for useful discussions and for providing GISS-ModelE3 outputs.

References

Bodas-Salcedo, A., Hill, P. G., Furtado, K., Williams, K. D., Field, P. R., Manners, J. C., et al. (2016). Large contribution of supercooled liquid clouds to the solar radiation budget of the Southern Ocean. *Journal of Climate*, *29*(11), 4213–4228. https://doi.org/10.1175/JCLI-D-15-0564.1

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*(6), 1735–1758. https://doi.org/10.1029/2019MS001688

Cesana, G., & Chepfer, H. (2012). How well do climate models simulate cloud vertical structure? A comparison between CALIPSO-GOCCP satellite observations and CMIP5 models. *Geophysical Research Letters*, *39*(20), 1–6. https://doi.org/10.1029/2012GL053153

Cesana, G., & Chepfer, H. (2013). Evaluation of the cloud thermodynamic phase in a climate model using CALIPSO-GOCCP. *Journal of Geophysical Research: Atmospheres*, *118*(14), 7922–7937. https://doi.org/10.1002/jgrd.50376

Cesana, G., Kay, J. E., Chepfer, H., English, J. M., & de Boer, G. (2012). Ubiquitous low-level liquid-containing Arctic clouds: New observations and climate model constraints from CALIPSO-GOCCP. *Geophysical Research Letters*, *39*(20), 1–6. https://doi.org/10.1029/2012GL053385

Cesana, G., & Storelvmo, T. (2017). Improving climate projections by understanding how cloud phase affects radiation. *Journal of Geophysical Research*, *122*(8), 4594–4599. https://doi.org/10.1002/2017JD026927

Cesana, G., & Waliser, D. E. (2016). Characterizing and understanding systematic biases in the vertical structure of clouds in CMIP5/CFMIP2 models. *Geophysical Research Letters*, *43*(19), 10538–10546. https://doi.org/10.1002/2016GL070515

Cesana, G., Waliser, D. E., Jiang, X., & Li, J. L. F. (2015). Multimodel evaluation of cloud phase transition using satellite and reanalysis data. *Journal of Geophysical Research*, *120*(15), 7871–7892. https://doi.org/10.1002/2014JD022932

Cesana, G. V., Ackerman, A. S., Fridlind, A. M., Silber, I., & Kelley, M. (2021). Snow reconciles observed and simulated phase partitioning and increases cloud feedback. *Geophysical Research Letters*, *48*(20), 1–11. https://doi.org/10.1029/2021gl094876

Cesana, G. V., & del Genio, A. D. (2021). Observational constraint on cloud feedbacks suggests moderate climate sensitivity. *Nature Climate Change*, *11*(3), 213–218. https://doi.org/10.1038/s41558-020-00970-y

Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J. L., Minnis, P., et al. (2010). The GCM-oriented CALIPSO cloud product (CALIPSO-GOCCP). *Journal of Geophysical Research*, *115*(5), D00H16. https://doi.org/10.1029/2009JD012251

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., & Sèze, G. (2008). Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophysical Research Letters*, *35*(15), L15704. https://doi.org/10.1029/2008GL034207

Choi, Y. S., Ho, C. H., Kim, S. W., & Lindzen, R. S. (2010). Observational diagnosis of cloud phase in the winter Antarctic atmosphere for parameterizations in climate models. *Advances in Atmospheric Sciences*, *27*(6), 1233–1245. https://doi.org/10.1007/s00376-010-9175-3

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*(2). https://doi.org/10.1029/2019MS001916

Elsaesser, G. S., O'Dell, C. W., Lebsock, M. D., Bennartz, R., Greenwald, T. J., & Wentz, F. J. (2017). The multisensor advanced climatology of liquid water path (MAC-LWP). *Journal of Climate*, *30*(24), 10193–10210. https://doi.org/10.1175/JCLI-D-16-0902.1

Frey, W. R., & Kay, J. E. (2018). The influence of extratropical cloud phase and amount feedbacks on climate sensitivity. *Climate Dynamics*, *50*(7–8), 3097–3116. https://doi.org/10.1007/s00382-017-3796-5

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*(14), 8329–8337. https://doi.org/10.1029/2019GL083978

Held, I. M., Guo, H., Adcroft, A., Dunne, J. P., Horowitz, L. W., Krasting, J., et al. (2019). Structure and performance of GFDL's CM4.0 climate model. *Journal of Advances in Modeling Earth Systems*, *11*(11), 3691–3727. https://doi.org/10.1029/2019MS001829

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

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

Kay, J. E., Bourdages, L., Miller, N. B., Morrison, A., Yettella, V., Chepfer, H., & Eaton, B. (2016). Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations. *Journal of Geophysical Research: Atmospheres*, *121*(8), 4162–4176. https://doi.org/10.1002/2015JD024699

Kelley, M., Schmidt, G. A., Nazarenko, L. S., Bauer, S. E., Ruedy, R., Russell, G. L., et al. (2020). GISS-E2.1: Configurations and climatology. *Journal of Advances in Modeling Earth Systems*, *12*(8). https://doi.org/10.1029/2019MS002025

Komurcu, M., Storelvmo, T., Tan, I., Lohmann, U., Yun, Y., Penner, J. E., et al. (2014). Intercomparison of the cloud water phase among global climate models. *Journal of Geophysical Research: Atmospheres*, *119*(6), 3372–3400. https://doi.org/10.1002/2013JD021119

Lebsock, M., & Su, H. (2014). Application of active spaceborne remote sensing for understanding biases between passive cloud water path retrievals. *Journal of Geophysical Research: Atmospheres*, *119*(14), 8962–8979. https://doi.org/10.1002/2014JD021568

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

Madeleine, J. B., Hourdin, F., Grandpeix, J. Y., Rio, C., Dufresne, J. L., Vignon, E., et al. (2020). Improved representation of clouds in the atmospheric component LMDZ6A of the IPSL-CM6A Earth system model. *Journal of Advances in Modeling Earth Systems*, *12*(10). https://doi.org/10.1029/2020MS002046

Matus, A. V., & L'Ecuyer, T. S. (2017). The role of cloud phase in Earth's radiation budget. *Journal of Geophysical Research*, *122*(5), 2559–2578. https://doi.org/10.1002/2016JD025951

McCoy, D. T., Hartmann, D. L., & Grosvenor, D. P. (2014). Observed Southern Ocean cloud properties and shortwave reflection. Part II: Phase changes and low cloud feedback. *Journal of Climate*, *27*(23), 8858–8868. https://doi.org/10.1175/jcli-d-14-00288.1

McCoy, D. T., Hartmann, D. L., Zelinka, M. D., Ceppi, P., & Grosvenor, D. P. (2015). Mixed-phase cloud physics and Southern Ocean cloud feedback in climate models. *Journal of Geophysical Research*, *120*(18), 9539–9554. https://doi.org/10.1002/2015JD023603

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., et al. (2021). Observations of clouds, aerosols, precipitation, and surface radiation over the Southern Ocean: An overview of CAPRICORN, MARCUS, MICRE, and SOCRATES. *Bulletin of the American Meteorological Society*, *102*(4), E894–E928. https://doi.org/10.1175/BAMS-D-20-0132.1

Myers, T. A., Scott, R. C., Zelinka, M. D., Klein, S. A., Norris, J. R., & Caldwell, P. M. (2021). Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity. *Nature Climate Change*, *11*(6), 501–507. https://doi.org/10.1038/s41558-021-01039-0

Nam, C., Bony, S., Dufresne, J., & Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophysical Research Letters*, *39*(21), 1–7. https://doi.org/10.1029/2012GL053421

Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J. F., et al. (2020). The CNRM global atmosphere model ARPEGE-climat 6.3: Description and evaluation. *Journal of Advances in Modeling Earth Systems*, *12*(7). https://doi.org/10.1029/2020MS002075

Rogers, R. R., & Yau, M. K. (1989). *A short course in cloud physics* (3rd ed.). Pergamon Press.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., et al. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876–2908. https://doi.org/10.1002/qj.3598

Stramler, K., del Genio, A. D., & Rossow, W. B. (2011). Synoptically driven Arctic winter states. *Journal of Climate*, *24*(6), 1747–1762. https://doi.org/10.1175/2010JCLI3817.1

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*(2), 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*(2–3), 113–126. https://doi.org/10.1007/s00382-006-0127-7

Wang, H., & Su, W. (2013). Evaluating and understanding top of the atmosphere cloud radiative effects in intergovernmental panel on climate change (IPCC) fifth assessment report (AR5) coupled model intercomparison project phase 5 (CMIP5) models using satellite observations. *Journal of Geophysical Research: Atmospheres*, *118*(2), 683–699. https://doi.org/10.1029/2012JD018619

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*(1), 1–12. https://doi.org/10.1029/2019GL085782

Zhang, M. H., Lin, W. Y., Klein, S. A., Bacmeister, J. T., Bony, S., Cederwall, R. T., et al. (2005). Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *Journal of Geophysical Research*, *110*(D15), 1–18. https://doi.org/10.1029/2004JD005021

References From the Supporting Information

Cesana, G., Chepfer, H., Winker, D., Getzewich, B., Cai, X., Jourdan, O., et al. (2016). Using in situ airborne measurements to evaluate three cloud phase products derived from CALIPSO. *Journal of Geophysical Research*, *121*(10), 5788–5808. https://doi.org/10.1002/2015JD024334

Swart, N. C., Cole, J. N. S., Kharin, V. V., Lazare, M., Scinocca, J. F., Gillett, N. P., et al. (2019). The Canadian Earth system model version 5 (CanESM5.0.3). *Geoscientific Model Development*, *12*(11), 4823–4873. https://doi.org/10.5194/gmd-12-4823-2019

Tatebe, H., Ogura, T., Nitta, T., Komuro, Y., Ogochi, K., Takemura, T., et al. (2019). Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geoscientific Model Development*, *12*(7), 2727–2765. https://doi.org/10.5194/gmd-12-2727-2019

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., et al. (2018). The met office global coupled model 3.0 and 3.1 (GC3.0 and GC3.1) configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380. https://doi.org/10.1002/2017MS001115

Wu, T., Lu, Y., Fang, Y., Xin, X., Li, L., Li, W., et al. (2019). The Beijing climate center climate system model (BCC-CSM): The main progress from CMIP5 to CMIP6. *Geoscientific Model Development*, *12*(4), 1573–1600. https://doi.org/10.5194/gmd-12-1573-2019

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., et al. (2019). The meteorological research institute Earth system model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan*, *97*(5), 931–965. https://doi.org/10.2151/jmsj.2019-051