Assessment of CMIP6 Cloud Fraction and Comparison with Satellite Observations

P. Preetham Vignesh¹, Jonathan H. Jiang², P. Kishore³, Hui Su², Timothy Smay¹, Nicholas Brighton¹, and Isabella Velicogna²,³

¹Grade 12, University High School, Irvine, CA, USA, ²Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA, ³Department of Earth System Science, University of California, Irvine, CA, USA

Abstract The seasonal and regional variations of cloud fractions are compared across two generations of global climate model ensembles, specifically, the Coupled Model Intercomparison Project-5 (CMIP5) and CMIP6, through the historical period in terms of skills and multimodel agreement. We find a wider spread of historical cloud fraction changes in the CMIP6 than was simulated by the CMIP5. The global mean cloud fractions of CMIP6 increased by about 4.5% from the CMIP5, which attributed to greater changes in the northern hemisphere than in the southern hemisphere. The CMIP6 cloud fractions in recent years are validated with the CALIPSO_CLOUSAT observations to understand the cloud fraction uncertainties in CMIP6 models. The CMIP6 ensemble mean of cloud fractions compares well with the observations with a mean difference of 0.5% in lower altitudes. The CMIP6 cloud fractions are higher than the observations at higher latitudes in both hemispheres in the upper troposphere, and the biases vary from one model to another. The spatial difference between the ensemble and observations is further revealed over the tropics: where the model displays a 3% higher bias. In addition, we observed a significant trend occurring in the northern hemisphere since the mid-20th century using calculations of cloud fraction trends based on the robust regression technique. Finally, we reduce the differences between the model and observations by applying a simple regression technique. The results exemplify that the model and modified observations compare well, with the root mean square value decreased by nearly 28%, and the correlation increased significantly.

1. Introduction

Clouds play an important role in the climate system by affecting the hydrological cycle and radiative balances (Mulmenstadt et al., 2015; Stubenrauch et al., 2013). Variations of clouds over horizontal and vertical extents have the potential to either amplify or reduce the Earth’s atmosphere–Earth system and the global hydrological cycle through their radiative balance effects (Stephens et al., 2012; Trenberth et al., 2009). The Coupled Model Intercomparison Project (CMIP), launched in 1995, generated several atmospheric parameters concerning present and future climates (Meehl et al., 2005). Tremendous improvements have been made from the CMIP Phase 1 (CMIP1, Lambert & Boer, 2001) to the CMIP Phase 6 (CMIP6), with lots of changes in the physical and cloud parameterization schemes and with higher model resolutions. Spatial and vertical distributions of clouds overlap some of the uncertainties in determining the cooling/heating profile by the radiative and evaporative/precipitable processes (e.g., Stephens & Webster, 1984; Stephens et al., 2002). Amongst a large variety of observations, these observations are not regular and continuous over the globe. Satellite-based cloud products are the best way to validate global cloud parameters, and they provide a continuous and comprehensive vertical distribution of global cloud datasets. Satellite measurements are the only source that we can use to retrieve horizontal and vertical cloud information on a global scale. We can use this valuable information on clouds to evaluate different models.

In the past, the model datasets are evaluated using satellite-based datasets for the integrated column water path (e.g., Eliasson et al., 2011), ice clouds (e.g., Jiang et al., 2012; Li et al., 2012), and cloud fraction (e.g., Cesana & Waliser, 2016; Marchand et al., 2009; Sun et al., 2015). Several studies indicate improvements, variability, or uncertainty in the historical models from the CMIP5 model compared to the CMIP3 model but only for specific climate variables or regions. (e.g., Cesana & Waliser, 2016; Monerie et al., 2012; Rupp et al., 2013). However, some studies reported that no improvements were made when CMIP5 simulations are compared to CMIP3 simulations (e.g., Knutti & Sedlacek, 2012). The past studies have indicated that
multimodel spread is one of the major sources of uncertainty (Hawkins & Sutton, 2011). Therefore, besides comparing cloud parameters with satellite observations, the intercomparison between the two-generation individual and multimodel ensemble is necessary and is important. This applies especially for future projections and for many more modifications in the cloud schemes and their related processes to be carefully examined in the new generation models. Simulations of more than a hundred years would be needed to assess the model performance, as well as significant changes and improvements to the CMIP6 model compared to CMIP5 simulations in terms of cloud parameters. Meehl et al. (2014) also mentioned that the next generation climate models were necessary for developing strategies for the assessment of CMIP6.

Active and passive sensors can detect clouds; these observed changes concern a subset of all clouds, but passive sensors do not detect all types of clouds. Lidars and radars are active sensors that can generally detect most clouds, even small cumulus clouds, and their vertical distribution in high resolutions with less influence of the surface (Di Michele et al., 2012; Miller et al., 2014; Wang et al., 2014; Wylie et al., 2007). There have been many studies using satellite measurements to reduce the model biases over different regions using different methods. Satellites provide the most comprehensive and continuous record for global cloud products. Satellite observations of cloud parameters are playing an important role in evaluating models on how well clouds are represented and will help to improve the next generation global circulation models (GCMs). However, we did not find any single comprehensive study on the intercomparison of the CMIP6 models with the earlier GCMs of CMIP5 and validation with satellite observations for spatio-temporal, regional, and seasonal assessment of the cloud frequency.

In this study, we investigate the performance of the total area cloud fraction (TCF) and layered cloud fractions (LCF) of the sixth phase of the CMIP6 with that of the previous phase, the fifth generation of global coupled models (CMIP5). Through an ensemble of 10 CMIP5 and CMIP6 GCM models’ outputs and satellite observations, we quantitatively assess the current climate simulations and provide the information useful for model improvements. This comparison enables us to assess the progress of simulating clouds in the new generation of CMIP6 models. The CMIP6 TCF and LCF are compared with CALIPSO_CLOUDSAT observations. TCF and LCF cloud parameters are particularly interesting because they largely impact the radiation budget and the climate in model simulations. We also attempt to identify TCF at different latitude bands in the northern hemisphere (NH) and the southern hemisphere (SH), and the LCF biases at different altitudes are also investigated. Finally, we attempt to adjust the spatial TCF using a simple regression technique.

The structure of this paper is organized as follows. Section 2 uses the CMIP5, CMIP6 models, and satellite TCF and LCF data are shown, and the methodology of analyzing this data is described as well. Section 3 compares the global and latitudinal distribution of CMIP5 and CMIP6 TCF and validates the CMIP6 TCF and LCF with satellite observations. Major differences along with associated similarities of cloud fractions are also discussed. In addition to this, we attempt a performance test of the adjustment using CMIP6 and satellite cloud fractions with currently available datasets. The last section, section 4, summarizes our findings and discusses them.

2. Data and Methods

2.1. Historical CMIP5 and CMIP6 Models

The climate models with numerous experiments have each provided a valuable repository of a wide range of climate model data over 20 different institutions (Taylor et al., 2012). In this study, we focus on the TCF and the LCF from the fifth and sixth phases of the CMIP (O'Neill et al., 2016; Taylor et al., 2012), along with satellite cloud datasets. The CMIP6 historical period begins in 1850 and extends to the end of 2014, and CMIP5 covers the period from 1850 to 2005. The sixth phase of CMIP provides a variety of mechanisms, a variety of common experiments, the Diagnostic Evaluation and Characterization of Klima experiments (Eyring et al., 2016). CMIP6 formulated a centralized activity of parametrizations and encompassed several individual CMIP6-endorsed MIPS, Diagnostic Evaluation and Characterization of Klima, and CMIP6 historical simulations, which helps climate science researchers. Table 1 provides the CMIP5 and CMIP6 model names along with their horizontal resolutions.
2.2. CALIPSO and CloudSat

The satellite can provide vertical profiles of clouds over global cloud profiles (either direct or retrieved) from physical and empirical methods. The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite is one of the satellites in the National Aeronautics and Space Administration A-Train of Earth-observing remote sensing satellites (Stephens et al., 2002). It was launched simultaneously with the CloudSat satellite and flew in a 705 km orbit with a 1330 local time (equatorial crossing time). On average, CALIPSO is 15 s behind CloudSat, and the period of the study pointed to 0.3° and 0.16° from the nadir and along the ground track, respectively. The CALIPSO and CloudSat (CALCLD hereafter) satellite cloud observations are merged (day and night time measurements) and validated (Kay & Gettelman, 2009; Mace et al., 2009). For our analysis, we collected the CALCLD TCF and LCF datasets: 2 × 2 longitude-latitude monthly grids from July 2006 to February 2011. The CALIPSO and CloudSat missions have provided the three-dimensional distribution of clouds, and these global cloud datasets are the best way to evaluate climate model simulations. The vertical height profiles of layered cloud fraction grids are sorted into three vertical layers: low (height <3.36 km), middle (3.36 km < height < 6.72 km), and high (height >6.72 km) (Rossow et al., 1996; Cesana & Waliser, 2016). The uncertainty of the CALCLD TCF and LCF data is estimated at ~5% (Marchand et al., 2009). Additional uncertainties due to satellite sampling might exist. TCF and LCF from the combined CloudSat/CALIPSO simulator CF may be a more fair choice for comparison with observations because instrument sensitivity is one of the most important factors in cloud detection that determines TCF and LCF. Su et al. (2013) showed that the LCF from the CALIPSO simulators are generally lower than those from the standard retrieval products. At the time of our analysis, the CloudSat/CALIPSO simulator data from most CMIP6 models were not available; therefore, we shall bear in mind the caveats when interpreting the model-observation discrepancies.

2.3. Methods

The historical simulations of the 145-year period (1861–2005) for CMIP5 and of the 154-year period (1861–2014) for CMIP6 are used. Different historical CMIP5 and CMIP6 model outputs are available on different spatial resolutions, as shown in Table 1. These historic models outputs and observations are bilinearly interpolated to a common spatial scale of 1 × 1 longitude-latitude resolution. Using the model and the observational cloud datasets, we have calculated seasonal and yearly means at each grid. We estimate the spatial relative differences in percentages between the two model datasets of the CMIP5 and CMIP6 at each grid point in this study, which is given by diff = 100.*(CMIP6-4-CMIP5)/CMIP6. For the purpose of long-term trend analysis, we utilized a robust regression technique. This robust regression technique was introduced by Holland and Welsch (1977); this technique is based on iteratively reweighted least squares regression, an improvement to least squares regression and is less affected by outliers (O’Leary, 1990). The Mann-Kendall test is used for finding the statistical significance of the trend at each grid point, and it confirms the existence of a negative or positive trend for a given confidence level (Mann, 1945).

| Table 1 | List of IPCC Global Climate Models of CMIP6 and CMIP5 and Grid Resolutions Used in this Study |
|---------|--------------------------------------------------|
| CMIP5  | Model name | CF Grid | LCF Grid | CMIP6 | Model name | CF Grid |
| 1   | BCC-CSM2-MR | Y 320 x 160 | Y 320 x 160 x 46 | bcc-csm1-1-m | N 320 x 160 |
| 2   | BCC-ESM 1 | Y 128 x 64 | Y 128 x 64 x 26 | bcc-csm1-1 | Y 128 x 64 |
| 3   | CESM2 | Y 288 x 192 | Y 288 x 192 x 32 | ccsM4 | Y 288 x 192 |
| 4   | CNRM-CM6-1 | Y 256 x 128 | Y 256 x 128 x 91 | CNRM-CM5 | Y 256 x 128 |
| 5   | CNRM-ESM 2-1 | Y 256 x 128 | Y 256 x 128 x 91 | CNRM-CM5-2 | Y 256 x 128 |
| 6   | GFDL-CM4 | Y 256 x 180 | N 256 x 180 x – | GFDL-CM3 | Y 256 x 90 |
| 7   | GISS-E2-1-G | Y 144 x 90 | Y 144 x 90 x 29 | GISS-E2-H | Y 144 x 90 |
| 8   | IPSL-CM6A-LR | Y 144 x 143 | Y 144 x 143 x 79 | IPSL-CM5A-LR | Y 96 x 96 |
| 9   | MIROC6 | Y 256 x 128 | Y 256 x 128 x 81 | MIROC5 | Y 256 x 128 |
| 10  | MRI-ESM 2-0 | N 320 x 160 | N 320 x 160 x 80 | MRI-ESM 1 | N 320 x 160 |

Abbreviations: CMIP = Coupled Model Intercomparison Project; LCF = layered cloud fraction.
3. Results and Discussion

3.1. Comparison of CMIP5 and CMIP6 TCF

This section discusses the comparison between CMIP5 and CMIP6’s TCF along with spatial differences. Figure 1 shows the annual mean of CMIP6 (left column), CMIP5 (middle column), and the differences (right column) of cloud fraction for the period from 1861 to 2005. A large difference is noticed in the CESM2 model, and moderate differences are noticed in the CNRM-CM6-1 and CNRM-ESM 2-1 models. The two models’ (GFDL-CM3 and GISS-E2-H) cloud fractions of CMIP5 are higher than CMIP6’s cloud fractions. The annual average differences in cloud fractions are found in between -8 and 10%. The differences between CMIP6 and CMIP5 are somewhat smaller at the tropical region (20°S–20°N). However, CMIP6 cloud fractions are still very large at the extratropics and midlatitude regions. The spatial biases vary with each model, and larger differences (>8%) are present in the CESM2 model when compared to the other models over all the regions.

Before proceeding to compare the CMIP6 models’ cloud fraction with satellite observations, we will first discuss the CMIP5 and CMIP6 seasonal spatial structures and their relative differences, annual mean differences in cloud fraction, bias, and latitudinal structures. Figure 2 displays the seasonal spatial patterns of CMIP6 and CMIP5 cloud fractions along with the relative difference of cloud fractions during the period from 1861 to 2005. The seasonal spatial structures look similar in all seasons, with moderate magnitude variations observed over high-latitude regions in the NH and SH. In order to understand more precisely where the CMIP6 cloud fractions are spread, we examined the relative difference between the ensemble of CMIP6 and CMIP5 models at different seasons as estimated and shown in Figures 2i–2l. The relative differences of cloud fractions are not that significant during each season. Southern and northern high-latitude regions indicate that the most evident CMIP6 cloud fractions are nearly 20% higher than that of the CMIP5 model values over the spring and summer seasons. CMIP5 cloud fractions are higher at tropical latitudes (25°S–25°N) but are highly concentrated over the ocean region than over the land region in all the seasons. From our analysis, we find a consistency increase in the CMIP6 cloud fractions over the period (1861–2005); however, consistency is lacking from model to model over the short-term time period.

Figure 3 redisplay the annual mean histograms, the bias in percentage, and the zonal mean, which are estimated using CMIP5 and CMIP6 historical ensemble means of TCF during the period from 1861 to 2005. It is observed from CESM2 that the CMIP6 annual mean values are nearly 25% more than the CMIP5 cloud fraction mean values (Figure 3a). The GFDL-CM4 cloud fractions of CMIP5 and CMIP6’s mean values of cloud fractions are a little higher than the multimodel mean values. All CMIP6 models show slightly higher cloud fraction values than those of the CMIP5 models. The horizontal blue and red dotted lines in Figure 3a are for multimodel mean values of total cloud fractions of CMIP5 and CMIP6 ensembles, respectively. The total cloud fraction CMIP6 multimodel mean values are nearly 4.5% higher than the CMIP5 multimodel mean values. Figure 3b shows that the annual mean cloud fraction bias bins averaged between the 80°S and 80°N latitude region of the CMIP6 and CMIP5 models from 1861 to 2005. The six models display positive bias, which indicates that CMIP6 has a larger average TCF than that of CMIP5. The percentage of bias lies between -4 and 7%, and this shows that there are not many bias variations observed between the two models.

Based on the monthly spatial distributions of the percentage occurrence of total cloud fractions presented in Figure 1, we estimated the zonal means, which were determined for 1° latitude bins. Figure 3c represents the climatological zonal mean values of TCF of CMIP5 and CMIP6 ensembles. Shaded regions for the higher CMIP6 (light green color) and lower CMIP5 (light blue color) ensembles represent one standard deviation projected by the eight models for each latitude bin. Figure 3c displays some different characteristics when compared with Figure 1. The CMIP6 zonal mean value has been shifted 5% higher than that of the CMIP5 cloud fraction zonal mean value. The mean value of TCF in the NH is little higher than that of the SH in the tropics in both the CMIP5 and CMIP6. However, the zonal mean value of TCF in the NH high-latitudes (>60°N) is lower than that of TCF in the SH high-latitudes (>60°S). This is because the monthly variations of TCF in the SH high-latitudes are stronger than in the NH high-latitudes. The maximum TCF mean differences occur between 55°S and 65°S in the NH, whereas in the SH, the TCF mean difference is nearly constant between 10°S and 65°S. In the CMIP5 monthly TCF standard deviations are larger than in the CMIP6. Results suggest little change in the cloud frequency’s central tendency or the variability of historical skills across the two-generation models. The changes across the two generations of individual models of CMIP5 and CMIP6 with regards to the TCF mean and standard deviation statistics are displayed in Figure 3c.
Figure 1. Spatial multi-annual mean cloud fractions from individual CMIP6, CMIP5, and differences of cloud fractions averaged during the period from 1861 to 2005. CMIP, Coupled Model Intercomparison Project.
Tables 2 and 3, respectively. The CMIP6 ensemble mean values are a little higher when compared to the CMIP5 ensemble in all latitudinal regions. Intermodel mean variations are more prominent in CMIP5 than in CMIP6’s individual models. The maximum values are observed over middle- and high-latitudes (50°N–80°N and 50°S–80°S) in the NH and SH in both generations of models. These mean values indicate that the CMIP6’s TCF is greater than that of the CMIP5 model values for the same parameter. The global mean expresses that the CMIP6 TCF values are slightly greater than that of the CMIP5. This exemplifies that individual models of the CMIP6 have been improved by correcting errors in cloud fractions, which is attributed to the lower standard deviations in TCF values in CMIP6 models. These statistical analyses expressed that the CMIP6 TCF values displayed better skills when compared to the CMIP5 TCF values.

Figure 2. Spatial distribution of cloud fraction of winter, spring, summer, and fall seasons depicted by the ensemble CMIP6 (left column), CMIP5 (middle column), and the relative differences (100*(CMIP6 – CMIP5)/CMIP5) of cloud fractions during the years 1861–2005 (right column). CMIP, Coupled Model Intercomparison Project.
Figure 3. (a) Annual global (80oS–80oN) total area cloud fractions mean of CMIP5 and CMIP6, (b) annual mean differences between CMIP6 and CMIP5, and (c) zonal mean plot of CMIP5 and CMIP6 cloud fractions; the shaded area corresponds to the standard deviations during the period 1861 to 2005. CMIP, Coupled Model Intercomparison Project.

Figure 4 shows the annual mean TCF anomalies relative to the 1861–1900 mean of TCF over the different latitude bands for CMIP5 and CMIP6 multimodel ensemble. The cloud fraction annual anomalies for six latitude regions are shown: the tropics (20oS–20oN), the NH and SH latitude bands are (20oN–50oN), (50oN–80oN), (20oS–50oS), (50oS–80oS), and the global range of latitudes (80oS–80oN), respectively. The TCF latitudinal mean changes time series is shown by the red color (CMIP5) and the blue color (CMIP6), as well as the locally weighted regression technique “LOESS” for the entire time series with a smoothing parameter $\alpha = 0.75$, and the resultant curves, which are over-plotted, are shown in Figure 4. Before the 1880s, all annual anomalies display a similar pattern in both ensemble models. Both the CMIP5 and CMIP6 ensemble anomalies display a similar pattern in the 20oS–20oN and the 50oS–80oS latitude bands,

| CMIP6 | 20oS–20oN | 20oN–50oS | 50oS–80oS | 20oS–50oS | 50oS–80oS | 80oS–80oN |
|-------|-----------|-----------|-----------|-----------|-----------|-----------|
| BCC-CSM2-MR | 66.411 (0.497) | 54.534 (0.364) | 68.777 (0.447) | 56.170 (0.425) | 69.831 (0.507) | 63.338 (0.236) |
| BCC-ESM 1 | 61.731 (0.327) | 51.278 (0.342) | 67.089 (0.337) | 55.791 (0.489) | 62.123 (0.723) | 62.189 (0.189) |
| CESM2 | 65.947 (0.549) | 59.357 (0.620) | 86.841 (0.471) | 68.990 (0.475) | 78.939 (0.190) | 73.837 (0.223) |
| CNRM-CM6-1 | 65.424 (0.332) | 55.238 (0.423) | 73.454 (0.509) | 58.703 (0.311) | 78.939 (0.306) | 66.356 (0.215) |
| CNRM-ESM 2-1 | 65.525 (0.294) | 54.972 (0.388) | 73.187 (0.422) | 58.392 (0.350) | 78.029 (0.335) | 66.049 (0.167) |
| GFDL-CM4 | 63.811 (0.416) | 59.137 (0.415) | 84.499 (0.289) | 63.443 (0.342) | 85.031 (0.259) | 70.875 (0.184) |
| GISS-E2-1-G | 58.558 (0.564) | 52.163 (1.036) | 75.348 (0.920) | 51.373 (0.341) | 71.876 (0.256) | 61.995 (0.469) |
| IPSL-CM6A-LR | 51.282 (0.606) | 54.146 (0.646) | 80.313 (0.522) | 65.773 (0.430) | 85.983 (0.250) | 66.751 (0.308) |
| MIROC6 | 60.792 (0.308) | 51.351 (0.450) | 62.686 (0.481) | 59.943 (0.310) | 69.782 (0.381) | 60.850 (0.214) |
| Ensemble | 62.209 (0.278) | 54.565 (0.489) | 75.342 (0.529) | 59.782 (0.412) | 75.645 (0.318) | 65.781 (0.448) |

Abbreviation: CMIP = Coupled Model Intercomparison Project.
Table 3
Total Area Cloud Fractions (%) Mean and Standard Deviations of Each Individual and Ensemble CMIP5 Model for Different Latitude Bands During 1861 to 2005

| Model                  | 20°S–20°N | 20°N–50°N | 50°N–80°N | 20°S–50°S | 50°S–80°S | 80°S–80°N |
|------------------------|-----------|-----------|-----------|-----------|-----------|-----------|
| bcc-csm1-1             | 57.182 (0.342) | 52.649 (0.421) | 67.935 (0.548) | 51.893 (0.307) | 69.931 (0.314) | 59.937 (0.370) |
| CCSM4                  | 47.531 (0.378) | 39.989 (0.354) | 46.869 (0.802) | 45.657 (0.290) | 53.037 (0.605) | 46.647 (0.438) |
| CNRM-CM5-2             | 57.390 (0.395) | 48.927 (0.477) | 64.321 (0.549) | 54.800 (0.445) | 73.600 (0.341) | 59.652 (0.491) |
| CNRM-CM5               | 57.384 (0.397) | 48.967 (1.473) | 64.861 (0.508) | 54.811 (0.415) | 73.788 (0.326) | 59.800 (0.454) |
| GFDL-CM3               | 69.974 (0.417) | 64.826 (1.531) | 87.985 (0.329) | 67.820 (0.449) | 86.950 (0.331) | 75.452 (0.394) |
| GISS-E2-H              | 63.641 (0.428) | 55.681 (0.559) | 74.181 (0.299) | 51.909 (0.306) | 70.907 (0.388) | 63.651 (0.330) |
| IPSL-CM5A-LR           | 49.500 (0.723) | 53.600 (0.913) | 68.839 (0.601) | 60.410 (0.667) | 79.645 (0.264) | 61.739 (0.439) |
| MIROC5                 | 58.467 (0.359) | 48.466 (0.867) | 57.613 (0.665) | 56.286 (0.331) | 64.217 (0.457) | 57.038 (0.411) |
| Ensemble               | 57.875 (0.412) | 51.750 (0.387) | 66.516 (0.561) | 56.750 (0.349) | 71.289 (0.289) | 60.625 (0.398) |

Abbreviation: CMIP = Coupled Model Intercomparison Project.

Figure 4. Annual mean changes averaged over different latitudinal bands (20°S–20°N, 20°N–50°N, 50°N–80°N, 20°S–50°S, and 50°S–80°S), and global (80°S–80°N) under CMIP6 (1861–2014), and CMIP5 (1861–2005) compared to the reference period (1861–1900).
but both bands start to undergo a change in the overall trend around 1950. After 1950, both ensemble models continue to decrease and increase at the 20°S–20°N and 50°S–80°S latitude bands, respectively. The 50°S–80°S band annual anomalies rapidly increase from 1950 in both models. In both models, the TCF significantly decreases in the 20°S–50°S latitude region with more pronounced changes in the CMIP6 model. The global (80°S–80°N) TCF anomalies of CMIP5 and CMIP6 are quite different due to the changes that appeared in the NH middle- and high-latitude regions of the CMIP6 ensemble. Figure 4 displays that the NH middle- and high-latitude TCF annual CMIP5 and CMIP6 anomalies are different. The CMIP6 ensemble's TCF rapidly increases in the middle- and high-latitudes, and drastic changes are seen in the NH when compared to the SH. Meanwhile, SH CMIP6 TCF anomalies display somewhat similar behavior to that of the CMIP5 ensemble annual anomalies. The CMIP6 interannual variability of TCF is higher than that of the CMIP5 TCF, particularly in the NH middle- and high-latitudes. Noteworthy changes have been noticed in the CMIP5 and CMIP6's TCF from the 1960s, and this may be due to the anthropogenic forcings that occurred after the year 1960 (Acharya & Sreekesh, 2013; Moorthy et al., 2013).

3.2. Comparison of Total Cloud Frequency in CMIP6 and Satellite Observations

In this section we evaluate the eight CMIP6 model TCF datasets with the observed CALIPSO_CLOUDSAT (CALCLD) satellite gridded TCF spatial variability and correlation during the period from July 2006 to February 2011. The annual mean bias and correlation values are evaluated at each grid point during the period of observations. These results are depicted in Figure 5. Bias spatial maps are simple but are more informative, since they display the spatial variability of the two different sets of mean bias. The negative bias is observed in most of the regions in the tropics, extra tropics, and the middle-latitude NH and SH regions. In some regions, a positive bias is observed over high-latitudes 60°N–80°N and the 60°S–80°S regions. The positive bias indicates that the CMIP6 TCF is higher than that of the CALCLD datasets. The average positive bias is about 8.5% over high-latitudes, and the average negative bias is about 4.3% between the 60°S and 60°N latitude region. The spatial correlation between the two datasets shows good correlation in most of the regions. However, some negative correlation is also found in some small regions. The average correlation coefficient between the CMIP6 ensemble and CALCLD is nearly 0.65. The CMIP6 ensemble shows an increasing trend, but the CMIP5 shows a decreasing trend in TCF values. These results indicate that the historical TCF differences between CMIP5 and CMIP6 are noticed more in the NH than the SH.

The Taylor diagram (Gleckler et al., 2008; Taylor, 2001) can provide a quantitative statistical summary of how well the spatial or temporal patterns match between the ensemble simulations and observations in terms of their correlations coefficients, their root mean square difference, and the simulated-to-observed ratio of their variances. If the standard deviation of the model is the same as that of the observations, then the radius will be 1. Figure 6 shows the correlation coefficient and the standard deviation values used to produce the Taylor diagram over the six different regions (20°S–20°N, 20°N–50°N, 50°N–80°N, 20°S–50°S, 50°S–80°S, and 80°S–80°N) between individual CMIP6 model and CALCLD observations. The standard deviations are a little lower in the 20°S–50°S region compared to the other latitude regions. The red filled circle indicates the ensemble mean of CMIP6. In the tropical (20°S–20°N) region the spread is large, and the ensemble mean is apart from the models. The standard deviation and correlation can provide an overall evaluation of the models due to the clustering of points. Some models show a lower correlation, that is, less than 0.4 in 20°S–20°N and 20°S–50°S latitude regions; meanwhile, other models show lower standard deviations. The high correlation (i.e., greater than 0.6) between the model and observations indicates that the seasonal cycles of the model-simulated cycles are phased correctly. Each model value is closer to the reference data value and lower standard deviations imply that the CMIP6 model is performing relatively well, which might improve parameterizations and the added time span matches with the observation period.

3.3. Altitude and Latitude Structure of CMIP6 and CALCLD Observations

In this section we describe the relative zonal mean difference of each model with CALCLD observations at each altitude level during the observational period. We estimate the relative difference between the model and the observations, which is given by diff = 100 × (CMIP6 − CALCLD) /CMIP6 LCF at each grid point. The latitude-altitude cross-section difference between CALCLD and an individual mean of CMIP6, the multimodel mean, the multimodel standard deviation, and the CALIPSO_CLOUDSAT mean are shown in Figure 7. The zonal mean of the high-cloud frequency (HCF) time series is shown in the upper part of
each subplot (in black line). The bias is similar and consistent in all models, and the CMIP6 models display negative bias in the lower heights over the tropics, extra tropics, and the midlatitude regions. A large relative difference is noticed in the upper height levels between the 15- and 20-km altitude level, which represents that the model underestimates LCF between the 50°S and 50°N latitude region. The majority of CMIP6 model cloud frequencies show a larger CALCLD at higher altitudes (12 and 18 km), over higher latitudes in both hemispheres, and it indicates that the model overestimates the LCF compared to the observations. The positive relative difference ranges between 1 and 2% and varies from model to model. Comparing the multimodel and satellite cloud frequency, the satellite observational values appear to be a little higher in the upper troposphere level between 50°S and 50°N. The multimodel HCF maximum zonal mean cloud frequencies are observed over the tropics, but the HCF values are 10% less than the satellite observations. These differences may be due to models using different cloud schemes and different forcings, which affect each model simulation. The effect of each forcing may be model-dependent. The systematic cloud biases are common to most models, and these depend on the cloud types and specific physical parameterizations (Zhang et al., 2005). Jiang et al. (2012) concluded that intermodel spread and errors in cloud water content are greatest in the upper troposphere. Lock (2009) mentioned that some models considered the

Figure 5. Contour plot of (a) bias and (b) correlation between CMIP6 and CALCLD cloud fractions over the period from July 2006 to February 2011. CMIP, Coupled Model Intercomparison Project.
planetary boundary layer scheme that models simulated relatively better in low-frequency clouds than in high-frequency clouds.

This section discusses the comparison between the CMIP6 ensemble and CALCLD for four different layered cloud frequencies (low, middle, high, and total columns). Figure 8 shows the longitude-latitude structures of relative mean bias, ensemble standard deviation, mean, and CALCLD mean for four different layered frequencies. The different types of cloud frequencies are discussed in section 2. The maximum cloud frequencies are observed at middle and high latitudes in both hemispheres at the total column in both the ensemble and model datasets. The total column cloud fractions are highly concentrated in the SH (40oS–60oS) than in the NH (40oN–60oN) in both the model and the observations. Another maximum CF is observed over the intertropical convergence zone in both observations; however, the CMIP6 ensemble CF intensity is lower than in the satellite observations. The intertropical convergence zone, located between 5oN and 10oN latitudes over the Pacific and Atlantic regions, plays a major role in the atmospheric energy balance (Waliser and Gautier, 1993). The lowest CFs of about ~15% are observed in the middle layer and are followed by the high and low-layered structures. The greatest standard deviations are observed in the total column of the CMIP6 ensemble when taking into account the other three cloud frequency structures. More significant standard deviations are observed in the SH than in the NH, and moderate standard deviations are found in tropical regions, especially in total column cloud frequencies. The relative bias between the CMIP6 ensemble and satellite observations is shown in the first column of Figure 8. The most significant positive bias between the CMIP6 and satellite total column CF datasets occurs in the high-latitude region in both hemispheres.

Figure 6. Taylor diagram for mean cloud fraction between individual CMIP6 model and CALCLD during the period from July 2006 to February 2011 for different latitude bands (a) 20oS–20oN, (b) 20oN–50oN, (c) 50oS–80oS, (d) 20oS–50oS, (e) 50oS–80oS, and (f) 80oS–80oN. Each solid filled circle (blue) indicates the individual model, and the red filled circle corresponds to the CMIP6 Ensemble. CMIP, Coupled Model Intercomparison Project.
Note that the blue (red) color denotes CMIP6 lower (higher). The spatially integrated latitudinal root mean square error (RMSE) and the correlation between each of the CMIP6 and CALCLD’s low-, middle-, and high-layered cloud fractions of the three NH latitudes and global (80oS–80oN) statistics are highlighted as depicted in Table 4. The model (MRI-ESM 2-0) RMSE is larger over the 50oN–80oN latitude region. From Table 4, the ensemble layered CF correlations are in good agreement with the observations, and the maximum correlation (0.7) is observed in the global LCF.

To deduce more information on the altitude dependence of the CF, we show the CF mean profiles along with standard deviations over five different latitude bands using CMIP6 and CALCLD observations in Figure 9. The altitude CF profiles gross features are displayed more prominently than those seen in the contour plots (Figure 8). The horizontal bars (red and blue) on the plot represent month-to-month variability with respect to the mean value. The vertical height profiles of CMIP5 and observations’ CFs look similar; however, some differences are observed in both the model and observations in all latitude bands. The maximum amplitudes

Figure 7. Relative zonal mean difference of individual CMIP6 model and CALCLD latitude-altitude structures, and ensemble CMIP6 model mean, standard deviation, and CALCLD mean during the period from July 2006 to February 2011. The time series of zonal mean of high-layered cloud frequency is shown at the top of each plot. The horizontal black dashed line indicates the low- and middle-level clouds. CMIP = Coupled Model Intercomparison Project; LCF = layered cloud fraction.
are observed around the lower troposphere heights in the tropical region (20°S–20°N) at about ~17% in both the model and the observation, whereas in the midlatitudes (20°N–50°N and 20°S–50°S) in both hemispheres the maximum CF is concentrated at around 10 km at about 15%. CALCLD values show a larger strength in CF values than the CMIP6 model values below 7 km. There are larger standard deviations in the CALCLD CF values than the model CF values in the lower heights of both hemispheres. Su et al. (2013) evaluated that model cloud vertical structures of deep convective clouds do not reach as high in altitude as compared to those of the observations. In most of the altitudes, the standard deviations for CALCLD are found to be higher than for the model. In both hemispheres, the cloud frequencies decrease with height and reach a minimum of 2% at 12.5 km over the high-latitude regions (50°N–80°N and 50°S–80°S). Comparison of the model and the observed cloud fractions suggests that there is a reasonable agreement between the two sets of values.

### 3.4. Historical CMIP6 Cloud Fraction Spatial Trends

The spatial distribution of seasonal and annual trends in TCF is simulated by the CMIP6 ensemble during the period 1901–2014 for different periods and is depicted in Figure 10. The trends are estimated using robust regression analysis. Min et al. (2011) and Basha et al. (2017) mentioned that anthropogenic forcings are
Table 4
Northern Hemisphere Latitudinal RMSE and Correlation Values of Low-, Middle-, and High-Cloud Frequencies (%) Between Each Ensemble CMIP6 Model and CALCLD Observations During the Period from July 2006 to February 2011

| Model/Observation | 20oS–20oN: RMSE (Correlation) | 20oN–50oS: RMSE (Correlation) | 50oN–80oS: RMSE (Correlation) | 80oS–80oN: RMSE (Correlation) |
|-------------------|-------------------------------|-------------------------------|-------------------------------|-------------------------------|
|                   | Low LCF | Mid LCF | High LCF | Low LCF | Mid LCF | High LCF | Low LCF | Mid LCF | High LCF | Low LCF | Mid LCF | High LCF | Low LCF | Mid LCF | High LCF | Low LCF | Mid LCF | High LCF |
| BCC-CSM2-1MR      | —       | 0.21(0.5) | 0.41(0.7) | 1.28(0.3) | 1.23(0.8) | 0.47(0.6) | —       | 1.46(0.7) | 1.3(0.3) | 1.32(0.3) | 0.45(0.4) | 0.43(0.2) |    |
| BCC-ESM 1         | 0.23(0.7) | 0.27(0.5) | 0.58(0.6) | 1.01(0.9) | 1.73(0.8) | 1.31(0.5) | 1.87(0.8) | 1.89(0.8) | 1.03(0.3) | 0.67(0.8) | 0.37(0.7) | 0.81(0.3) | 0.37(0.7) | | | | | | | |
| CESM2              | 0.87(0.3) | 1.23(0.4) | 1.12(0.3) | 1.54(0.4) | 1.87(0.3) | 1.21(0.4) | 2.13(0.3) | 2.23(0.3) | 2.54(0.2) | 1.87(0.3) | 1.34(0.3) | 2.12(0.3) | 0.81(0.3) | | | | | | | |
| CNRM-CM6-1        | 0.65(0.4) | 1.12(0.4) | 1.01(0.5) | 0.89(0.5) | 1.43(0.5) | 1.34(0.6) | 1.34(0.5) | 1.45(0.6) | 1.76(0.4) | 1.34(0.5) | 0.56(0.6) | 1.56(0.4) | 0.81(0.3) | 0.37(0.7) | | | | | | | |
| CNRM-ESM 2-1      | 0.68(0.4) | 1.34(0.5) | 1.10(0.5) | 1.04(0.4) | 1.54(0.5) | 1.89(0.6) | 1.34(0.6) | 1.34(0.5) | 1.89(0.5) | 1.29(0.4) | 1.09(0.5) | 1.89(0.4) | 0.81(0.3) | 0.37(0.7) | | | | | | | |
| GISS-E2-1-G       | 0.20(0.3) | 0.21(0.3) | 0.39(0.6) | 0.53(0.9) | 0.84(0.8) | 0.31(0.6) | 1.41(0.9) | 1.74(0.8) | 1.02(0.3) | 0.35(0.8) | 0.42(0.5) | 0.17(0.3) | 0.81(0.3) | | | | | | | |
| IPSL-CM6A         | 0.67(0.3) | 0.89(0.4) | 1.18(0.6) | 0.34(0.6) | 1.29(0.7) | 0.97(0.6) | 0.67(0.7) | 0.87(0.7) | 0.88(0.4) | 0.98(0.6) | 0.78(0.5) | 0.45(0.5) | 0.81(0.3) | | | | | | | |
| MIROC6            | 2.77(0.2) | 0.25(0.3) | 0.21(0.3) | —        | 0.65(0.8) | 0.29(0.7) | —        | 1.51(0.3) | 0.78(0.2) | 2.54(0.2) | 0.19(0.3) | 0.17(0.2) | 0.81(0.3) | 0.37(0.7) | | | | | | | |
| MRI-ESM 2-0       | 1.01(0.5) | 0.38(0.4) | 0.32(0.2) | 4.4(0.7)  | 1.16(0.8) | 0.79(0.4) | 9.7(0.7)  | 1.37(0.8) | 1.18(0.3) | 5.3(0.8)  | 1.34(0.5) | 0.33(0.3) | 0.81(0.3) | | | | | | | |
| Ensemble          | 1.12(0.4) | 0.76(0.5) | 0.69(0.6) | 1.89(0.6) | 1.39(0.6) | 0.89(0.5) | 3.11(0.5) | 1.49(0.6) | 1.89(0.5) | 1.94(0.7) | 1.29(0.6) | 0.98(0.6) | 0.81(0.3) | 0.37(0.7) | | | | | | | |

Abbreviations: CMIP = Coupled Model Intercomparison Project; LCF, layered cloud fraction; RMSE, root mean square error.

Figure 9. Mean vertical height profiles of cloud fractions (a) 20oS–20oN, (b) 20oN–50oS, (c) 50oN–80oS, (d) 20oS–50oS, (e) 50oS–80oS, and (f) 80oS–80oN of CMIP6 and CALCLD observations during the period from July 2006 to February 2011. CMIP = Coupled Model Intercomparison Project.
increased from the second half of the 20th century. We also found increasing/decreasing TCF trends from the second half of the century (Figure 3). So, we have separated the TCF into 1901–1960, 1961–2014, and the entire period (1901–2014). A 95% significance of identifying positive and negative trends is assessed in each grid over the globe using Student’s t-test. We considered four seasons for estimating the trends: winter (December, January, and February), spring (March, April, and May), summer (June, July, and August), and fall (September, October, and November). An increasing positive trend is observed over winter, fall, and in the annual NH land region during the period 1901 to 1960. However, a negative TCF trend is more pronounced in the 30°N–60°N latitude region during 1961–2014, and maximum negative

Figure 10. Geographic distribution of cloud fraction trends for annual and seasons depicted by the ensemble CMIP6 model data during the years 1901–1960 (left column), 1961–2014 (middle column), and 1901–2014 (right column). CMIP = Coupled Model Intercomparison Project.
trends are observed over Europe, North America, and the Atlantic Ocean in the summer and fall. During recent decades, 1960 to 2014, there is an increasing TCF trend over the Indian Ocean, India, the Northern part of South America, and the Pacific Ocean during all the seasons. The changes of TCF trends worldwide may be because anthropogenic forcings are steadily increased since the middle of the 20th century (Fischer & Knutti, 2015; Min et al., 2011 and Sachindra et al., 2016). Wang et al. (2016) pointed out that the changes in anthropogenic emissions in the 20th century and large anthropogenic emission increases are located more in the NH than in the SH. A strong positive trend is observed over Antarctica during all the seasons: the maximum in winter (~2.75%/dec) and fall (~3.2%/dec) seasons in the second half of the 20th century. An important point to be noted is that the TCF trends are decreasing and that this is observed in the second half

Figure 11. Spatial structures of the (a) slope and (b) intercept, defined by the equation (1). CMIP = Coupled Model Intercomparison Project.
compared to the first half of the 20th century, especially over the NH. In annual trends, a positive (negative) TCF is observed in several parts overland (Ocean) during the 1901–2014 period.

3.5. Preparation of Modified CALCLD

When comparing satellite measurements to the CMIP6 historical simulation, the spatial comparison maps often show that simulations tend to be high or low and that the biases vary by location, season, and variables used. To establish a relationship between model and observations, one class of simple and statistical technique is given by Yatagai et al. (2014). The modified CALCLD (m-CALCLD) monthly CF is reproduced by the following equation:

\[ R_{\text{model}} = a \cdot R_{\text{CALCLD}} + b, \]

where \( R_{\text{CALCLD}} \) and \( R_{\text{model}} \) are monthly CALIPSO_CLOUDSAT and CMIP6 ensemble’s CF, respectively. The regression coefficients \( a \) and \( b \) are the slope and intercept, respectively. Here, \( a \) is the ratio of CALCLD to the CMIP6 historical model. Figure 10 shows the slope and intercept of the model and observations. Two rectangular boxes (black color) are drawn by considering the appropriate region, where the slope

Figure 12. Time series (a) correlation coefficient, (b) root mean square error (RMSE), between the CMIP6 and CALCLD (blue) and between the CMIP6 and m-CALCLD (green). Cloud frequency spatial structures for (c) ensemble CMIP6, (d) CALCLD, and (e) m-CALCLD observations during the period from July 2007 to February 2011 for the selected region (left side box) shown in Figure 10. CMIP = Coupled Model Intercomparison Project; TCF = total area cloud fraction.
and intercept's variations between the model and observations are the greatest; these boxes cover the eastern and western longitude regions of Figure 10. The region inside the rectangle box is considered for modification of the CALCLD TCF datasets using equation (1). The monthly time series of correlation, RMSE, spatial maps of the model, CALCLD, and m-CALCLD CFs are depicted in Figures 12 and 13. The monthly time series correlation and RMSE between CMIP6 and CALCLD (m-CALCLD) were estimated over the rectangle box (Figure 11). The RMSE values decreased by nearly 35% of the actual RMSE values (Figure 12b), and the correlation coefficient is much better when m-CALCLD is compared with CALCLD, and this indicates that the adjustment values are more comparable with CMIP6's CF values. The spatial pattern of m-CALCLD displays a very good performance when compared with the ensemble model's CFs, since the correlation between the model and m-CALCLD is 0.90, and the RMSE is 7.84. The spatial CFs of m-CALCLD perform better in most of the regions with the model, indicating the improved performance on spatial and temporal scales. The monthly time series and correlation coefficients of the rectangle box (right side of Figure 11) are shown in Figure 13. The spatial maps of CMIP6 and m-CALCLD TCFs express similar behavior, since the correlation increased from 0.63 to 0.84, and RMSE decreased from

![Figure 13. Time series (a) correlation coefficient, (b) root mean square error, between the CMIP6 and CALCLD (blue) and between the CMIP6 and m-CALCLD (green). Cloud frequency spatial structures for (c) ensemble CMIP6, (d) CALCLD, and (e) m-CALCLD observations during the period from July 2007 to February 2011 for the selected region (right side box) shown in Figure 10. CMIP = Coupled Model Intercomparison Project; TCF = total area cloud fraction.](image-url)
significantly decreases in the 20oS–50oS latitude regions with more pronounced changes in the CMIP6 model. The global (80oS–80oN) TCF anomalies CMIP5 and CMIP6 are quite different due to changes appearing in the NH middle- and high-latitude regions of the CMIP6 ensemble. The cloud fractions of CMIP6 are compared with the CALCLD observations during the period 2006–2011. The standard error is less than 0.5 in all latitude bands, and the correlation is approximately 0.67 when comparing the CMIP6 ensemble and the CALCLD TCF in between the 80oS and 80oN band.

Latitudinal changes are estimated relative to 1861–1900 from global (CMIP5 and CMIP6) models for the historical changes at six different latitude regions. Both of the ensemble model anomalies are quite different especially over the NH than the SH. Both model anomalies' differences are small during the first half and significantly increased or decreased in the second half of the 20th century. In both models, the TCF significantly decreases in the 20oS–50oS latitude regions with more pronounced changes in the CMIP6 model. The statistical analysis (Tables 2 and 3) indicates that CMIP5 TCF values are somewhat larger than the CMIP5 ensemble values, which is likely due to the improvement in cloud parameterization schemes that have improved the TCFs. In general, all models cannot accurately simulate the temporal TCF and LCF characteristics in all latitude regions. Similarly, Lauer and Hamilton (2013) compared CMIP5 and CMIP3, and they reported that CMIP5 models perform slightly better than CMIP3 models, especially in cloud climatologies in some regions, but intermodel differences are still large in the CMIP5 simulations. The CMIP6 ensemble's cloud fraction mean differences are observed to be nearly 4.5% higher than the CMIP5 ensemble's mean differences, and the correlation coefficient is nearly 0.77.

In addition, we compared the LCF of individual CMIP6 models with the CALCLD, and we observe that the zonal mean’s relative difference is greater at the upper troposphere altitudes (12–18 km) in each model over the 50oS–50oN latitude region, and this indicates that the models underestimated LCF when compared to the observations. Tsushima et al. (2013) found that the systematic low bias in the cloud fraction in the tropics could be underestimated in the models. The percentage of relative mean difference between the CALCLD and individual CMIP6 models at most of the individual heights vary between –5 and 2% in the 2 to 19-km altitude levels. Cesana and Waliser (2016)) compared the LCF from CMIP5 and CALIPSO-GOCCP datasets, and they found the maximum differences above 12 km, yielding results similar to our study. However, they found that the CMIP5 models overestimated LCF when compared to the LCF of observations in the regions above 12 km. The different resolutions are a partial cause of these differences. The cloud model simulations developed at different centers may have changed cloud parameterizations at each version of the CMIP models, and this may be the cause of these differences. The CMIP6 ensemble clearly captures the cloud distribution from the ground to the upper troposphere region.

The annual and seasonal TCF trends were estimated using the robust regression technique at each grid point for three different time periods (1901–1960, 1961–2014, and 1901–2014), and from this analysis, we considered the trends above 95% significance level. The significance analysis (Student's t-test) helps to identify where the TCF trends are either increasing or decreasing. It is worth mentioning here that the significant increasing or decreasing trends are observed in 1961–2014 compared to 1901–1960. The increasing trend can be seen in all seasons over the land regions except during the summer season over 1901 to 2014.
Over the past decade, a number of studies have attempted to reduce the uncertainty as well as achieving better agreement over different regions in models and observations using different techniques. We have made an attempt to minimize the uncertainties as well as agreement between the CMIP6 and m-CALCLD spatial maps. The correlation coefficient increased from 0.736 to 0.902; this is highly significant when taking into consideration that the RMSE decreased from 12.2 to 7.8% (Figures 11 and 12), which indicates that a strong association is present between both the model and the observations. This study will be helpful to the modeling groups regarding the enhancement of cloud parameterization and development of upcoming GCM’s models. In future studies, we will collect more datasets with longer term observations, enabling us to have a better understanding of the cloud characteristics.

Data Statement

All the climate model data used for this research can be downloaded from the PCMDI (Program for Climate Model Diagnosis & Intercomparison) website at https://pcmdi.llnl.gov/CMIP6/. The CloudSat/CALIPSO data can be downloaded from http://www.cloudsat.cira.colostate.edu/. For additional questions regarding the data sharing, please contact the corresponding author at jonathan.h.jiang@jpl.nasa.gov.

References

Acharya, P., & Sreekesh, S. (2013). Seasonal variability in aerosol optical depth over India: A spatio-temporal analysis using the MODIS aerosol product. *International Journal of Remote Sensing*, 34(13), 4832–4849. https://doi.org/10.1080/01431161.2013.782114

Basha, G., Kishore, P., Ratnam, M. V., Jayaraman, A., Kouchak, A. A., & Taha, B. (2017). Historical and projected surface temperature over India during the 20th and 21st century. *Scientific Reports*, 7(1), 2987. https://doi.org/10.1038/s41598-017-02130-3

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, 10,358–10,546. https://doi.org/10.1002/2016GL070515

Di Michele, S., Ahlgrim, M., Forbes, R., Kulie, M., Bennartz, R., Janiskova, M., & Bauer, P. (2012). Interpreting an evaluation of the ECMWF global model with CloudSat observations: Ambiguities due to radar reflectivity forward operator uncertainties. *Quarterly Journal of the Royal Meteorological Society*, 138, 2047–2065.

Eliasson, S., Buehler, S. A., Milz, M., Erikson, P., & John, V. O. (2011). Assessing observed and modelled spatial distributions of ice water path using satellite data. *Atmospheric Chemistry and Physics*, 11, 375–391. https://doi.org/10.5194/acp-11-375-2011

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

Fischer, E. M., & Knutti, R. (2015). Anthropogenic contribution to global occurrence of heavy precipitation and high-temperature extremes. *Nature Climate Change*, 5, 560–564. https://doi.org/10.1038/nclimate2617

Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research, 113*, D06104. https://doi.org/10.1029/2007JD009872

Hawkins, E., & Sutton, R. (2011). The potential to narrow uncertainty in projections of regional precipitation change. *Climate Dynamics*, 37, 407–418. https://doi.org/10.1007/s00382-010-0810-6

Holland, P. W., & Welsch, R. E. (1977). Robust regression using iteratively reweighted least-squares. *Communications in Statistics, A6*, 813–827.

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*(D14), D14105. https://doi.org/10.1029/2011JD017237

Kay, J. E., & Gettelman, A. (2009). Cloud influence on and response to seasonal Arctic sea ice loss. *Journal of Geophysical Research, 114*(D18), D18204. https://doi.org/10.1029/2009JD011773

Knutti, R., & Sediacek, J. (2012). Robustness and uncertainties in the new CMIP5 climate model projections. *Nature Climate Change*, 3, 369–373.

Lambert, S. J., & Boer, G. J. (2001). CMIP: evaluation and intercomparison of coupled climate models. *Climate Dynamics*, 17, 83–106.

Lauer, A., & Hamilton, K. (2013). Simulating clouds with global climate models: A comparison of CMIP3 results with CMIP5 and satellite data. *Journal of Climate*, 26(11), 3823–3845. https://doi.org/10.1175/JCLI-D-12-0045.1

Li, J.-L., Waliser, D. E., Chen, W.-T., Guan, B., Kubar, T., Stephens, G., et al. (2012). An observationally based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data. *Journal of Geophysical Research, 117*(D16), D16105. https://doi.org/10.1029/2012JD017640

Lock, A. P. (2009). Factors influencing cloud area at the capping inversion for shallow cumulus clouds. *Quarterly Journal of the Royal Meteorological Society*, 135, 941–952. https://doi.org/10.1002/qj.424

Mace, G. G., Zhang, Q., Vaughan, M., Marchand, R., Stephens, G., Tropea, C., & Winker, D. (2009). A description of hydrometer layer occurrence statistics derived from the first year of merged Cloudsat and CALIPSO data. *Journal of Geophysical Research, 114*, D00A26. https://doi.org/10.1029/2009JD009755

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica*, 13(3), 245–259. https://doi.org/10.2307/1907187

Marchand, R., Beagley, N., & Ackerman, T. P. (2009). Evaluation of hydrometeor occurrence profiles in the multiscale modeling framework climate model using atmospheric classification. *Journal of Climate*, 22(17), 4557–4573.

Meeth, G. A., Covey, C., McAvaney, B., et al. (2005). Overview of the coupled model intercomparison project. *Bulletin of the American Meteorological Society*, 86, 89–93. https://doi.org/10.1175/BAMS-86-1-89
Rossow, W. B., Walker, A. W., Beuschel, D. E., & Roiter, M. D. (1996). International Cloud Climatology Project (ISCCP) documentation of

Miller, S. D., Weeks, C. E., Bullock, R. G., Forsythe, J. M., Kucera, P. A., Brown, B. G., et al. (2014). Model intercomparisons: Preparing

Monerie, P. A., Fontaine, B., & Roucou, P. (2012). Expected future changes in the African monsoon between 2030 and 2070 using some

Wylie, D., Eloranta, E., Spinhirne, J. D., & Palm, S. P. (2007). A comparison of cloud cover statistics from the GLAS Lidar with HIRS.

Yatagai, A., Krishnamurti Kumar, V., Mishra, A. K., & Simon, A. (2014). Use of APHRODITE rain gauge-based precipitation and TRMM

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., et al., & the CloudSat Science Team (2002). The CloudSat

Zhang, M. H., & coauthors (2005). Comparison clouds and their seasonal variations in 10 atmospheric general circulation models with

Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., et al. (2016). Evaluation of CMIP5 twentieth century climate simulations for the Pacific Northwest USA. 

Sachindra, D. A., Ng, A. W. M., Muthukumaran, S., & Perera, B. J. C. (2016). Impact of climate change on urban heat island effect and

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

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93(4), 485–498. https://doi.org/10.1073/BAMS-D-11-00094.1

Treinberth, K. E., Basu, J., & Kiehl, J. (2009). Earth's global energy budget. Bulletin of the American Meteorological Society, 90(3), 331–348. https://doi.org/10.1175/2008BAMS2634.1

Tsubshina, Y., Ringer, M. A., Webb, M. J., & Williams, K. D. (2013). Quantitative evaluation of the seasonal variations in climate model cloud regimes. Climate Dynamics, 41(9-10), 2679–2696. https://doi.org/10.1007/s00382-012-1609-4

Wang, F., Xin, Z., Wang, Z., Cheng, J., Zhang, S., & Yang, S. (2014). Evaluation of cloud vertical structure simulated by recent RCC_AGCM versions through comparison with CALIPSO-GOCCP data. Advances in Atmospheric Sciences, 31(3), 721–733. https://doi.org/10.1007/s00376-013-3092-3

Wang, H., Xie, S.-P., & Liu, Q. (2016). Comparison of climate response to anthropogenic aerosol versus greenhouse gas forcing: Distinct pattern. Journal of Climate, 29(14), 5175–5188. https://doi.org/10.1175/JCLI-D-16-0106.1

Wylie, D., Eloranta, E., Spinhirne, J. D., & Palm, S. P. (2007). A comparison of cloud cover statistics from the GLAS Lidar with HIRS. Journal of Climate, 20(19), 4989–4981. https://doi.org/10.1175/JCLI4269.1

Yatagai, A., Krishnamurti Kumar, V., Mishra, A. K., & Simon, A. (2014). Use of APHRODITE rain gauge-based precipitation and TRMM

Zhang, M.-H., & coauthors (2005). Comparison clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. Journal of Geophysical Research, 110(D15), D15S02. https://doi.org/10.1029/2004JD005021

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

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93(4), 485–498. https://doi.org/10.1073/BAMS-D-11-00094.1

Treinberth, K. E., Basu, J., & Kiehl, J. (2009). Earth's global energy budget. Bulletin of the American Meteorological Society, 90(3), 331–348. https://doi.org/10.1175/2008BAMS2634.1

Tsubshina, Y., Ringer, M. A., Webb, M. J., & Williams, K. D. (2013). Quantitative evaluation of the seasonal variations in climate model cloud regimes. Climate Dynamics, 41(9-10), 2679–2696. https://doi.org/10.1007/s00382-012-1609-4

Wang, F., Xin, Z., Wang, Z., Cheng, J., Zhang, S., & Yang, S. (2014). Evaluation of cloud vertical structure simulated by recent RCC_AGCM versions through comparison with CALIPSO-GOCCP data. Advances in Atmospheric Sciences, 31(3), 721–733. https://doi.org/10.1007/s00376-013-3092-3

Wang, H., Xie, S.-P., & Liu, Q. (2016). Comparison of climate response to anthropogenic aerosol versus greenhouse gas forcing: Distinct pattern. Journal of Climate, 29(14), 5175–5188. https://doi.org/10.1175/JCLI-D-16-0106.1

Wylie, D., Eloranta, E., Spinhirne, J. D., & Palm, S. P. (2007). A comparison of cloud cover statistics from the GLAS Lidar with HIRS. Journal of Climate, 20(19), 4989–4981. https://doi.org/10.1175/JCLI4269.1

Yatagai, A., Krishnamurti Kumar, V., Mishra, A. K., & Simon, A. (2014). Use of APHRODITE rain gauge-based precipitation and TRMM

Zhang, M.-H., & coauthors (2005). Comparison clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. Journal of Geophysical Research, 110(D15), D15S02. https://doi.org/10.1029/2004JD005021