JGR Atmospheres

RESEARCH ARTICLE
10.1029/2022JD038437

Key Points:

- We introduce a new approach to distinguish stratocumulus and shallow cumulus regimes over tropical oceans based on cloud cover
- The "too few, too bright" tropical low-cloud problem persists in 12 CMIP6 models within stratocumulus and shallow cumulus regimes
- Most CMIP6 models underestimate (overestimate) the relative frequency of occurrence of stratocumulus (shallow cumulus)

Supporting Information:

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

Correspondence to:

N. Črnivec,
nina.crnivec@gmail.com

Citation:

Črnivec, N., Cesana, G., & Pincus, R. (2023). Evaluating the representation of tropical stratocumulus and shallow cumulus clouds as well as their radiative effects in CMIP6 models using satellite observations. *Journal of Geophysical Research: Atmospheres*, *128*, e2022JD038437. https://doi.org/10.1029/2022JD038437

Received 23 DEC 2022
Accepted 9 NOV 2023

Evaluating the Representation of Tropical Stratocumulus and Shallow Cumulus Clouds As Well As Their Radiative Effects in CMIP6 Models Using Satellite Observations

Nina Črnivec[1,2], Grégory Cesana[1,2], and Robert Pincus[3]

[1]Center for Climate Systems Research, Columbia University, New York, NY, USA, [2]NASA Goddard Institute for Space Studies, New York, NY, USA, [3]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

Abstract Low clouds over tropical oceans reflect a great proportion of solar radiation back to space and thereby cool the Earth, yet this phenomenon has been poorly simulated in several previous generations of climate models. The principal aim of the present study is to employ satellite observations to evaluate the representation of marine tropical low clouds and their radiative effect at the top of the atmosphere in a subset of latest climate models participating in CMIP6. We strive for regime-oriented model validation and hence introduce a qualitative approach to discriminate stratocumulus (Sc) from shallow cumulus (Cu). The novel Sc-Cu categorization has a conceptual advantage of being based on cloud properties, rather than relying on a model response to a cloud-controlling factor. We find that CMIP6 models underestimate low-cloud cover in both Sc-regions and Cu-regions of tropical oceans. A more detailed investigation of cloud biases reveals that most CMIP6 models underestimate the relative frequency of occurrence (RFO) of Sc and overestimate RFO of Cu. We further demonstrate that tropical low cloudiness in CMIP6 models remains too bright. The regime-oriented validation represents the basis for improving parameterizations of physical processes that determine the cloud cover and radiative impact of Sc and Cu, which are still misrepresented in current climate models.

Plain Language Summary Similar as white snow and ice caps, bright low clouds have a high shortwave albedo, reflecting a huge amount of sunlight back to space and thereby helping us counteract global warming. The shadowing effect of bright low clouds is especially pronounced over tropical oceans, since equatorial regions of our planet receive most sunshine, which is in clear skies otherwise practically entirely absorbed within the contrastingly dark ocean. Climate models had traditionally struggled simulating these clouds by underestimating their areal extent and simultaneously overestimating their reflectivity. In other words, simulated clouds were commonly found to be "too few" and "too bright" compared to observations, which introduced a substantial uncertainty to climate projections. Herein, we proposed a novel approach to proficiently decompose tropical low cloudiness into stratocumulus and shallow cumulus regime, which is essential to provide a proper guidance for climate model development. We subsequently showed that the newest generation of climate models still suffers from the "too few, too bright" tropical low-cloud problem within both stratocumulus and shallow cumulus regimes, which thus needs to be further tackled with the greatest possible endeavor.

1. Introduction

Bright low clouds cover substantial areas of dark tropical oceans and play a critical role in regulating the Earth's radiative energy budget (Bony & Dufresne, 2005; Cesana & Del Genio, 2021; Schneider et al., 2017). They reflect a substantial portion of the incoming sunlight back to space and thus exert a profound cooling effect on the Earth's climate. Climate models, however, have a longstanding problem simulating these clouds, which limits our ability to accurately predict the amount of global warming caused by rising greenhouse gas emissions (Bony & Dufresne, 2005; Sherwood et al., 2020).

A major issue, which persisted in previous generations of climate models of the World Climate Research Programme's Coupled Model Intercomparison Project (CMIP), is known as the "too few, too bright" tropical low-cloud problem (Nam et al., 2012). In brief, climate models commonly underestimated the amount of tropical and subtropical low-level clouds (e.g., Cesana & Chepfer, 2012; Cesana & Waliser, 2016; Cesana, Waliser, et al., 2019; Teixeira et al., 2011) and simultaneously overestimated their reflectance (e.g., Karlsson et al., 2008;

Nam et al., 2012; Weare, 2004). In the present study, we revisit the "too few, too bright" tropical low-cloud problem in latest climate models participating in Phase 6 of CMIP (CMIP6; Eyring et al., 2016b), which are currently being evaluated as a community effort (Schuddeboom & McDonald, 2021; Tselioudis et al., 2021).

Satellite observations have been regularly exploited to assess the fidelity of climate models (e.g., Cesana & Chepfer, 2012; Cesana, Waliser, et al., 2019; Dolinar et al., 2015; Jiang et al., 2012; Pincus et al., 2008), being especially valuable due to their extensive coverage. Evaluating clouds (and other fields) simulated by global climate models (GCMs) using satellite observations, however, is challenging, because satellite-borne instruments do not directly measure meteorological quantities of interest as simulated by GCMs. In order to facilitate the comparison between observed and model-simulated fields, the Cloud Feedback Model Intercomparison Project (CFMIP) community introduced the CFMIP Observation Simulator Package (COSP; Bodas-Salcedo et al., 2011; Swales et al., 2018). Given the atmospheric data provided by a GCM, the COSP software reproduces observations of multiple instruments on board of various satellite missions such as CloudSat (Stephens et al., 2002) and Cloud-Aerosols Lidar and Infrared Pathfinder Satellite Observations (CALIPSO; Winker et al., 2010) within the A-train constellation (Stephens et al., 2002, 2018).

To summarize, the objective of the present study is to employ satellite observations to evaluate tropical low clouds together with their radiative effect at the top of the atmosphere (TOA) in a subset of CMIP6 models. In particular, we strive to assess model representation of individual low-cloud regimes commonly found in tropical areas of large-scale subsidence—namely the eastern ocean stratocumulus (Sc) and trade wind shallow cumulus (Cu). These essentially contrasting cloud types are driven by a distinct interplay of small-scale processes within the moist marine boundary layer—convection, turbulence, radiation, and cloud microphysics (Ackerman & Toon, 1996; Ackerman et al., 1993, 2000, 2009; Stevens et al., 2001), which are often poorly and inconsistently parameterized across climate models (Bony et al., 2015; Klein et al., 2017; Randall et al., 2003; Stevens & Bony, 2013). A cloud regime-oriented model evaluation would help identify shortcomings of physical parameterization schemes, which govern the formation and evolution of Sc and Cu clouds, and is thus a crucial first step toward more reliable climate change projections. Furthermore, Sc and Cu clouds exhibit a fundamentally contrasting response to the change in their controlling meteorological factors such as rising sea surface temperature and low-level inversion strength and are associated with different feedbacks (Cesana & Del Genio, 2021), which highlights the importance of a regime-based investigation. Accurate representation of geographical distributions of Sc and Cu clouds is thereby essential for realistic low-cloud feedbacks (Cesana & Del Genio, 2021).

To carry out such a cloud regime-oriented evaluation of climate models, however, one has to find a qualitative way to separate Sc from Cu, because CMIP diagnostics do not distinguish between stratiform and convective cloud covers. A classic way to separate Sc from Cu clouds in tropical and subtropical areas of large-scale subsidence is by means of the estimated inversion strength (EIS), which is known to be a good predictor of stratocumulus (Wood & Bretherton, 2006). Even though this approach generally works well in the real world, it has limitations when applied to climate models which misrepresent EIS. An important aim of the present work is to introduce a novel Sc-Cu categorization, which can be utilized to reliably separate contributions from Sc and Cu clouds in both observations and climate models.

The remainder of this paper is structured as follows. Section 2 presents observational, reanalysis, and CMIP6 data as well as various approaches to discriminate Sc from Cu. The evaluation of climate models using the Sc-Cu categorization introduced in this study is carried out in Section 3. A brief summary and concluding remarks are given in Section 4.

2. Data and Methods

2.1. Satellite Observations and Reanalysis Data

We utilize low-cloud cover (LCC) observations from the GCM-Oriented CALIPSO Cloud Product version 2.9 (CALIPSO-GOCCP; Cesana et al., 2016; Chepfer et al., 2010a), which was specifically designed to evaluate cloudiness simulated by GCMs using a lidar simulator. It is based on measurements taken by the CALIPSO Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP; Winker et al., 2010). We restrict our analysis to subsidence regimes over tropical and subtropical oceans (between 35°S and 35°N), where the pressure vertical velocity at 500 hPa exceeds 10 hPa day^{-1}. In these regions, the amount of high cloud is small and hence generates less attenuation of the lidar signal, thereby reducing the high-cloud shielding effect. Although

CALIPSO-GOCCP global and tropical LCCs are very similar to other satellite observations both quantitatively and qualitatively (Cesana et al., 2019a; Chepfer et al., 2010a; Stubenrauch et al., 2013), CALIPSO-GOCCP detects more broken low clouds (typically shallow Cu; Cesana et al., 2019a). Some misdiagnosing of fully attenuated pixels as being clear sky might cause underestimation of the cloud fraction in the lowermost levels of CALIPSO-GOCCP (up to 25% of the Arctic boundary-layer cloud cases). However, this potential misdiagnosing does not affect 2D LCCs used here, because the uppermost cloud level is always detected (Cesana et al., 2016).

To discriminate Sc from Cu, we use the recently created Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD) described in detail by Cesana et al. (2019a). This unique algorithm considers cloud morphology to classify low cloudiness into several categories including Sc, Cu, and various transitioning regimes (broken Sc, Cu under Sc, and Cu with stratiform outflow) at the orbital level. CASCCAD reports monthly values of cloud fraction over a 10-year period (2007–2017) and has a spatial resolution of 2.5° in both latitudinal and longitudinal directions.

The observed cloud-radiative effect (CRE) estimates at TOA are obtained from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) satellite product Ed. 4.1 (Loeb et al., 2018a). Specifically, the CERES-EBAF clear-sky and all-sky radiative fluxes are employed to compute the CRE (the uncertainties are discussed in Section 4 of Loeb et al. (2018a)). In the present study, we consider solely the shortwave (SW) CRE component, since low clouds contribute minorly to the longwave TOA radiation budget.

The middle-tropospheric pressure vertical velocity at 500 hPa, which is used to define the subsidence regimes, is derived from averaging three reanalysis data sets including Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA-2; Gelaro et al., 2017a), the fifth generation of ECMWF reanalysis (ERA-5; Hersbach et al., 2020) and the NCEP Department of Energy Atmospheric Model Intercomparison Project reanalysis (NCEP-DOE R-2; Kanamitsu et al., 2002). The same three reanalysis data sets are employed to compute EIS. All data were regridded to the CASCCAD spatial grid.

2.2. Climate Models

We analyze monthly mean output of CMIP6 climate model simulations, which relate to the Atmospheric Model Intercomparison Project (AMIP). The AMIP experiments use a prescribed sea surface temperature (and sea ice concentrations), which enables us to evaluate the atmospheric component of the fully coupled atmosphere-ocean models which are employed for climate change prediction. The LCC and CRE errors thereby originate from incomplete parameterizations of atmospheric processes (e.g., convection, cloud microphysics, boundary-layer turbulence, radiation) and not from sea surface temperature fields.

The CALIPSO lidar simulator (Chepfer et al., 2008) integrated in COSP is employed to consistently compare low cloudiness in observations and climate models. We investigate the realism of 12 model configurations stemming from different modeling centers given in Table 1, which provided the output of CALIPSO lidar simulator. To address the atmospheric variability within a vast GCM grid column, the COSP instrument simulator operates on a multitude of homogeneous subcolumns to reproduce satellite pixel variability. These subcolumns are normally produced within COSP in accordance with GCM's assumptions for subgrid cloud structure utilizing the Subgrid Cloud Overlap Profile Sampler (Webb et al., 2001). A few GCMs from the analyzed set employ COSP2 (Swales et al., 2018), whereby subcolumns can be adopted directly from a GCM, since they are often stochastically generated (Raisanen et al., 2004) within model's radiation scheme such as the commonly employed McICA algorithm (Pincus et al., 2003).

We employ 8 years of CMIP6 simulations (2007–2014), which overlap with the CASCCAD temporal range and regrid them to the observational spatial grid.

2.3. Approaches to Discriminate Sc From Cu

As pointed out in Section 1, it is challenging to evaluate the representation of Sc and Cu in climate model output, since only a single LCC variable is archived in the CMIP database. In the following, we first summarize the traditional approach to separate Sc from Cu based on environmental characteristics and subsequently present an alternative, new categorization.

Table 1
List of Analyzed CMIP6 Models

Model acronym	Modeling center	References
NorESM2-LM	Norwegian Climate Center, Norway	Seland et al. (2020)
MIROC6	Center for Climate System Research, Japan	Tatebe et al. (2019)
GISS-E2-1-G	NASA Goddard Institute for Space Studies, USA	Kelley et al. (2020)
BCC-CSM2-MR	Beijing Climate Center, China	Wu et al. (2019)
MRI-ESM2-0	Meteorological Research Institute, Japan	Yukimoto et al. (2019)
GFDL-CM4	NOAA Geophysical Fluid Dynamics Laboratory, USA	Held et al. (2019)
IPSL-CM6A-LR	Institute Pierre Simon Laplace, France	Boucher et al. (2020)
CNRM-CM6-1	National Centre for Meteorological Research, France	Voldoire et al. (2019)
CESM2	National Center for Atmospheric Research, USA	Danabasoglu et al. (2020)
E3SM-1-0	US Department of Energy, USA	Golaz et al. (2019)
HadGEM3-GC31-LL	Met Office Hadley Centre, UK	Andrews et al. (2020)
CanESM5	Centre for Climate Modeling and Analysis, Canada	Swart et al. (2019)

2.3.1. Categorization Based on EIS

A traditional approach to differentiate between individual low-cloud regimes in tropical and subtropical areas governed by large-scale subsidence is by means of some measure of lower-tropospheric stability (e.g., Myers et al., 2021; Nam et al., 2012), since regions of the main stratocumulus decks off the west coast of the continents are associated with stronger atmospheric stability than shallow cumulus areas in the trade winds. A convenient parameter to describe low-level atmospheric stability is the EIS (Wood & Bretherton, 2006) and an EIS threshold of 1 K—despite being imperfect (Cesana & Del Genio, 2021)—was commonly employed in previous work to determine whether (model) grid box is classified as being either Sc-dominated or Cu-dominated (e.g., Myers et al., 2021).

This categorization has additional shortcomings when applied to climate models. Figure 1 (left) shows probability density of EIS in averaged reanalysis (combining MERRA-2, ERA-5, NCEP data) and CMIP6 models. All analyzed models systematically underestimate EIS implied by reanalyses, although they relatively well capture the shape of the EIS distribution. This large underestimation of EIS in CMIP6 models implies that the traditional Sc-Cu categorization utilizing a fixed EIS threshold of 1 K tends to allocate an insufficient amount of LCC to the Sc component, while attributing an excessive amount of LCC to the Cu component in models.

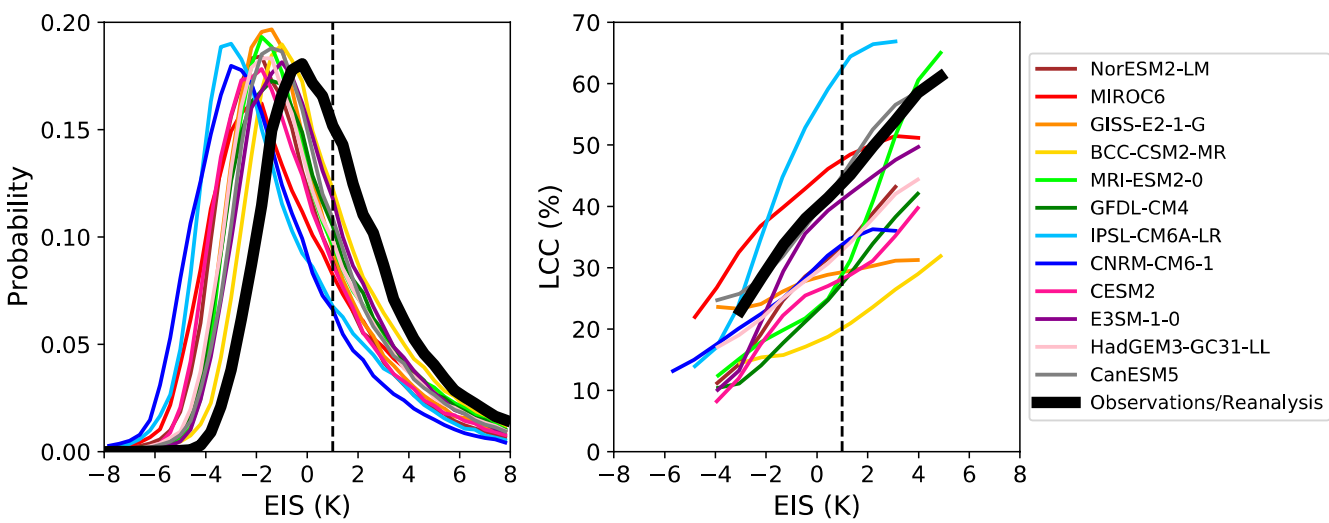

Figure 1. (left) Probability density functions of estimated inversion strength in subsidence regimes of tropical oceans derived from reanalyses and CMIP6 models. (right) Relationship between LCC and EIS in observations/reanalyses and CMIP6 models. The vertical dashed line at EIS of 1 K marks the common threshold used to discriminate Sc from Cu.

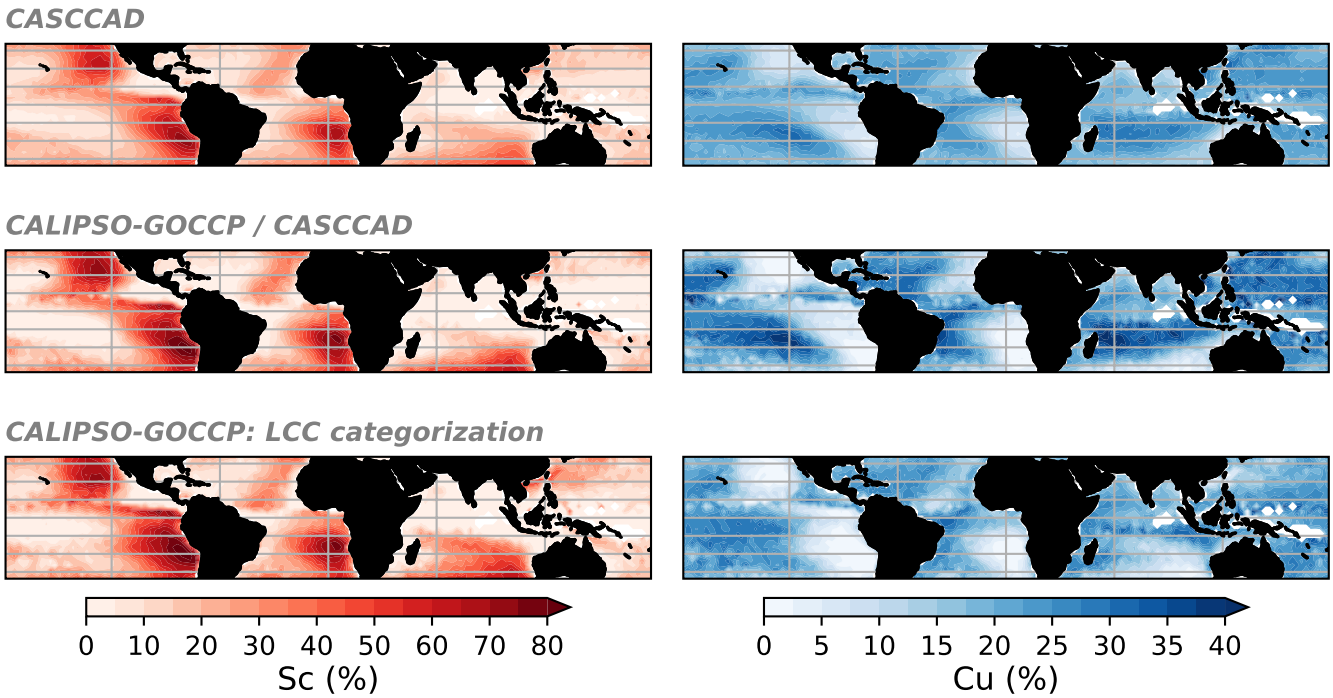

Figure 2. Reconstructed cloud-type distributions obtained with novel Sc-Cu categorization utilizing low-cloud cover (LCC) threshold being applied to CALIPSO-GOCCP observations of LCC (shown in bottom row). Benchmark Sc-cloud and Cu-cloud cover components (shown in middle row) are derived from CALIPSO-GOCCP LCC based on geographical locations of CASCCAD Sc and Cu (CASCCAD variables are shown in top row). The correlation coefficient between the cloud cover reconstructed using the LCC categorization and its benchmark counterpart is 0.93 for Sc and 0.66 for Cu. Note that only subsidence regimes of tropical and subtropical oceans are analyzed (data voids near Indonesia correspond to regions of persistent deep convection).

Figure 1 (right) additionally visualizes LCC as a function of EIS derived from observations/reanalyses and CMIP6 models. Observed LCC increases with an approximately constant rate of about 5% per K of EIS rise. Models have a differing ability to reproduce the rate of this increase, whereby multiple models strongly misrepresent the LCC-EIS relationship, making EIS not the ideal choice for discriminating among cloud types.

2.3.2. Categorization Based on LCC

Herein, we propose a new approach to discriminate Sc from Cu, which is based on cloud properties and thus alleviates the aforementioned problem. This categorization originates from the idea that overcast Sc scenarios typically have larger cloud cover than broken Cu cloud fields (Figure 2). In particular, we utilize the averaged LCC in analyzed tropical subsidence oceanic regions in each monthly time step (LCC_{thres}) to determine whether a grid box is dominated by Sc or Cu. The observed probability density functions (PDFs) of Sc-cloud and Cu-cloud cover derived from CALIPSO-GOCCP/CASCCAD namely exhibit a crossover approximately at this threshold. The observed (benchmark) Sc-cloud and Cu-cloud cover components (Figure 2, middle row) were thereby obtained from CALIPSO-GOCCP LCC data utilizing the relative frequency of occurrence (RFO) of a given cloud type derived from CASCCAD variables (e.g., $RFO_{Sc} = FO_{Sc}/(FO_{Sc} + FO_{Cu})$; where FO denotes the frequency of occurrence, and $RFO_{Cu} = 1 - RFO_{Sc}$) to determine whether grid box is classified as being either Sc-dominated ($RFO_{Sc} > 0.5$) or Cu-dominated ($RFO_{Sc} \leq 0.5$). It should be noted that we thereby incorporated CASCCAD transitioning regimes into the Sc component, whereby their contribution to Sc-cloud cover is small (Cesana et al., 2019a). This action is further justified by the fact that Sc-Cu transitioning categories ("broken Sc," "Cu under Sc," "Cu with stratiform outflow") overall show a greater resemblance to the pure Sc-than to the pure Cu-category when comparing their morphological characteristics (Cesana et al., 2019a) as well as optical properties (Pincus et al., 1999).

Hence, a (model) grid box is classified as being Sc-dominated if its LCC exceeds LCC_{thres} and as Cu-dominated otherwise. The Sc-cloud and Cu-cloud cover components are thus constructed from an arbitrary LCC data set (i.e., GCM output and CALIPSO-GOCCP observations of LCC) as follows:

$$LCC_{Sc}(lat, lon, t) = \begin{cases} LCC(lat, lon, t), & \text{if } LCC(lat, lon, t) > LCC_{thres}(t) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$LCC_{Cu}(lat, lon, t) = \begin{cases} LCC(lat, lon, t), & \text{if } LCC(lat, lon, t) \leq LCC_{thres}(t) \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where lat and lon denote the latitude and longitude of a grid box and t is the monthly time step. It should thereby be emphasized that we utilize a model-dependent LCC threshold when employing a new categorization in GCMs. This implies that we utilize LCC in analyzed tropical subsidence oceanic regions in each monthly time step from each particular GCM to compute its averaged value, which is used as threshold in Equations 1 and 2. It should also be noted that the corresponding Sc and Cu components of the SW CRE are obtained in an analogous manner (from the CRE field following Equations 1 and 2 using the same LCC_{thres} to separate Sc from Cu). These definitions of Sc and Cu components of LCC and SW CRE are employed throughout this work when analyzing geographical distributions, whereas we utilize only within-regime values (i.e., excluding zeros in regions dominated by Cu and Sc, respectively) in the subsequent analysis.

Figure 2 evaluates geographical distributions of Sc-cloud and Cu-cloud cover created with the novel Sc-Cu categorization using LCC threshold being applied to CALIPSO-GOCCP observations of LCC. The benchmark Sc-cloud and Cu-cloud cover components are obtained from the same observational LCC data set (CALIPSO-GOCCP) utilizing RFO of a given cloud type derived from CASCCAD variables to determine whether grid box is Sc-dominated ($RFO_{Sc} > 0.5$) or Cu-dominated ($RFO_{Sc} \leq 0.5$) as explained above. Geographical distributions of both cloud types reconstructed with the LCC categorization show a good match with their benchmark counterparts derived from CALIPSO-GOCCP/CASCCAD.

Furthermore, the two categorizations employing LCC and EIS threshold to discriminate Sc from Cu perform similarly well when applied to observational LCC data set (Text S1 in Supporting Information S1). There is indeed some ambiguity about the outcome of the two categorizations when employed in climate models. It should finally be noted that the categorization based on LCC threshold has been designed to distinguish Sc-dominated and Cu-dominated regimes in climate models that produce stratiform and cumuliform clouds. However, in the extreme case where a climate model would severely lack a particular cloud type this should be instantly apparent simply from inspecting geographical distribution of LCC because in regions over tropical and subtropical oceans Sc and Cu clouds are normally quite well geographically separated (Sc decks are typically found off the west coast of the continents, while Cu clouds are located further west over the open ocean). In such situation the exact evaluation of a missing cloud-type component is not of interest at all (model should be flagged as "highly erroneous…"), neither does one need a methodology to distinguish Sc and Cu (since the entire LCC can be attributed to the other, nonmissing cloud type).

We next proceed with a regime-oriented validation of CMIP6 models using the new Sc-Cu categorization (Section 3). Finally, Text S3 in Supporting Information S1 briefly compares the two categorizations in models.

3. Results and Discussion

3.1. Geographical Distributions of LCC and SW CRE

Figure 3 compares geographical distributions of low-cloud type climatology in subsidence areas of tropical oceans as simulated by the CMIP6 multimodel mean with CALIPSO-GOCCP observations. It is clear that latest climate models still struggle representing low cloudiness: on average, LCC in both Sc-regions and Cu-regions is strongly underestimated. The absolute bias reaches up to 30% in Sc-regions and up to 15% in Cu-regions. The CMIP6 models, on average, approximately capture geographical locations of the two regimes; namely stratocumulus decks off the west coast of the continents and shallow cumuli scattered further west over the open ocean. Nevertheless,

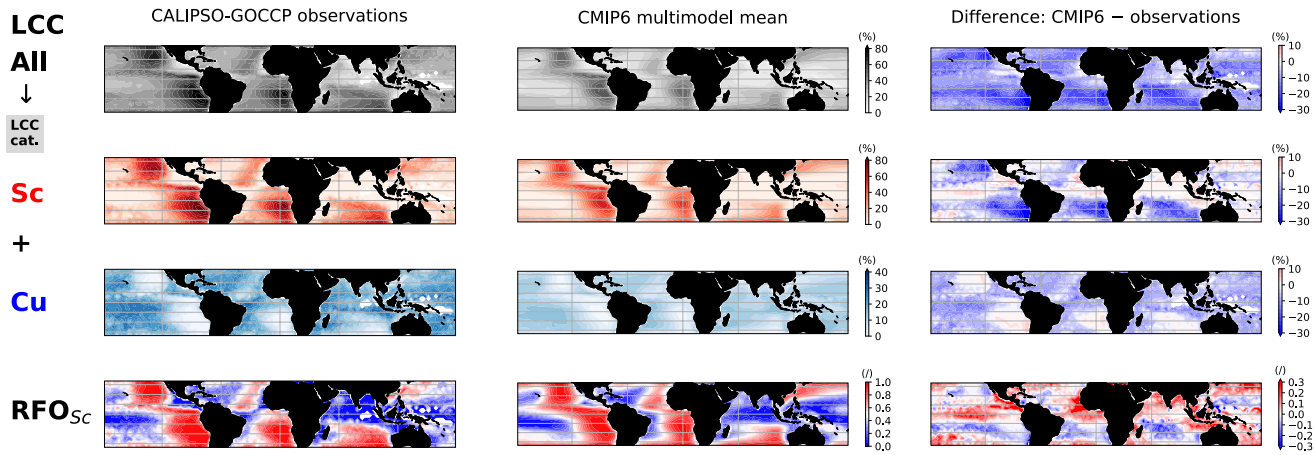

Figure 3. Geographical distributions of low-cloud cover (LCC) and individual cloud-type components constructed with the new categorization based on LCC threshold ("LCC cat.") in CALIPSO-GOCCP observations and CMIP6 multimodel mean. Bottom left and middle panels visualize RFO_{Sc}: regions shaded red are dominated by Sc, while regions shaded blue are dominated by Cu. The right column shows the corresponding difference between the CMIP6 multimodel mean and observations.

the inspection of the RFO of a given cloud type (e.g., RFO_{Sc}, Figure 3; bottom row) reveals areas where models simulate inadequate amount of a particular cloud type relative to low clouds compared to observations.

Figure 4 compares the corresponding SW CRE at TOA as simulated by the CMIP6 multimodel mean with CERES observations. The negative CRE field implies a cooling effect of low clouds on climate throughout tropical oceans, although the pattern is by far not uniform. Instead, it markedly reflects the presence of individual cloud regimes: it stems predominantly from Sc-regions, whereas it is smaller in Cu-regions. The CMIP6 models, on average, overestimate the magnitude of the observed negative CRE throughout the majority of tropical oceans. Figures S3 and S4 in Supporting Information S1 additionally show geographical distributions of LCC and CRE decomposed into Sc and Cu components as simulated by individual CMIP6 models.

Figure 5 evaluates spatial patterns of simulated LCC and SW CRE against observations with the aid of Taylor diagrams (Taylor, 2001). They concurrently display multiple metrics including normalized standard deviation and spatial correlation coefficient, whereby the anomalies are computed relative to the total mean (8-year period). It is apparent that CMIP6 models struggle capturing geographical distributions of LCC and CRE in both Sc-regions and Cu-regions of tropical oceans. Nevertheless, the correlation coefficient between the modeled and observed field is mostly higher in Sc-regions than in Cu-regions, which is evident for both LCC and CRE. Noteworthy, all models underestimate the observed variability of LCC in both Sc-regions and Cu-regions, with the exception of IPSL-CM6A-LR, which is in closest agreement with observations. The variability of simulated CRE generally shows a better match with observations, whereby normalized standard deviations commonly lie in the range between 0.8 and 1.2. The most noticeable outlier is MIROC6, which significantly overestimates

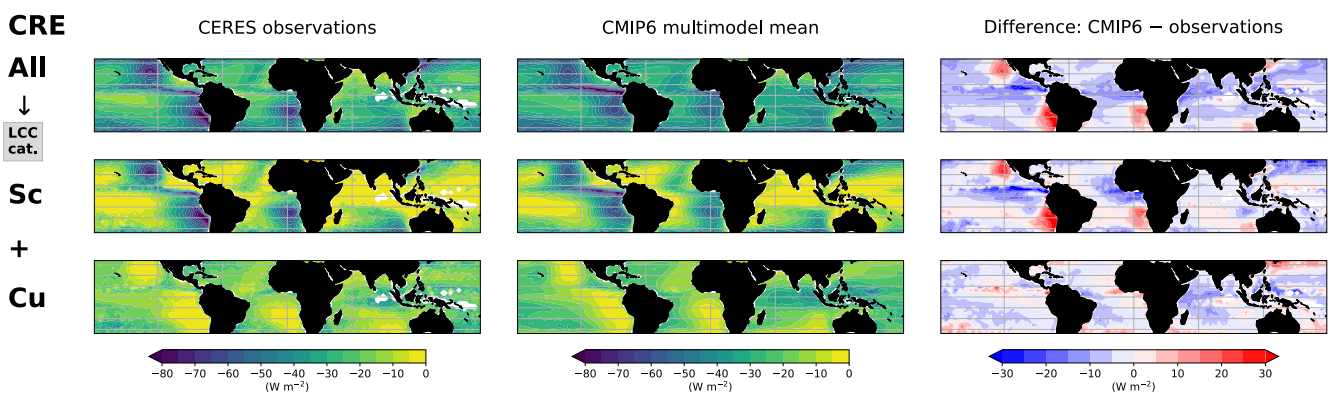

Figure 4. Geographical distributions of shortwave (SW) cloud-radiative effect (CRE) associated with each low-cloud regime in CERES observations and CMIP6 multimodel mean as well as the corresponding difference.

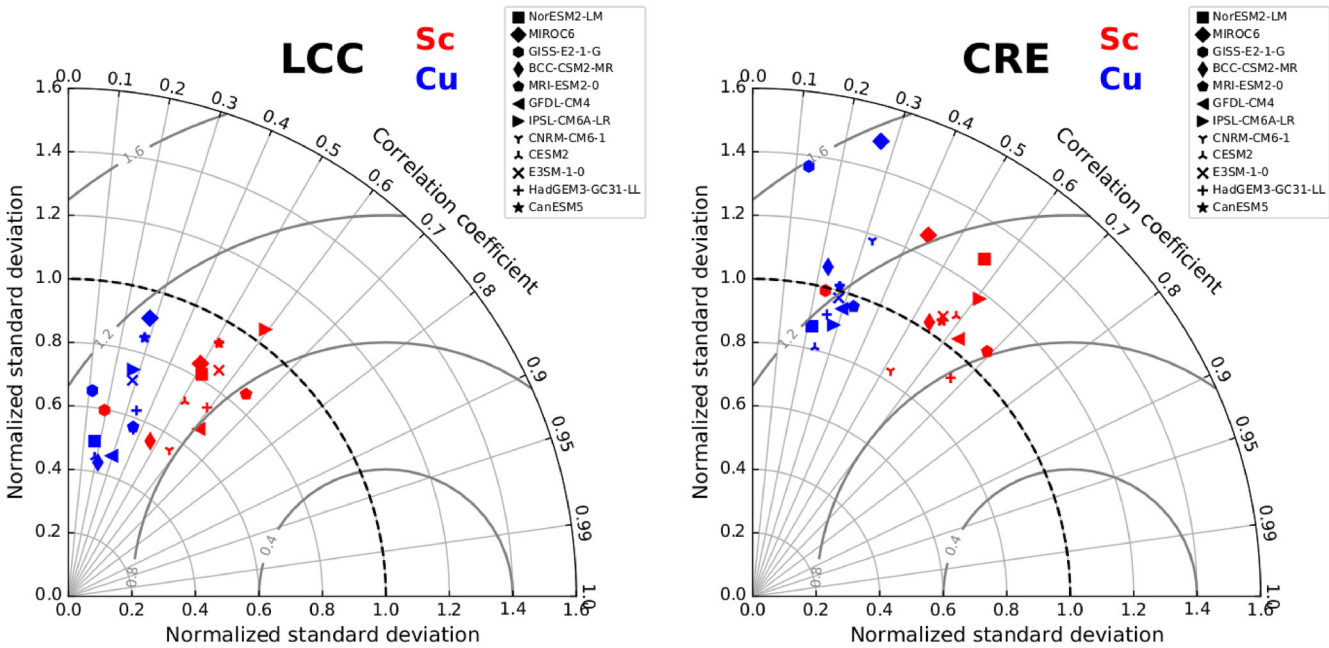

Figure 5. Taylor diagrams evaluating simulated low-cloud cover (LCC, left) and shortwave (SW) cloud-radiative effect (CRE, right) in Sc-regions and Cu-regions of tropical oceans.

the variability in both Sc-regions and Cu-regions. We identify further contrasting findings for Sc-regimes and Cu-regimes, whereby models mostly underestimate the observed variability of CRE in Cu-regions, whereas they tend to overestimate it in Sc-regions. These results exemplify there are other factors than LCC which profoundly affect the CRE bias.

3.2. Relationship Between LCC and SW CRE

Figure 6 (top) additionally displays PDFs of LCC in subsidence regimes of tropical oceans. The division of the latter into Sc and Cu components reveals that models simulate strongly biased LCC distribution within each of the two regimes. Specifically, the observed PDF peaks at ∼55% and 35% within Sc-regions and Cu-regions, respectively. Contrarily, model PDFs generally peak in the range between 20% and 40% in Sc-regions, whereas they predominantly peak between 5% and 30% in Cu-regions. We further note that the shape of simulated distributions is highly variable across the CMIP6 ensemble. The best-performing model is CanESM5, which fairly well captures the observed distributional peak and shape in Sc-regions and Cu-regions. Figure S5 in Supporting Information S1 offers an alternative visualization of these results.

We next investigate the relationship between LCC and SW CRE shown in Figure 6 (bottom). As expected the observed amount of reflected sunlight and hence the magnitude of negative SW CRE increases with increasing LCC. Noteworthy, the observed rate of increase is larger in Sc-regions than in Cu-regions. The observed nonlinear relationship between LCC and CRE is generally attributed to the increasing liquid water path with increasing LCC. Črnivec and Mayer (2019), as an illustration, investigated a shallow cumulus cloud field rising into stratocumulus and showed that both cloud cover and liquid water path concurrently increase with the simulation time.

Remarkably, all climate models overestimate the magnitude of negative SW CRE at a given LCC (by up to a factor of 2–3), which is evident throughout the entire range of LCC. This is consistent with findings of Nam et al. (2012) and implies that tropical low cloudiness in CMIP6 models remains too bright. The latter bias might be attributed to the inappropriate amount of averaged liquid water content, which is thus likely overestimated in models, as well as to other factors such as the parameterization of subgrid cloud variability and optical properties within the radiation scheme. CanESM5, which best captures the observed PDF of LCC, also exhibits the smallest bias in reflectance within both Sc-regions and Cu-regions. This is again in line with Nam et al. (2012), who pointed out that previous version of the Canadian GCM (CanAM4) incorporating

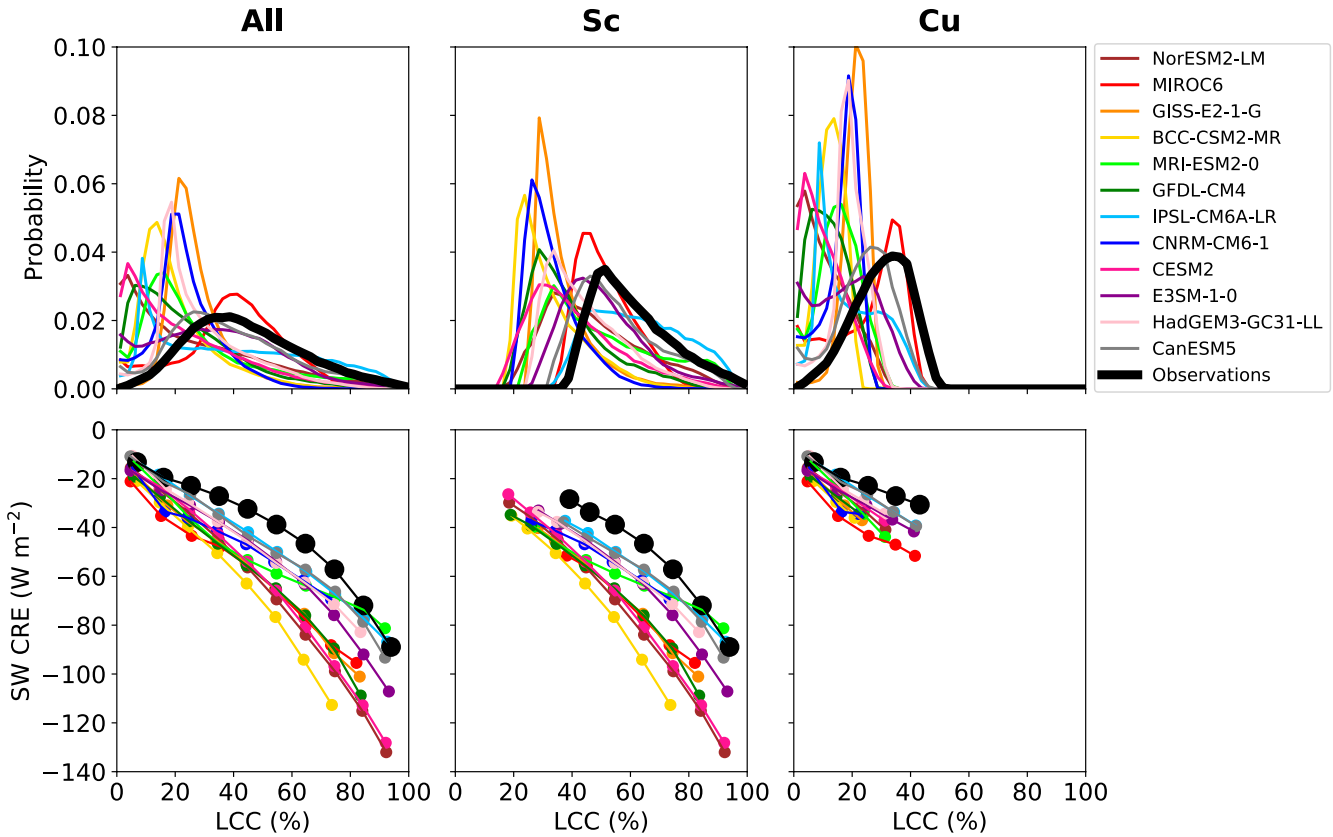

Figure 6. (top) Probability density functions of low-cloud cover (LCC) derived from observations and CMIP6 models in subsidence regimes of tropical oceans as well as separately in regions dominated by Sc and Cu. (bottom) The corresponding mean 2D histograms of LCC and shortwave (SW) cloud-radiative effect (CRE).

proficient parameterizations of subgrid cloud structure, was the model that minimized the overestimation of negative CRE among their set of analyzed CMIP5 members. BCC-CSM2-MR exhibits the largest CRE bias in Sc-regions, while in Cu-regions MIROC6 shows the largest discrepancy from observations. Studies exploiting a rich combination of measurements (Pincus et al., 1999) and large-eddy simulations (Črnivec & Mayer, 2020, 2021) reveal that Sc and Cu own distinct internal inhomogeneity characteristics. The contrasting heterogeneity of stratiform and convective clouds is not yet properly addressed within radiation schemes of current GCMs.

3.3. LCC Error Decomposition

Following the methodology of Schuddeboom and McDonald (2021), LCC associated with a particular cloud regime (r) can be expressed as

$$LCC_r = \overline{LCC}_r RFO_r \tag{3}$$

where the relative frequency of occurrence RFO_r is the rate at which given cloud regime occurs; and $\overline{LCC}_r$ represents the averaged cloud cover in regime r. This value is an average over the analyzed time period (8 years) and geographical regions (subsidence regimes over tropical oceans)—considering only grid boxes with a given cloud regime (either Sc or Cu in our case). To gain further insight into cloud biases, it is therefore convenient to decompose the total (overall) LCC model error of a specific cloud regime into various components as follows:

$$\delta LCC_r = \overline{LCC}_r^{obs} \Delta RFO_r + RFO_r^{obs} \Delta\overline{LCC}_r + \Delta\overline{LCC}_r \Delta RFO_r, \tag{4}$$

where $\Delta RFO_r = RFO_r^{\mathrm{mod}} - RFO_r^{obs}$ and $\Delta\overline{LCC}_r = \overline{LCC}_r^{\mathrm{mod}} - \overline{LCC}_r^{obs}$ capture the difference between model and observations. The three summands on the right-hand side of Equation 4 thus represent the error due to RFO, the error due to $\overline{LCC}$ and finally the error covariance term for a given cloud regime.

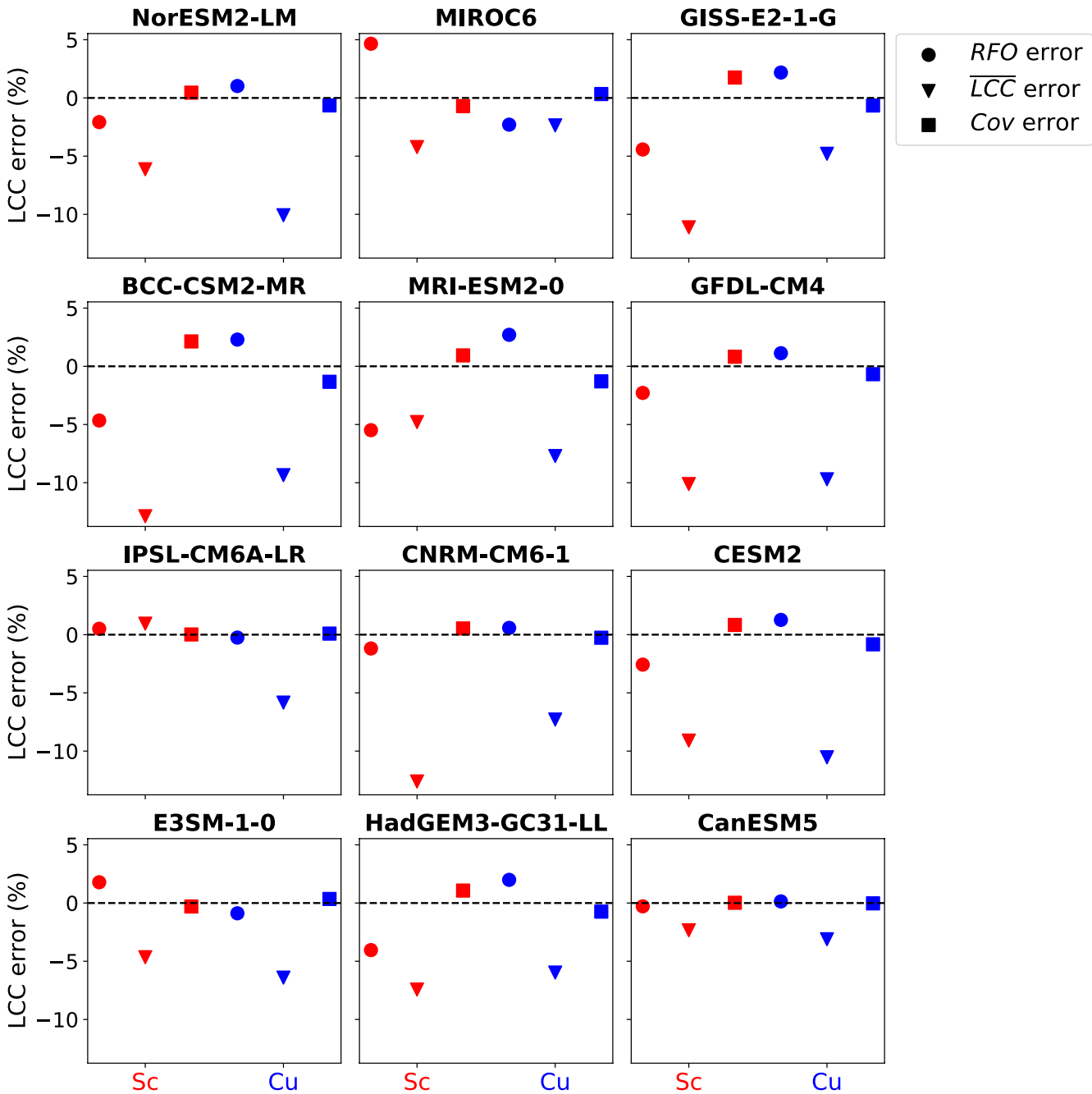

Figure 7. The low-cloud cover (LCC) error in subsidence areas of tropical oceans within Sc-regimes and Cu-regimes decomposed into errors stemming from RFO, $\overline{LCC}$ and covariance.

Figure 7 shows results of the overall LCC error decomposition according to Equation 4 within each of the two regimes (Sc, Cu), highlighting the complex nature of cloud biases among individual CMIP6 ensemble members. As anticipated the error covariance is generally small, therefore solely RFO and $\overline{LCC}$ errors are discussed in the following (see also Figure S6 in Supporting Information S1 for an alternative visualization of these two terms). Despite the fact that each model is subjected to its unique problematics, it is possible to draw several interesting conclusions.

The majority of models underestimate RFO of Sc, whereas they overestimate RFO of Cu (MRI-ESM2-0, GISS-E2-1-G, BCC-CSM2-MR, HadGEM3-GC31-LL, CESM2, GFDL-CM4, NorESM2-LM). The apparent outlier in this regard is MIROC6, which is in line with Williams and Tselioudis (2007), who showed that two previous versions of MIROC6 also considerably overestimated RFO of tropical stratocumulus and simultaneously

lacked shallow cumulus regime. There are a few models where the RFO error of both Sc and Cu is essentially zero (particularly CanESM5, IPSL-CM6A-LR, CNRM-CM6-1).

Remarkably, all models underestimate $\overline{LCC}$ within Sc-regions and Cu-regions (except IPSL-CM6A-LR, which however exhibits a negligible bias in Sc-region). Climate models that most strongly underestimate the mean Sc-cloud and Cu-cloud cover are BCC-CSM2-MR, CNRM-CM6-1, CESM2, GFDL-CM4, GISS-E2-1-G, and NorESM2-LM. In this subset of models, the $\overline{LCC}$ error exceeds 10% within either of the two regimes.

In the following discussion, we strive to provide some physical explanations for the aforementioned erroneous model behavior. Climate models frequently lack the inclusion of realistic moist processes within their planetary boundary layer (PBL) parameterizations (e.g., Cesana, Del Genio, Ackerman, et al., 2019), which affects their ability to sustain low clouds. As an illustration, the problem with the lack of Sc in GISS-E2-1-G was largely resolved in the next iteration of the GISS model, whereby the moist turbulence scheme of Bretherton and Park (2009) was newly implemented. One should also keep in mind that the underestimation of low cloudiness can partially stem from the shielding effect of high clouds, which might be overestimated in some models compared to observations, although we filter subsidence regimes to minimize this problem.

The best-performing model overall is CanESM5, which exhibits a zero bias in RFO and only a minor bias in $\overline{LCC}$ within both Sc-regions and Cu-regions. Recall that CanESM5 remarkably well matches the observed relationship between LCC and EIS (Figure 1). The IPSL-CM6A-LR model, moreover, has a negligible error in RFO and $\overline{LCC}$ within the Sc-regime. This is in agreement with Madeleine et al. (2020) who showed that the representation of low-level clouds (and their reflectance) in the IPSL-CM6 model has considerably improved compared to the previous model version participating in CMIP5. However, it is important to bear in mind that this apparent model improvement could be a consequence of model tuning, whereby the same observational data sets of LCC and CRE employed for the present model validation were used at IPSL to adjust free model parameters in an attempt to match the observed fields of clouds and radiation (Hourdin et al., 2019). Figure 1 indeed reveals that in the IPSL-CM6A-LR model LCC grows too strongly with EIS, which acts to offset the great lack of EIS in this model, so that the LCC error is eventually small.

As previously noted, contemporary climate models frequently employ the McICA radiation scheme of Pincus et al. (2003). The McICA stands for the Monte Carlo Independent Column Approximation, whereby simultaneous sampling of subgrid cloud state and spectral interval is exploited to speed up the computationally expensive calculation of full one-dimensional (ICA) radiative transfer. Whereas McICA is indeed a flexible and computationally efficient radiation scheme, it thus comes with a potentially detrimental side effect, which is commonly referred to as the McICA noise. We revisited a question, whether the McICA noise could lead to notable LCC biases in analyzed CMIP6 models containing the McICA algorithm (NorESM2-LM, GFDL-CM4, CESM2, E3SM-1-0, HadGEM3-GC31-LL, CanESM5). We found no evidence that the random noise generated by McICA is responsible for a notable underestimation of LCC (Figure S10 in Supporting Information S1), which is consistent with previous work examining older generation of GCMs (e.g., Barker et al., 2008). We also found no relationship between low-cloud biases in the present-day climate and equilibrium climate sensitivity (Figure S11 in Supporting Information S1).

4. Summary and Conclusions

Low clouds are ubiquitous in the tropics and intensely cool the Earth's climate, thus it is of tremendous importance to properly capture this effect in climate models. The overall objective of this study was to employ satellite observations to evaluate the representation of marine tropical stratocumulus and shallow cumulus and their impact on the Earth's radiation budget in a subset of latest climate models in the present-day climate. To that end, we first introduced a new approach to discriminate Sc from Cu based on a dynamic LCC threshold. The new Sc-Cu categorization proved to work well when applied to CALIPSO-GOCCP observations of LCC, validated against proper Sc and Cu components derived from the CASCCAD. Compared to the traditional approach for establishing low-cloud regimes utilizing a fixed threshold of EIS, the new Sc-Cu categorization is more reliable when analyzing climate models, since the latter systematically underestimate EIS implied by multiple reanalysis data sets and generally misrepresent the relationship between EIS and LCC.

Utilizing the newly proposed Sc-Cu categorization we then assessed models' fidelity to represent Sc and Cu together with their radiative effect at the TOA in the present-day climate. We thereby analyzed a suite of 12

state-of-the-art climate models stemming from various modeling centers participating in Phase 6 of CMIP. We restricted our analysis on subsidence regimes over tropical oceans, where low clouds are not obscured by midlevel and high-level clouds. We found that CMIP6 models underestimate the cloud cover in both Sc-dominated and Cu-dominated regions of tropical oceans. A more detailed inspection of cloud biases revealed that most climate models underestimate RFO of Sc and overestimate RFO of Cu. We further showed that tropical low cloudiness in CMIP6 models remains too bright.

The results of the present study are in line with a recent work by Konsta et al. (2022), who demonstrated that low-level marine tropical clouds in six CMIP6 models are too few and too bright, but also too compact and too homogeneous. The present study extends the results of Konsta et al. (2022) by evaluating 12 CMIP6 models and by discriminating stratocumulus and shallow cumulus regimes. All in all, these findings imply that contemporary climate models are still subjected to notable biases in clouds and radiation, which should fuel further climate model development.

Data Availability Statement

The CALIPSO-GOCCP data set (Chepfer et al., 2010b) was downloaded from the CFMIP-Obs website (http://climserv.ipsl.polytechnique.fr/cfmip-obs/Calipso_goccp.html). The CASCCAD data set (Cesana et al., 2019b) is available at NASA GISS website (https://data.giss.nasa.gov/clouds/casccad/). CERES-EBAF 4.0 data (Loeb et al., 2018b) were downloaded from the CERES website (https://ceres.larc.nasa.gov/data/#energy-balanced-and-filled-ebaf). ERA5 data (Hersbach et al., 2019) were downloaded from the Copernicus Climate Data Store (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-pressure-levels-monthly-means?tab=-form). NCEP-DOE R2 data (Kanamitsu et al., 2014) were downloaded from the NOAA ESRL Physical Sciences Division website (https://www.cpc.ncep.noaa.gov/products/wesley/reanalysis2/). MERRA-2 data (Gelaro et al., 2017b) were downloaded from the NASA Goddard Earth Sciences Data and Information Services Center (https://disc.gsfc.nasa.gov/datasets/M2IMNPASM_5.12.4/summary). The CMIP6 simulation output (Eyring et al., 2016a) was downloaded from ESGF (https://esgf-node.llnl.gov/search/cmip6/).

Acknowledgments
NC, GC, and RP were supported by the NOAA's Climate Program Office within the Modeling, Analysis, Predictions, and Projections (MAPP)'s Climate Sensitivity Task Force (Grant NA20OAR4310390). We would like to thank Andrew Ackerman for fruitful discussions and insightful comments, which improved the quality of this research. We furthermore thank Michael Puma for helpful discussions at the final stage of this work. We thank climate modeling centers and the World Climate Research Programme's Working Group on Coupled Modeling for providing data of CMIP simulations. Nina Črnivec would like to thank Andrew Gettelman, Neil Swart, Jason Cole, Seiji Yukimoto, Miho Sekiguchi, Olivier Boucher, Jean-Louis Dufresne, Chris Golaz, Wuyin Lin, Tongwen Wu, Øyvind Seland, David Paynter, and Jing Feng for helpful discussions about multiple aspects of models' radiation schemes including the usage of the McICA algorithm. Finally, we thank reviewers for their constructive comments.

References

Ackerman, A. S., & Toon, O. B. (1996). Unrealistic desiccation of marine stratocumulus clouds by enhanced solar absorption. *Nature*, *380*(6574), 512–515. https://doi.org/10.1038/380512a0

Ackerman, A. S., Toon, O. B., & Hobbs, P. V. (1993). Dissipation of marine stratiform clouds and collapse of the marine boundary layer due to the depletion of cloud condensation nuclei by clouds. *Science*, *262*(5131), 226–229. https://doi.org/10.1126/science.262.5131.226

Ackerman, A. S., Toon, O. B., Stevens, D. E., Heymsfield, A. J., Ramanathan, V., & Welton, E. J. (2000). Reduction of tropical cloudiness by soot. *Science*, *288*(5468), 1042–1047. https://doi.org/10.1126/science.288.5468.1042

Ackerman, A. S., van Zanten, M. C., Stevens, B., Savic-Jovcic, V., Bretherton, C. S., Chlond, A., et al. (2009). Large-eddy simulations of a drizzling, stratocumulus-topped marine boundary layer. *Monthly Weather Review*, *137*(3), 1083–1110. https://doi.org/10.1175/2008MWR2582.1

Andrews, M. B., Ridley, J. K., Wood, R. A., Andrews, T., Blockley, E. W., Booth, B., et al. (2020). Historical simulations with HadGEM3-GC3.1 for CMIP6. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001995. https://doi.org/10.1029/2019MS001995

Barker, H. W., Cole, J. N. S., Morcrette, J.-J., Pincus, R., Räisänen, P., von Salzen, K., & Vaillancourt, P. A. (2008). The Monte Carlo Independent Column Approximation: An assessment using several global atmospheric models. *Quarterly Journal of the Royal Meteorological Society*, *134*(635), 1463–1478. https://doi.org/10.1002/qj.303

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J., Klein, S. A., et al. (2011). COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023–1043. https://doi.org/10.1175/2011BAMS2856.1

Bony, S., & Dufresne, J.-L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. *Geophysical Research Letters*, *32*, L20806. https://doi.org/10.1029/2005GL023851

Bony, S., Stevens, B., Frierson, D., Jakob, C., Kageyama, M., Pincus, R., et al. (2015). Clouds, circulation and climate sensitivity. *Nature Geoscience*, *8*(4), 261–268. https://doi.org/10.1038/NGEO2398

Boucher, O., Servonnat, J., Albright, A. L., Aumont, O., Balkanski, Y., Bastrikov, V., et al. (2020). Presentation and evaluation of the IPSL-CM6A-LR climate model. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS002010. https://doi.org/10.1029/2019MS002010

Bretherton, C. S., & Park, S. (2009). A new moist turbulence parameterization in the community atmosphere model. *Journal of Climate*, *22*(12), 3422–3448. https://doi.org/10.1175/2008JCLI2556.1

Cesana, G. V., & 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*, L20803. https://doi.org/10.1029/2012GL053153

Cesana, G. V., 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: Atmospheres*, *121*, 5788–5808. https://doi.org/10.1002/2015JD024334

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

Cesana, G. V., Del Genio, A. D., Ackerman, A. S., Kelley, M., Elsaesser, G., Fridlind, A. M., et al. (2019). Evaluating models' response of tropical low clouds to SST forcings using CALIPSO observations. *Atmospheric Chemistry and Physics*, *19*(5), 2813–2832. https://doi.org/10.5194/acp-19-2813-2019

Cesana, G. V., Del Genio, A. D., & Chepfer, H. (2019a). The Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD). *Earth System Science Data*, *11*(4), 1745–1764. https://doi.org/10.5194/essd-11-1745-2019

Cesana, G. V., Del Genio, A. D., & Chepfer, H. (2019b). Cumulus And Stratocumulus CloudSat-CALIPSO Dataset (CASCCAD) [Dataset]. NASA. Retrieved from https://data.giss.nasa.gov/clouds/casccad/

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

Cesana, G. V., Waliser, D. E., Henderson, D., L'Ecuyer, T. S., Jiang, X., & Li, J.-L. (2019). The vertical structure of radiative heating rates: A multimodel evaluation using A-train satellite observations. *Journal of Climate*, *32*, 1573–1590. https://doi.org/10.1175/JCLI-D-17-0136.1

Chepfer, H., Bony, S., Winker, D., Cesana, G. V., Dufresne, J., Minnis, P., et al. (2010a). The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP). *Journal of Geophysical Research*, *115*, D00H16. https://doi.org/10.1029/2009JD012251

Chepfer, H., Bony, S., Winker, D., Cesana, G. V., Dufresne, J., Minnis, P., et al. (2010b). GCM-Oriented CALIPSO Cloud Product [Dataset]. CALIPSO-GOCCP. Retrieved from http://climserv.ipsl.polytechnique.fr/cfmip-obs/Calipso_goccp.html

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

Črnivec, N., & Mayer, B. (2019). Quantifying the bias of radiative heating rates in numerical weather prediction models for shallow cumulus clouds. *Atmospheric Chemistry and Physics*, *19*(12), 8083–8100. https://doi.org/10.5194/acp-19-8083-2019

Črnivec, N., & Mayer, B. (2020). The incorporation of the tripleclouds concept into the δ-Eddington two-stream radiation scheme: Solver characterization and its application to shallow cumulus clouds. *Atmospheric Chemistry and Physics*, *20*(17), 10733–10755. https://doi.org/10.5194/acp-20-10733-2020

Črnivec, N., & Mayer, B. (2021). Towards an improved treatment of cloud-radiation interaction in weather and climate models: Exploring the potential of the tripleclouds method for various cloud types using libRadtran 2.0.4. *Geoscientific Model Development*, *14*(6), 3663–3682. https://doi.org/10.5194/gmd-14-3663-2021

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

Dolinar, E. K., Dong, X., Xi, B., Jiang, J. H., & Su, H. (2015). Evaluation of CMIP5 simulated clouds and TOA radiation budgets using NASA satellite observations. *Climate Dynamics*, *44*(7–8), 2229–2247. https://doi.org/10.1007/s00382-014-2158-9

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., et al. (2016a). World climate Research Programme (WCRP) Coupled Model Intercomparison Project 6 (CMIP6) model output [Dataset]. Retrieved from https://esgf-node.llnl.gov/search/cmip6/

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016b). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017a). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 1631–1644. https://doi.org/10.1175/JCLI-D-16-0758.1

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017b). NASA Global Modeling and Assimilation Office (GMAO) MERRA-2 instM_3d_asm_Np: 3d, monthly mean, instantaneous, pressure-level, assimilation, assimilated meteorological fields V5.12.4 (M2IMNPASM) [Dataset]. NASA. https://doi.org/10.5067/2E096JV59PK7

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., et al. (2019). The DOE E3SM Coupled Model Version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089–2129. https://doi.org/10.1029/2018MS001603

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*, 3691–3727. https://doi.org/10.1029/2019MS001829

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2019). ERA5 monthly averaged data on pressure levels from 1940 to present [Dataset]. Copernicus Climate Change Service (C3S) Climate Data Store. https://doi.org/10.24381/cds.6860a573

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

Hourdin, F., Rio, C., Grandpeix, J.-Y., Madeleine, J.-B., Cheruy, F., Rochetin, N., et al. (2019). LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. *Journal of Advances in Modeling Earth Systems*, *12*, e2019MS001892. https://doi.org/10.1029/2019MS001892

Jiang, J. H., Su, H., Zhai, C., Perun, V. S., Del Genio, A., Nazarenko, L. S., et al. (2012). Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-train" satellite observations. *Journal of Geophysical Research*, *117*, D14105. https://doi.org/10.1029/2011JD017237

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., et al. (2014). National Centers for Environmental Prediction/DOE AMIP-II Reanalysis [Dataset]. NOAA. Retrieved from https://www.cpc.ncep.noaa.gov/products/wesley/reanalysis2/

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), 1631–1644. https://doi.org/10.1175/BAMS-83-11-1631

Karlsson, J., Svensson, G., & Rodhe, H. (2008). Cloud radiative forcing of subtropical low level clouds in global models. *Climate Dynamics*, *30*(7–8), 779–788. https://doi.org/10.1007/s00382-007-0322-1

Kelley, M., Schmidt, G. A., Nazarenko, L., 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*, e2019MS002025. https://doi.org/10.1029/2019MS002025

Klein, S. A., Hall, A., Norris, J. R., & Pincus, R. (2017). Low-cloud feedbacks from cloud-controlling factors: A review. *Surveys in Geophysics*, *38*(6), 1307–1329. https://doi.org/10.1007/s10712-017-9433-3

Konsta, D., Dufresne, J.-L., Chepfer, H., Vial, J., Koshiro, T., Kawai, H., et al. (2022). Low-level marine tropical clouds in six CMIP6 models are too few, too bright but also too compact and too homogeneous. *Geophysical Research Letters*, *49*, e2021GL097593. https://doi.org/10.1029/2021GL097593

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018a). 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

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., et al. (2018b). NASA Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) top-of-atmosphere (TOA) Edition-4.0 data product [Dataset]. NASA. Retrieved from https://ceres.larc.nasa.gov/data/#energy-balanced-and-filled-ebaf

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*, e2020MS002046. https://doi.org/10.1029/2020MS002046

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.-L., & Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in CMIP5 models. *Geophysical Research Letters*, *39*, L21801. https://doi.org/10.1029/2012GL053421

Pincus, R., Barker, H. W., & Morcrette, J. J. (2003). A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *Journal of Geophysical Research*, *108*(D13), 4376. https://doi.org/10.1029/2002JD003322

Pincus, R., Batstone, C. P., Hofmann, R. J. P., Taylor, K. E., & Glecker, P. J. (2008). Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *Journal of Geophysical Research*, *113*, D14209. https://doi.org/10.1029/2007JD009334

Pincus, R., McFarlane, S. A., & Klein, S. A. (1999). Albedo bias and the horizontal variability of clouds in subtropical marine boundary layers: Observations from ships and satellites. *Journal of Geophysical Research*, *104*(D6), 6183–6191. https://doi.org/10.1029/1998JD200125

Raisanen, P., Barker, H. W., Khairoutdinov, M. F., Li, J., & Randall, D. A. (2004). Stochastic generation of subgrid-scale cloudy columns for large-scale models. *Quarterly Journal of the Royal Meteorological Society*, *130*(601), 2047–2067. https://doi.org/10.1256/qj.03.99

Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society*, *84*(11), 1547–1564. https://doi.org/10.1175/Bams-84-11-1547

Schneider, T., Teixeira, J., Bretherton, C. S., Brient, F., Pressel, K. G., Schär, C., & Siebesma, A. P. (2017). Climate goals and computing the future of clouds. *Nature Climate Change*, *7*(1), 3–5. https://doi.org/10.1038/nclimate3190

Schuddeboom, A. J., & McDonald, A. J. (2021). The Southern Ocean radiative bias, cloud compensating errors, and equilibrium climate sensitivity in CMIP6 models. *Journal of Geophysical Research: Atmospheres*, *126*, e2021JD035310. https://doi.org/10.1029/2021JD035310

Seland, Ø., Bentsen, M., Olivié, D., Toniazzo, T., Gjermundsen, A., Graff, L. S., et al. (2020). Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geoscientific Model Development*, *13*(12), 6165–6200. https://doi.org/10.5194/gmd-13-6165-2020

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*, e2019RG000678. https://doi.org/10.1029/2019RG000678

Stephens, G., Winker, D., Pelon, J., Trepte, C., Vane, D., Yuhas, C., et al. (2018). CloudSat and CALIPSO within the A-train: Ten years of actively observing the Earth system. *Bulletin of the American Meteorological Society*, *99*(3), 569–581. https://doi.org/10.1175/BAMS-D-16-0324.1

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al. (2002). The CloudSat mission and the A-train: A new dimension of space-based observations of clouds and precipitation. *Bulletin of the American Meteorological Society*, *83*(12), 1771–1790. https://doi.org/10.1175/BAMS-83-12-1771

Stevens, B., Ackerman, A. S., Albrecht, B. A., Brown, A. R., Chlond, A., Cuxart, J., et al. (2001). Simulations of trade wind cumuli under a strong inversion. *Journal of the Atmospheric Sciences*, *58*(14), 1870–1891. https://doi.org/10.1175/1520-0469(2001)058⟨1870:SOTWCU⟩2.0.CO;2

Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, *340*(6136), 1053–1054. https://doi.org/10.1126/science.1237554

Stubenrauch, C. J., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G. V., Chepfer, H., et al. (2013). Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel. *Bulletin of the American Meteorological Society*, *94*(7), 1031–1049. https://doi.org/10.1175/BAMS-D-12-00117.1

Swales, D. J., Pincus, R., & Bodas-Salcedo, A. (2018). The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2. *Geoscientific Model Development*, *11*(1), 77–81. https://doi.org/10.5194/gmd-11-77-2018

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

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Teixeira, J., Cardoso, S., Bonazzola, M., Cole, J., Del Genio, A., DeMott, C., et al. (2011). Tropical and sub-tropical cloud transitions in weather and climate prediction models: The GCSS/WGNE Pacific Cross-section Intercomparison (GPCI). *Journal of Climate*, *24*(20), 5223–5256. https://doi.org/10.1175/2011JCLI3672.1

Tselioudis, G., Rossow, W. B., Jakob, C., Remillard, J., Tropf, D., & Zhang, Y. (2021). Evaluation of clouds, radiation, and precipitation in CMIP6 models using global weather states derived from ISCCP-H cloud property data. *Journal of Climate*, *34*, 7311–7324. https://doi.org/10.1175/JCLI-D-21-0076.1

Voldoire, A., Saint-Martin, D., Sénési, S., Decharme, B., Alias, A., Chevallier, M., et al. (2019). Evaluation of CMIP6 DECK experiments with CNRM-CM6-1. *Journal of Advances in Modeling Earth Systems*, *11*, 2177–2213. https://doi.org/10.1029/2019MS001683

Weare, B. (2004). A comparison of AMIP II model cloud layer properties with ISCCP D2 estimates. *Climate Dynamics*, *22*(2–3), 281–292. https://doi.org/10.1007/s00382-003-0374-9

Webb, M., Senior, C., Bony, S., & Morcrette, J. J. (2001). Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dynamics*, *17*(12), 905–922. https://doi.org/10.1007/s003820100157

Williams, K. D., & Tselioudis, G. (2007). GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response. *Climate Dynamics*, *29*(2–3), 231–250. https://doi.org/10.1007/s00382-007-0232-2

Winker, D. M., Pelon, J., Coakley, J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., et al. (2010). The CALIPSO mission: A global 3D view of aerosols and clouds. *Bulletin of the American Meteorological Society*, *91*(9), 1211–1230. https://doi.org/10.1175/2010BAMS3009.1

Wood, R., & Bretherton, C. S. (2006). On the relationship between stratiform low cloud cover and lower-tropospheric stability. *Journal of Climate*, *19*(24), 6425–6432. https://doi.org/10.1175/JCLI3988.1

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