Earth and Space Science

RESEARCH ARTICLE
10.1029/2020EA001520

Improvements in Cloud and Water Vapor Simulations Over the Tropical Oceans in CMIP6 Compared to CMIP5

Jonathan H. Jiang[1], Hui Su[1], Longtao Wu[1], Chengxing Zhai[1], and Kathleen A. Schiro[1,2]

[1]Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, [2]Department of Environmental Sciences, University of Virginia, Charlottesville, VA, USA

Key Points:

- We use satellite data to assess the representation of clouds and water vapor structures in 28 climate models that participate in the CMIP6
- We find measurable improvements in CMIP6 models relative to CMIP5 models for both clouds and water vapor
- In addition, we find that the models' equilibrium climate sensitivity is positively correlated with overall performance scores for the models

Supporting Information:

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

Correspondence to:

J. H. Jiang,
Jonathan.H.Jiang@jpl.nasa.gov

Citation:

Jiang, J. H., Su, H., Wu, L., Zhai, C., & Schiro, K. A. (2021). Improvements in cloud and water vapor simulations over the tropical oceans in CMIP6 compared to CMIP5. *Earth and Space Science*, *8*, e2020EA001520. https://doi.org/10.1029/2020EA001520

Received 18 OCT 2020
Accepted 10 MAR 2021

Abstract Clouds and water vapor are among the most difficult quantities for global climate models to simulate because they are affected by physical processes that operate over scales unresolved by current climate models. We use NASA satellite data to assess the representation of clouds and water vapor structures in 28 climate models that participate in the Coupled Model Intercomparison Project Phase 6 (CMIP6). Each model is assigned numerical scores based on its performance in simulating spatial mean, variance and pattern correlation of multi-year mean clouds and water vapor structures in lower, middle, upper troposphere, and near the tropopause over tropical oceans. We find measurable improvements in CMIP6 models relative to CMIP5 models for both clouds and water vapor. The differences between models and satellite observations and the spread across the models are reduced. In addition, we find that the models' equilibrium climate sensitivity (ECS) is correlated with overall performance scores for both CMIP5 and CMIP6 models, with a weaker correlation in CMIP6, suggesting that the models that capture better tropical clouds and water vapor distributions tend to have higher ECS. The physical processes responsible for the apparent correlation between ECS and model performance score warrant further study.

Plain Language Summary An improved understanding of Earth's climate requires realistic climate model simulations. Currently there are more than 20 climate models in use around the world that form the basis of the Intergovernmental Panel on Climate Change (IPCC) assessments report. One of the largest uncertainties in climate change projections arise from how the models handle the complex feedback mechanisms of clouds and water vapor. Using satellite observations, we made detailed and quantitative evaluation of cloud and water vapor structures from the latest versions of the climate models and their predecessors. Our analyses reveal that the latest models are measurably improved over their predecessors. Individual models were graded based on their performance in simulating clouds and water vapor. The analyses find that the models with better performance also tend to simulate higher surface temperature increase in response to a doubling of CO_2.

1. Introduction

Clouds and water vapor are important players in modifying surface warming caused by increasing greenhouse gases (IPCC AR5). However, their spatial and temporal variations in the atmosphere are driven by processes that occur over multiple scales, some of which are smaller than the grid sizes of global climate models (GCMs) and thus are difficult to simulate. Previous studies documented the large errors in simulated clouds and water vapor structures in GCMs and their relations with large-scale dynamic and thermodynamic conditions (e.g., Dolinar et al. 2015; Jiang et al. 2012; Lebsock & Su, 2014; Li et al. 2005; Su et al. 2006, 2013; Tian et al., 2013; Waliser et al. 2009). These model evaluation studies provided useful references for climate model improvements. For example, Jiang et al. (2012) evaluated a number of climate model simulations of clouds and water vapor profiles of models participating in the Coupled Model Intercomparison Project Phase 5 (CMIP5). They showed that GISS model E2 produced upper tropospheric ice water content (IWC) about a factor of 10 greater than satellite retrievals, which motivated the development of a new ice cloud parameterization scheme with updated ice particle fall velocity and particle size assumption in the GISS model that drastically reduced the IWC biases (Elsaesser et al. 2017).

For the Coupled Model Intercomparison Project Phase 6 (CMIP6), increased spatial (both horizontal and vertical) resolutions and arguably more sophisticated model physics are implemented in the newest generations of GCMs (Meehl et al. 2020). Satellite retrievals of clouds and water vapor have also been updated

recently with better accuracy and precision (e.g., Austin et al. 2018; Livesey et al. 2020; Tian & Hearty, 2020). A comprehensive assessment of CMIP6 model simulations of clouds and water vapor against the best-available observations is clearly needed to help gauge the fidelity of climate change projections and guide further model improvements.

Clouds and water vapor are an indispensable part of global energy and hydrological cycles and quantitative assessment of their representation is justified. However, model performance in simulating clouds and water vapor also exhibit interestingly direct relevance to constraining the estimates of equilibrium climate sensitivity (ECS), the equilibrium global-mean surface temperature increases under a doubling of CO_2. For instance, Fasullo and Trenberth (2012) showed that climatological subtropical tropospheric relative humidity (RH) is highly correlated with ECS and drier and less cloudy subtropics is associated with higher ECS. Su et al. (2014) found that the similarity between the modeled and observed zonal-mean structures of cloud fraction (CF) and RH are positively correlated with ECS in the CMIP5 models with better performing models having higher ECS. Qu et al. (2018) elucidated that the CF and RH model performance metrics constructed in Su et al. (2014) are highly correlated with shortwave cloud feedbacks and thus ECS. Hence, model performance scores in reproducing climatological clouds and water vapor structures may serve as emergent constraints on ECS (Sherwood et al. 2020). A recent study (Schlund et al. 2020) evaluated a number of CMIP5-based emergent constraints on ECS in CMIP6 models and found their correlations with ECS changed from CMIP5 to CMIP6. This suggests that there are substantial structural differences between the CMIP5 and CMIP6 models. An update about the model performance scores in clouds and water vapor and their relations to ECS is warranted.

In this study, we provide a quantitative assessment of CMIP6 simulated water vapor mixing ratio (H_2O) and cloud water content (CWC) structures using the best-available observations from NASA satellites. We follow the performance scoring system used in Jiang et al. (2012) and present both CMIP5 and CMIP6 results whenever possible. We focus on the climatological mean structures of H_2O and CWC in terms of spatial mean, variance and correlation between the modeled and observed fields. Four vertical pressure levels are evaluated separately. A final integrated score is computed for each model by averaging all individual scores for each statistical property of H_2O and CWC at the four pressure levels. The relationship between the final scores and models' ECS values is presented.

2. Data

2.1. The CMIP6 Climate Models

Totally 28 coupled atmosphere-ocean models from the CMIP6 archive available at the time of the analysis are included in this study (Table 1). The vertical profiles of clouds (ice and liquid) and water vapor from the historical runs are used. The cloud outputs used for this study are *cli* for ice water mixing ratio and *clw* for liquid water mixing ratio, both vertically resolved (Taylor et al., 2012, 2018). As defined, *cli* (*clw*) includes both large-scale and convective cloud, which is calculated as the mass of cloud ice (liquid) in the grid cell divided by the mass of air (including the water in all phases) in the grid cell. Precipitating hydrometeors are considered in the calculation only if they affect the calculation of radiative transfer in model, but not included in the model output of *cli* and *clw*.

The total cloud water content (CWC) is the sum of *clw* and *cli*. The water vapor data in the model outputs are the specific humidity profiles *hus*, or the water vapor mixing ratio. The models' ECS values are taken from Meehl et al. (2020).

2.2. The Satellite Data

For satellite observations, we use NASA's A-Train satellites (Aqua, Aura, and CloudSat) that provide nearly simultaneous and co-located measurements of cloud and moisture profiles (L'Ecuyer & Jiang, 2010). The measurement parameters used in this study are similar to the previous study of Jiang et al. (2012), which include (a) water vapor (H_2O) profiles from 1,000 hPa to 300 hPa from the Atmospheric Infrared Sounder (AIRS) onboard Aqua, (b) upper tropospheric H_2O at 215 hPa and above from the Microwave Limb Sounder (MLS), and (e) ice water content profiles (IWC) and liquid water content (LWC) profiles from CloudSat.

Table 1
CMIP6 Models Used in this Study

Model name (center, model)	Experiment	Resolution (H, V) grid (lat × lon), level (model top)	Reference	ECS (K)	Score
AS-RCEC, TaiESM1	historical	~0.9° × 1.25°(192 × 288), 30 levels (~2 hPa)	https://furtherinfo.es-doc.org/CMIP6.AS-RCEC.TaiESM1.historical.none.r1i1p1f1	N/A	0.653
BCC, CSM2-MR	historical	~1.12° × 1.125° (160 × 320), 46 levels (1.46 hPa)	Wu et al. (2013, 2014, 2018); http://forecast.bcccsm.ncc-cma.net/htm	3.0	0.685
BCC, ESM1	historical	~2.79° × 2.8125° (64 × 128), 26 levels (2.19 hPa)	Wu et al. (2013, 2014, 2018); http://forecast.bcccsm.ncc-cma.net/htm	3.3	0.650
CAMS, CSM1-0	historical	~1.12° × 1.125° (160 × 320), 31 levels (10 hPa)	Rong et al. (2019), https://furtherinfo.es-doc.org/CMIP6.CAMS.CAMS-CSM1-0.historical.none.r1i1p1f1	2.3	0.645
CAS, FGOALS-g3	historical	~2.02° × 2° (80 × 180), 26 levels (2.19 hPa)	https://furtherinfo.es-doc.org/CMIP6.CAS.FGOALS-g3.historical.none.r1i1p1f1	N/A	0.702
CNRM, CM6-1-HR	historical	0.5° × 0.5° (360 × 720), 91 levels (78.4 km)	http://www.umr-cnrm.fr/cmip6/references	4.3	0.682
CNRM, CM6-1	historical	1.4° × 1.4° (128 × 256), 91 levels (78.4 km)	http://www.umr-cnrm.fr/cmip6/references	4.8	0.688
CNRM, ESM2-1	historical	1.4° × 1.4° (128 × 256), 91 levels (78.4 km)	http://www.umr-cnrm.fr/cmip6/references	4.8	0.715
CSIRO-ARCCSS, CM2	historical	1.25° × 1.875° (144 × 192), 85 levels (85 km)	https://furtherinfo.es-doc.org/CMIP6.CSIRO-ARCCSS.ACCESS-CM2.historical.none.r1i1p1f1	4.7	0.717
CSIRO-ARCCSS, ESM1-5	historical	1.25° × 1.875° (144 × 192), 38 levels (39.255 km)	https://furtherinfo.es-doc.org/CMIP6.CSIRO.ACCESS-ESM1-5.historical.none.r1i1p1f1	3.9	0.508
E3SM-Project, E3SM-1-0	historical	1° × 1° (180 × 360), 72 levels (0.1 hPa)	Golaz et al. (2019), https://furtherinfo.es-doc.org/CMIP6.E3SM-Project.E3SM-1-0.historical.none.r1i1p1f1	5.3	0.638
GFDL, ESM4	historical	1° × 1° (180 × 360), 49 levels (1 hPa)	https://furtherinfo.es-doc.org/CMIP6.NOAA-GFDL.GFDL-ESM4.historical.none.r1i1p1f1	2.6	0.724
GISS, E2-1-G	historical	2° × 2.5° (90 × 144), 40 levels (0.1 hPa)	https://data.giss.nasa.gov/modelE/cmip6	2.7	0.568
GISS, E2-1-H	historical	2° × 2.5° (90 × 144), 40 levels (0.1 hPa)	https://data.giss.nasa.gov/modelE/cmip6	3.1	0.578
GISS, E3trn705a	Historical	1° × 1.25° (180 × 288), 110 levels (0.0035 hPa)	Cesana et al. (2019), https://acp.copernicus.org/articles/19/2813/2019/	3.2	0.657
IPSL, CM6A-LR	historical	1.27° × 2.5° (143 × 144), 79 levels (40 km)	https://furtherinfo.es-doc.org/CMIP6.IPSL.IPSL-CM6A-LR.historical.none.r1i1p1f1	4.6	0.578
MIROC, ES2L	historical	2.79° × 2.8125° (64 × 128), 40 levels (3 hPa)	https://furtherinfo.es-doc.org/CMIP6.MIROC.MIROC-ES2L.historical.none.r1i1p1f2	2.7	0.619
MIROC, MIROC6	historical	1.4° × 1.4° (128 × 256), 81 levels (0.004 hPa)	https://furtherinfo.es-doc.org/CMIP6.MIROC.MIROC6.historical.none.r1i1p1f1	2.6	0.630
MOHC, HadGEM3-GC31-LL	historical	1.25° × 1.875° (144 × 192), 85 levels (85 km)	https://furtherinfo.es-doc.org/CMIP6.MOHC.HadGEM3-GC31-LL.historical.none.r1i1p1f3	5.6	0.716
MOHC, UKESM1-0-LL	historical	1.25° × 1.875° (144 × 192), 85 levels (85 km)	https://furtherinfo.es-doc.org/CMIP6.MOHC.UKESM1-0-LL.historical.none.r1i1p1f2	5.3	0.681
MPI, ESM1-2-LR	historical	1.865° × 1.875° (96 × 192), 47 levels (0.01 hPa)	Mauritsen et al. (2019), https://furtherinfo.es-doc.org/CMIP6.MPI-M.MPI-ESM1-2-LR.historical.none.r1i1p1f1	3.0	0.706
MRI, ESM2.0	historical	~1.12° × 1.125° (160 × 320), 80 levels (0.01 hPa)	https://furtherinfo.es-doc.org/CMIP6.MRI.MRI-ESM2-0.historical.none.r1i1p1f1	3.2	0.718
NCAR, CESM2	historical	~0.9° × 1.25° (192 × 288), 32 levels (2.25 hPa)	https://furtherinfo.es-doc.org/CMIP6.NCAR.CESM2.historical.none.r1i1p1f1	5.2	0.739
NCAR, CESM2-WACCM	historical	~0.9° × 1.25° (192 × 288), 70 levels (4.5 e−6 hPa)	https://furtherinfo.es-doc.org/CMIP6.NCAR.CESM2-WACCM.historical.none.r1i1p1f1	4.8	0.743

Table 1
Continued

Model name (center, model)	Experiment	Resolution (H, V) grid (lat × lon), level (model top)	Reference	ECS (K)	Score
NCC, NorESM2-LM	historical	~1.9° × 2.5° (96 × 144), 32 levels (3 hPa)	https://furtherinfo.es-doc.org/CMIP6.NCC.NorESM2-LM.historical.none.r1i1p1f1	2.5	0.712
NIMS-KMA, KACE1.0-G	Historical	1.25° × 1.875° (144 × 192), 85 levels (85 km)	Lee et al. (2019), https://furtherinfo.es-doc.org/CMIP6.NIMS-KMA.KACE-1-0-G.historical.none.r1i1p1f1	4.5	0.745
NUIST, NESM3	historical	~1.9° × 1.875° (96 × 192), 47 levels (1 hPa)	Cao et al. (2018), https://furtherinfo.es-doc.org/CMIP6.NUIST.NESM3.historical.none.r1i1p1f1	4.7	0.643
SNU, SAM0-UNICON	historical	~0.9° × 1.25° (192 × 288), 30 levels (2 hPa)	Park et al. (2019), https://furtherinfo.es-doc.org/CMIP6.SNU.SAM0-UNICON.historical.none.r1i1p1f1	3.7	0.633

Note. Their ECS values and integrated model performance scores for climatological water vapor and clouds over the tropical oceans are shown in the last two columns.
Abbreviation: ECS, equilibrium climate sensitivity.

For the tropospheric water vapor mixing ratio (H_2O) up to 300 hPa, AIRS version 6, Level 3 product is used, which is similar to the AIRS Obs4MIPs V2 Data (Tian & Hearty, 2020). The data are on a global 1° × 1° latitude-longitude grid, with vertical pressure levels from 1,000 hPa to 300 hPa. The estimated uncertainty is 25% including both random and sampling errors. For water vapor at pressure <300 hPa, we use version Aura MLS 4.2 Level 2 (Livesey et al., 2020) H_2O data set. The MLS measurement uncertainties (including biases) for H_2O in the tropical and subtropical regionals are 20% near 215 hPa and within 10% at 100 hPa and higher altitudes.

For CWC, CloudSat IWC and LWC from the 2B-CWC-RO (version R05) data set are used. The retrieved IWC and LWC include contributions from precipitating particles. We thus construct noPcp IWC/LWC at each grid box by removing the precipitating CWC profiles (rain, snow, drizzle and graupel) indicated by precipitation flags in the CloudSat 2C-PRECIP-COLUMN product (Haynes et al., 2009). The uncertainty of both IWC and LWC is about a factor of 2 due to the particle size assumptions used in the retrieval. Therefore, the range of observed CWC is within 0.5× to 2.0× of the retrieved values.

All the above-mentioned data sets were re-gridded onto a common 144 (longitude) × 91 (latitude) grid and 40 pressure levels as done for the model outputs. The 40 pressure levels are from the surface to 24 hPa, with intervals of 50 hPa in the middle troposphere and finer in the boundary layer and near the tropopause. The original CMIP6 outputs are interpolated linearly with respect to log-pressure from their native vertical levels to standard pressure levels. We have carried out sensitivity studies to test different vertical interpolation methods and compare the results. We find the maximum error due to the vertical interpolation is <20%, mostly near the tropopause. The model results used for comparison with satellite data are multiyear averages from the CMIP6 "historical" runs, which were generated from coupled atmospheric-ocean experiments under historical forcings (Eyring et al. 2016). We use the 20-year averages from 1995 to 2014 to represent the present-day climatological mean. The multi-year mean satellite measurements used in evaluating the models are averages over the following time periods: 4.5 years (June 2006 to December 2010) for CloudSat; 13 years (August 2004 to September 2016) for AIRS and MLS. Although these time periods do not overlap with those of the model outputs, no significant trends in clouds and water vapor are found in the model outputs from 1995 to 2015. The A-Train satellites are sun-synchronous with equatorial crossings at ~1:30 p.m. and ~1:30 a.m., which leads to some sampling biases associated with diurnal variations. To reduce the effects of diurnal sampling biases, we focus on the tropical oceans (30ºN to 30ºS) when quantitatively scoring the model performances, as diurnal variations are much smaller over ocean than over land. In previous studies (e.g., Jiang et al. 2015), we estimated the magnitude of diurnal bias in earlier versions of CMIP5 models, as well as reanalysis data by comparing standard monthly mean IWCs from the GCMs with the monthly mean IWCs constructed by sampling 3-hourly model outputs onto A-Train satellite tracks. We found that the differences between two monthly means over the tropical ocean are generally <2% compared to up to 200% differences over land. We thus assume that diurnal variation introduces a bias of less than 2% in the tropical oceanic means, significantly smaller than the measurement uncertainties. In addition, as

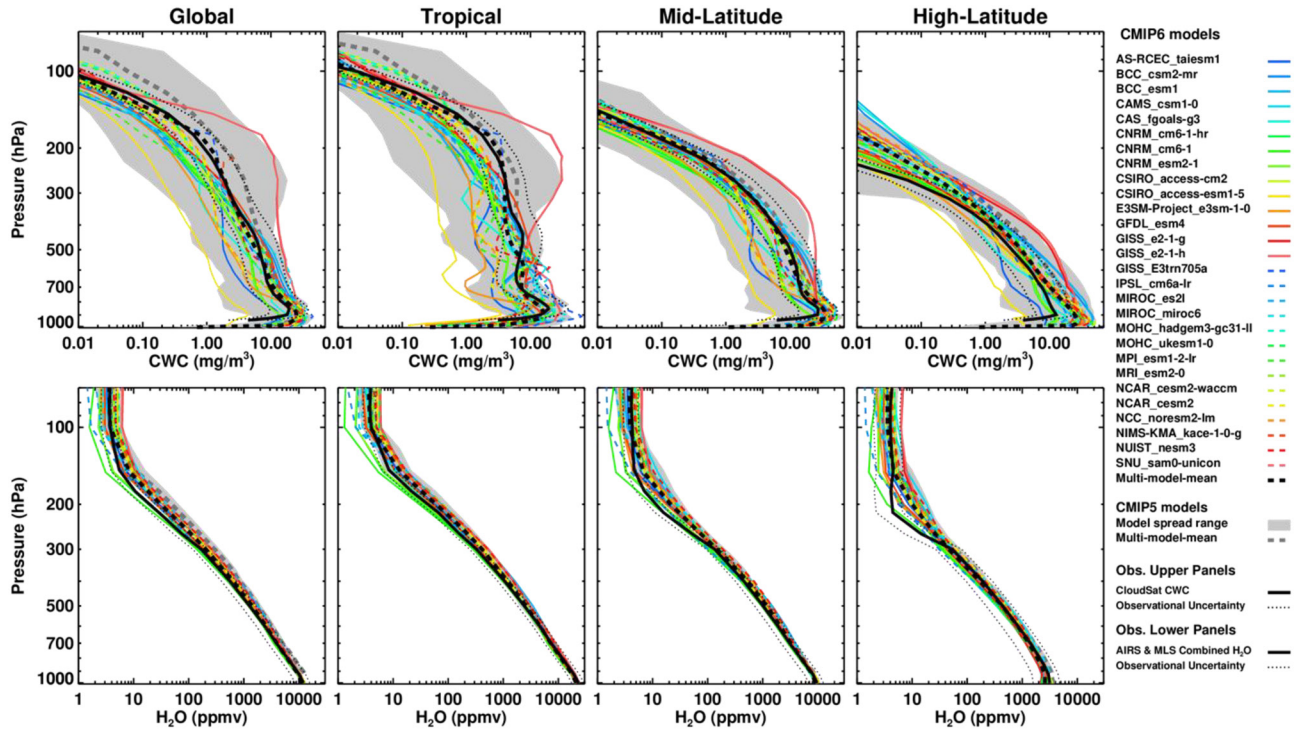

Figure 1. The colored lines show the multiyear mean vertical profiles of CWC (top-panels) and H_2O (lower-panels) from the 28 CMIP6 models. The CloudSat CWC and Aqua-AIRS and Aura-MLS H_2O profiles are indicated by the solid black lines. Observational uncertainty limits are indicated by the dotted black lines. The MMMs of the 28 CMIP6 models are shown by the dashed black lines. The light-gray shades mark the spread of CWC and H_2O profiles for the CMIP5 models and the CMIP5 MMMs are shown by the dashed dark-gray lines. CWC, cloud water content; MMM, multi-model mean.

satellite cannot accurately retrieve clouds and moisture near the surface, we limit our analysis to altitudes of 900 hPa and above.

3. Results

We first examine the vertical structures of clouds and moisture averaged over the globe. The colored lines in Figure 1 show the multiyear mean vertical profiles of CWC (top-panels) and H_2O (lower-panels) from the 28 CMIP6 models. The observational CWC profiles from CloudSat for CWC and H_2O profiles from AIRS and MLS, are indicated by the solid black lines. Observational uncertainty limits are indicated by the dotted black lines. The ensemble-mean values of the 28 CMIP6 models, that is, the multi-model mean (MMM), are shown by the dashed black lines. For comparison, the spread of CWC and H_2O profiles from the CMIP5 models are shown by the light-gray shade, and the CMIP5 MMMs are shown by the dashed dark-gray lines.

For clouds (Figure 1, top-panel), there are large spreads among CMIP6 model CWCs, although the MMM CWCs are very close to the observed values at all altitudes and latitude regions, with overall differences between MMM and observation less than 20%. When compared with CMIP5 (Jiang et al. 2012), the CMIP6 CWC spreads are reduced in the upper troposphere (pressure <400 hPa), especially in the tropics. The exceptions are the two GISS e2 models GISS e2-1-g and GISS_e2-1-h—their CWCs between 500 and 200 hPa are about five times larger than the MMM and observation. Another difference between the CMIP6 and CMIP5 results is that the CWCs at pressure <500 hPa are reduced in all CMIP6 models, except for GISS e2-1-g and GISS_e2-1-h, resulting in smaller MMM in CMIP6 than in CMIP5. Near the tropopause (pressure <150 hPa), the CWCs from CIROS-access-cm2 and the two MOHC models, MOHC_hadgem3-gc31-II and MOHC_ukesm1-0, are about 2–4 times larger than the MMM and the observations.

For H_2O (Figure 1, bottom-panel), the MMMs are very close to the observed values at all altitudes and latitude regions, with the mean difference of <5%. Similar to CMIP5, the CMIP6 H_2O model spread are <20% in the mid- and lower troposphere, but more than 200% above 200 hPa altitude in the upper troposphere

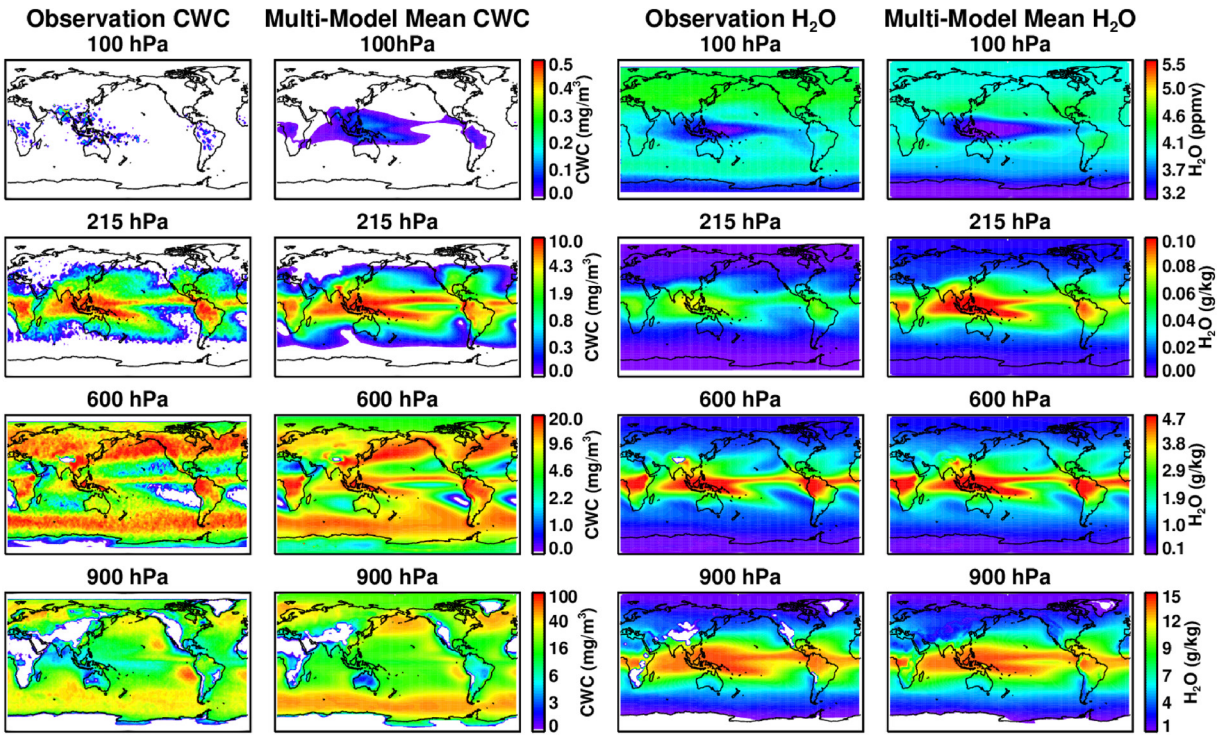

Figure 2. Multiyear mean CWC (left two panels) and H_2O (right two panels) maps for satellite observations and CMIP6 multi-model-mean. The CWC observations are from the CloudSat non-precipitating CWC data; the H_2O observations are from Aqua AIRS below 300 hPa altitude and from Aura MLS above the 300 hPa altitude. CWC, cloud water content; MLS, Microwave Limb Sounder.

and lower stratosphere. Model differences from the observations are small (<10%) in the middle and lower troposphere, but range from 1% to 100% near 100 hPa. The most notable result is that the CMIP6 models, on average, are drier than the CMIP5 models in the upper troposphere. In CMIP5, modeled H_2O profiles are mostly biased high compared to the observations in the mid- and upper troposphere between 150 and 700 hPa in all latitude bands. This bias has been corrected in the CMIP6 models. In CMIP6, however, the H_2O spread in the upper troposphere has increased, compared to CMIP5.

The similarity between the CMIP6 MMM and observations can be further illustrated by the maps of the CWC and H_2O multi-year means (Figure 2). Four pressure levels (100, 215, 600, and 900 hPa) are displayed. At 100 hPa, both the observations and CMIP6 MMM show that high CWC are collocated with low H_2O in the tropical western Pacific with a high (low) bias in the modeled CWC (H_2O) at 100 hPa. At 215 hPa, high CWC and H_2O are associated with tropical deep convection (Su et al. 2006). It is clear that the MMM H_2O is biased high compared to the MLS observation. At 600 hPa, high values of CWC are mostly in the mid- and high latitudes and the inter-tropical convergence zone (ITCZ). The modeled CWC is lower than the observed CWC in the mid- and high-latitudes but higher than the observed in the ITCZ. At 900 hPa, low clouds are distributed over the west coast of American continents in the subtropics, the Southern Ocean and the North Pacific and Atlantic oceans. The CMIP6 MMM appears to have a low bias in the subtropical low cloud CWC but a high bias in high-latitude CWC at 900 hPa compared to the CloudSat data. For H_2O at 900 hPa and 600 hPa, the ensemble means are very close to the observed.

The CWC and H_2O multi-year mean maps for individual models are shown in the Supplementary Figure S1. It is clear that the differences between the models and the differences between each model and the observations are huge. However, as shown in Figure 2, the ensemble means from all models effectively average out individual models' biases, delivering a reasonable representation of the climatological clouds and water vapor distributions.

To quantitively assess each model's performance against the observations, we focus on the tropical oceanic regions (30ºN to 30ºS) and use the scoring system as in Jiang et al. (2012). We focus on the models'

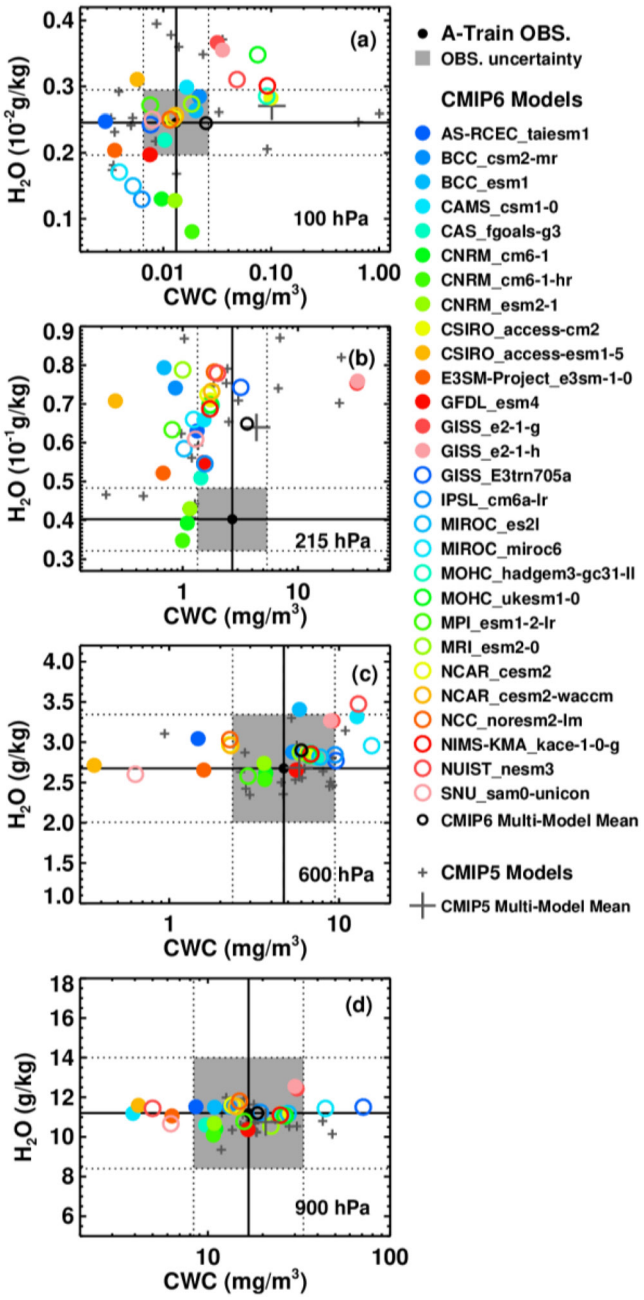

Figure 3. Scatterplots of multiyear means H_2O versus CWC averaged over the tropical oceans (30ºN–30ºS) at (a) 100, (b) 215, (c) 600, and (d) 900 hPa pressure levels. Black dots are the observational multiyear means with the gray area indicating the observational uncertainties. Colored dots and open-circles are the multiyear means from the CMIP6 models. Black open-circles represent the CMIP6 MMMs. Small gray crosses are the CMIP5 results, and the CMIP5 MMM is shown by the large gray cross. CWC, cloud water content; MMM, multi-model mean.

performances in simulating spatial mean and variance of CWC and H_2O and their spatial correlations with the observations.

Figure 3 shows scatterplots of the multi-year mean tropical oceanic H_2O versus CWC at four pressure levels: 100, 215, 600, and 900 hPa.

At 100 hPa, the CMIP6 model spread for CWC is about 0.2%–500% of the observations, smaller than the CMIP5 model spread of 0.3%–1,500%. The CMIP6 MMM CWC is about 160% larger than the CloudSat CWC, but within the observational uncertainty of 200%. It is a factor of 10 smaller than the CMIP5 MMM. The MMM H_2O at 100 hPa is about 5% drier than the MLS H_2O, also within the observational uncertainty of 10%. The model spread of H_2O ranges from ~20% to 145% of the MLS observation.

At 215 hPa, all the CMIP6 models except the two GISS e2 models produce spatial-mean CWCs lower than CloudSat. The CMIP6 MMM CWC is about 140% larger than the observed, mainly due to the high biases in the two GISS e2 models. The CMIP6 MMM is close to the CMIP5 MMM and is within the observational uncertainty of 200%. For 215 hPa H_2O, most CMIP6 models overestimate the moisture concentration, except the three CNRM models, which are drier than the MLS observation. The CNRM_cm6-1-h has a dry bias larger than the observational uncertainty of 20%. Similar to the CMIP5 but slightly better, the CMIP6 MMM H_2O is still about 44% wetter than the MLS measurement, which is outside the observational uncertainty.

In the mid- (600 hPa) and lower- (900 hPa) troposphere, the CMIP6 model simulated H_2O is all within the observational uncertainty of 25%. CWCs from CMIP6 have larger spread than the CMIP5. Relative to the observed CWCs, the CMIP6 CWC spatial means are 4%–250% at 600 hPa and 10%–250% at 900 hPa. However, the MMM CMCs at 600 and 900 hPa are very close to the observed and within the observational uncertainties.

Figure 4 shows the Taylor diagrams that illustrate the spatial variances and correlations for the CWCs and H_2Os at 100, 215, 600, and 900 hPa pressure levels. The definitions of the symbols are the same as those in Figure 3.

For CWC at 100 hPa, there are large differences among CMIP6 models simulated spatial variance and spatial correlation with the observation. The CMIP6 CWC all have weak correlations <0.4 and large normalized root-mean-squared-error (RMSE) greater than the observed spatial standard deviation. The CMIP6 MMM CWC has a correlation of 0.3 and an RMSE of 0.9. In comparison, the CMIP5 MMM has a correlation of 0.84 and normalized RMSE (relative to the observed standard deviation) of ~2.

At 215 hPa, the CMIP6 CWCs yield spatial correlations with the observation around 0.8 to 0.9, with the highest being 0.93 and the lowest being 0.72. This result is similar to CMIP5, suggesting that the spatial locations of deep convection are well captured. However, all CMIP6 models except the two GISS e2 models have smaller standard variance and RMSE due to the generally smaller CWC values compared to the observation (see Figure 3). The CMIP6 modeled standard deviations and RMSE for CWC at both 600 hPa and 900 hPa are quite scattered with spatial correlations from 0.6 to 0.8 at 600 hPa and 0.3 to 0.9 at 900 hPa. Both GISS e2 models have negative spatial correlations at 900 hPa, which indicates a problem in simulating the locations of marine stratiform clouds (see Figure S1).

10.1029/2020EA001520

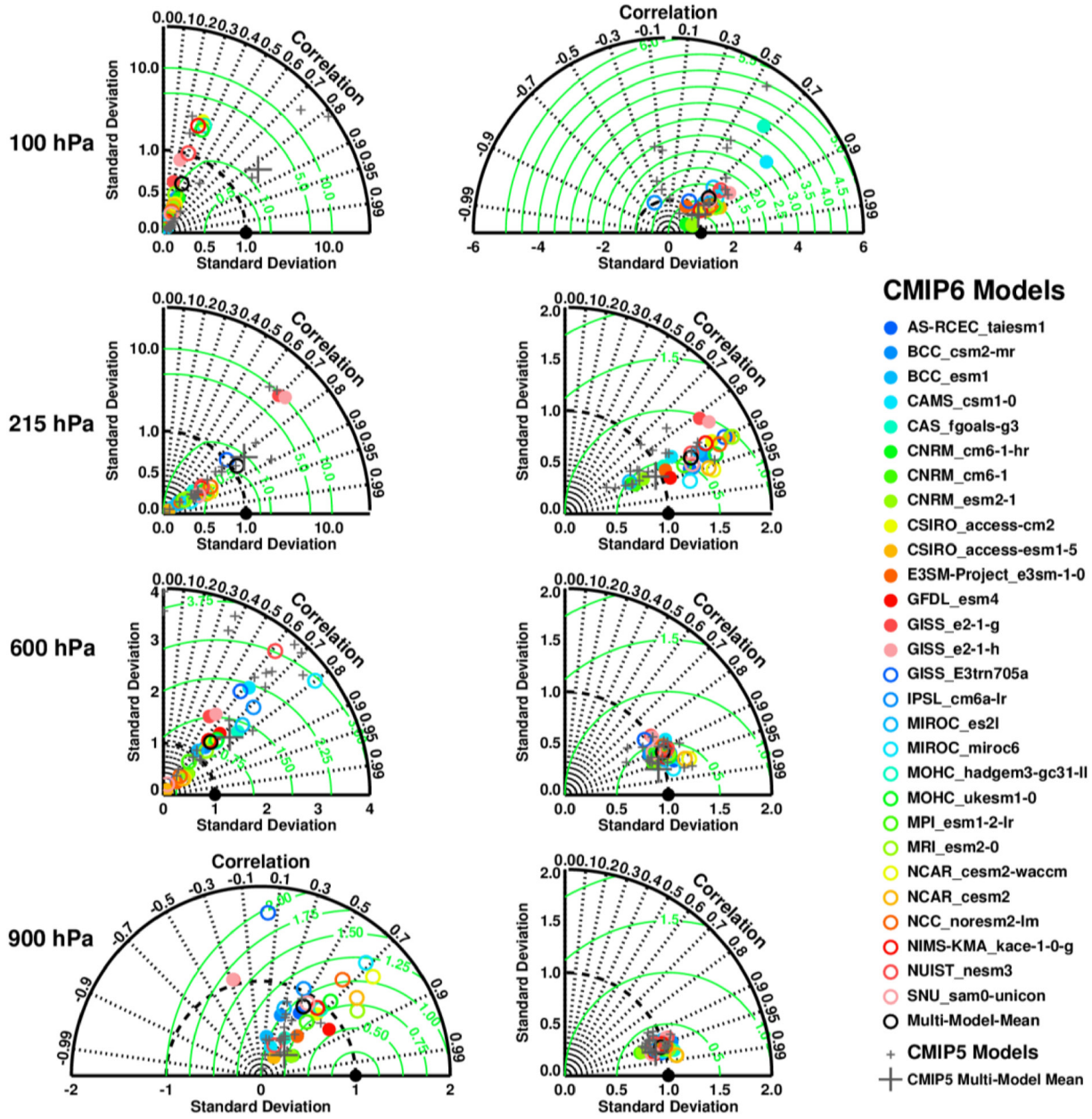

Figure 4. Taylor diagrams showing the tropical (30ºN–30ºS) oceanic multi-year mean CWC (left-column) and H_2O (right-column) simulations from the CMIP6 and CMIP5 models (definitions for the symbols are the same as those in Figure 3) as compared to the satellite observations (the black dot on the horizontal axis with the value of 1 = the standard deviation of the observed variable). The horizontal axis represents the fraction of the modeled spatial variation pattern that can be explained by the observed spatial pattern. The vertical axis represents the standard deviation of the modeled spatial pattern orthogonal to the observation, which is normalized by the observed standard deviation. The distance to the origin from each point in the Taylor Diagram corresponds to the spatial standard deviation of modeled variable and the distance of each point to the observed point (1, 0) on the *x* axis is the RMS of the difference between the modeled and observed quantities, as scaled by the green arc-lines. The correlation between the modeled and observed quantities is marked by the numbers on the black arc.

The CMIP6 modeled H_2O exhibits better performance than the CWC at the same levels. At 100 hPa, one model, the IPSL_cm6a-lr, produces a negative correlation with the MLS observation. The spreads of standard deviation and RMSE are large at 100 and 215 hPa. The spatial correlations range between 0.5 to 0.9 at 100 hPa and are greater than 0.9 at 215, 600, and 900 hPa. At 100 hPa, the spatial standard deviation of the modeled H_2O is about 0.5–5 times of the observed. At 215 hPa, the standard deviation of the modeled H_2O

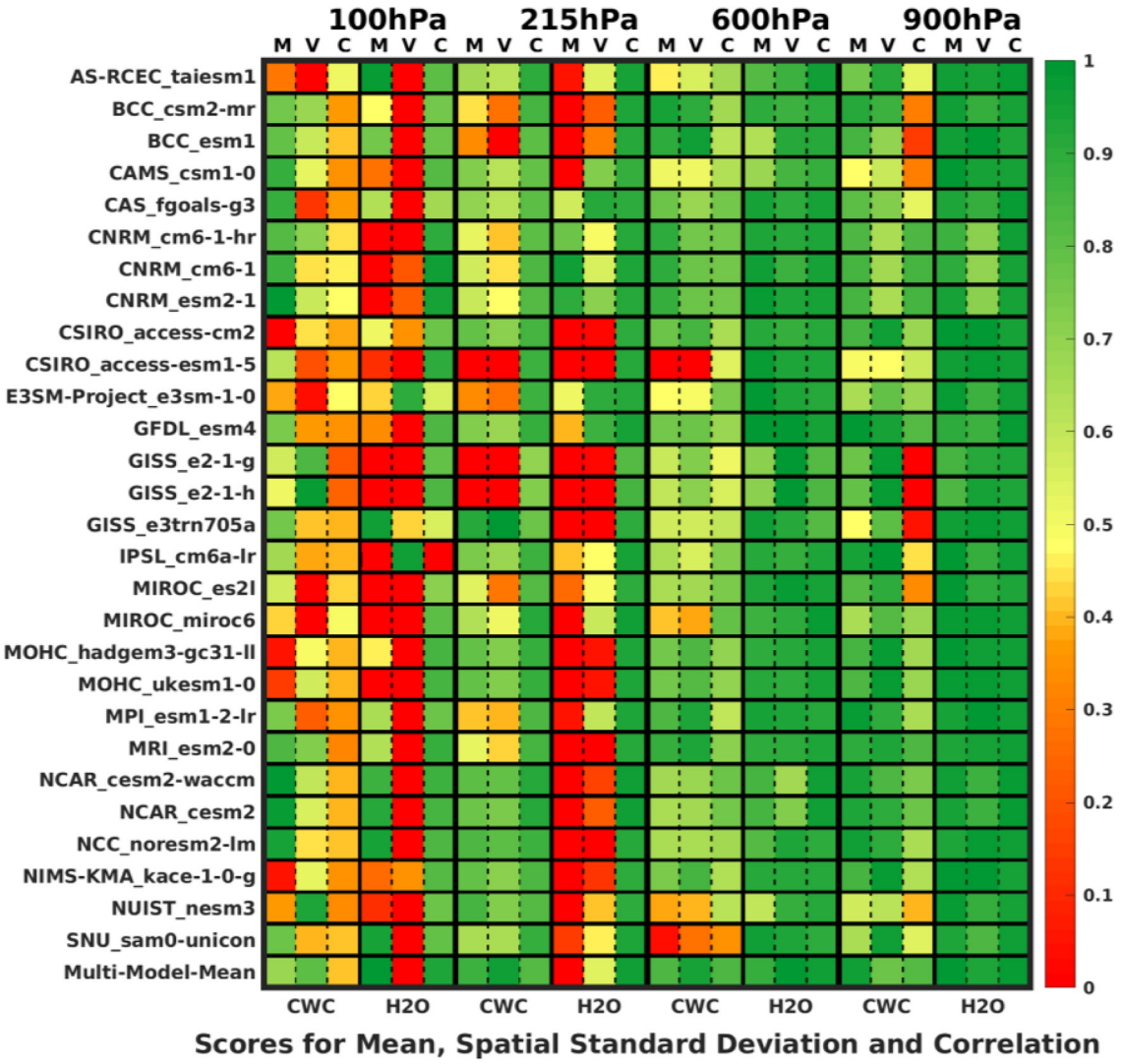

Figure 5. Color-coded summary of CMIP6 model performance scores for simulating CWC and H_2O at 100, 215, 600, and 900 hPa pressure levels, based on the scoring method of Jiang et al. (2012). For each model, M = spatial mean performance scores, V = spatial variance performance scores, and C = spatial correlation performance scores. The color-bar: 1 means perfect skill and 0 means no skill. CWC, cloud water content.

is about 0.5–1.5 times of the observed. At 600 and 900 hPa, the simulated and observed spatial variances are very comparable.

We now quantitatively score the model performances. The model performance is ranked following a scoring system in Jiang et al. (2012), in which the distances between the modeled and observed spatial mean, variance and correlation are scaled with respect to the observational data uncertainty to values between 0 and 1, with 1 being the perfect skill and 0 being no skill. The model performance scores are given at four vertical pressure levels, representing the atmospheric boundary layer (900 hPa), mid-troposphere (600 hPa), upper troposphere (215 hPa), and the tropopause layer (100 hPa) for both CWC and H_2O.

Figure 5 summarizes the 28 CMIP6 models' performances. We find that most CMIP6 models perform better at lower- to mid-tropospheric levels (900 and 600 hPa) than in the upper troposphere (215 and 100 hPa), especially for water vapor simulations. However, six models have larger errors in simulating the locations of low clouds, resulting in low spatial correlations for CWC at 900 hPa—BCC csm2-mr and esm1, CAMS csm1-0, GISS e2-1-g and e2-1-h, and MIROC es2l. We note that the new GISS-e3trn705a model has improved performance scores at all pressure levels for both water vapor and clouds. Most models do not simulate the observed spatial mean and variance of CWC or H_2O, or both, at 215 and 100 hPa very well. However,

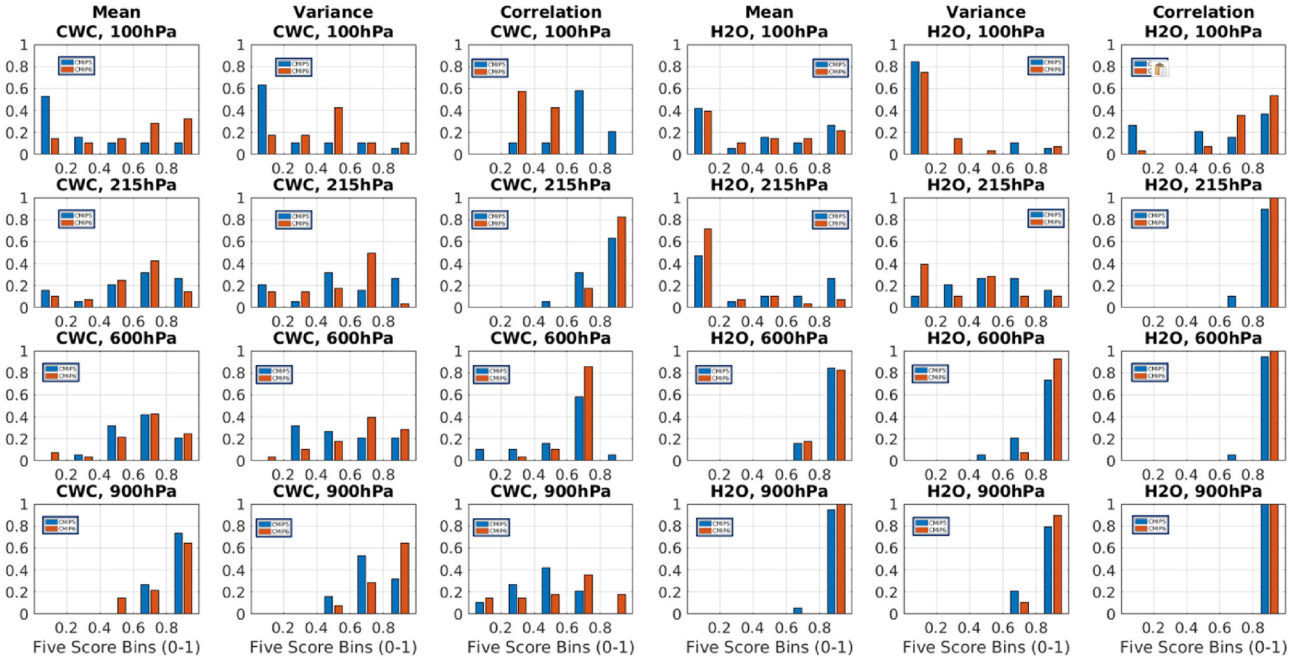

Figure 6. Histograms of CMIP6 (red) and CMIP5 (blue) model performance scores for spatial mean, variance and correlation of CWC (left three columns) and H_2O (right three columns) at the four pressure levels (rows). Five bins are used to compute the histograms: 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0. CWC, cloud water content.

they have better scores for spatial correlation at 215 hPa for both CWC and H_2O, and at 100 hPa for H_2O, indicating that models generally capture the climatological locations of deep convection but have difficulties in reproducing the magnitude of convective influence on the upper troposphere. The ensemble model means (the last row) exhibit relatively superior performance at all pressure levels, except for the 215 hPa spatial mean and 100 hPa spatial variance of H_2O, which are both below 0.2. The low scores for H_2O at these levels reflect the fact that most models have high biases at 215 hPa and low biases at 100 hPa in H_2O, compared to the Aura MLS observations.

To compare the model performance scores between CMIP5 and CMIP6, we show in Figure 6 the histograms of CMIP6 and CMIP5 models' performance scores for spatial mean, variance and correlation at each pressure level from 100 hPa to 900 hPa. Five bins are used to compute the histograms: 0–0.2, 0.2–0.4, 0.4–0.6, 0.6–0.8, and 0.8–1.0.

Overall, both CWC and H_2O show improvements in CMIP6, compared to CMIP5. Near the tropopause at 100 hPa, better spatial mean and variance in CWC and better spatial correlation in H_2O are shown. Despite the fact that the improvements in spatial mean for CWC and spatial mean and variance for H_2O are encouraging, the cloud and moisture amounts in deep convective regions are still poorly captured in CMIP6. At 215 hPa, there is improvement in CMIP6 H_2O spatial variance and correlation relative to CMIP5, but the spatial mean still has low scores due to large moist biases in the upper troposphere (also see Figure 2). The 215 hPa CWC shows notable improvement in correlation but most models have a low CWC bias, resulting in poor scores in spatial mean and variance. Overall, CMIP6 models show only slight improvement at 215 hPa.

There are clear improvements from CMIP5 to CMIP6 in the mid- and lower- troposphere (600 and 900 hPa) in both CWC and H_2O, especially at 600 hPa where the spatial mean, variance and correlation performance scores for both CWC and H_2O are improved in CMIP6. Similarly, at 900 hPa, overall improvements in both CWC and H_2O are shown, although the CMIP6 CWC spatial mean scores show little change from CMIP5 due to the large spread in simulated CWC (Figure 2).

Averaging the model performance scores for all pressure levels, we obtain an integrated score for each model, shown in Figure 7. There are 11 CMIP6 models scoring at 0.7 and above, including NIMS-KMA-kace, the two NCAR models, GFDL-esm4, MRI-esm2, CSIRO-access-cm2, MOHC-hadgem3, CNRM-esm2,

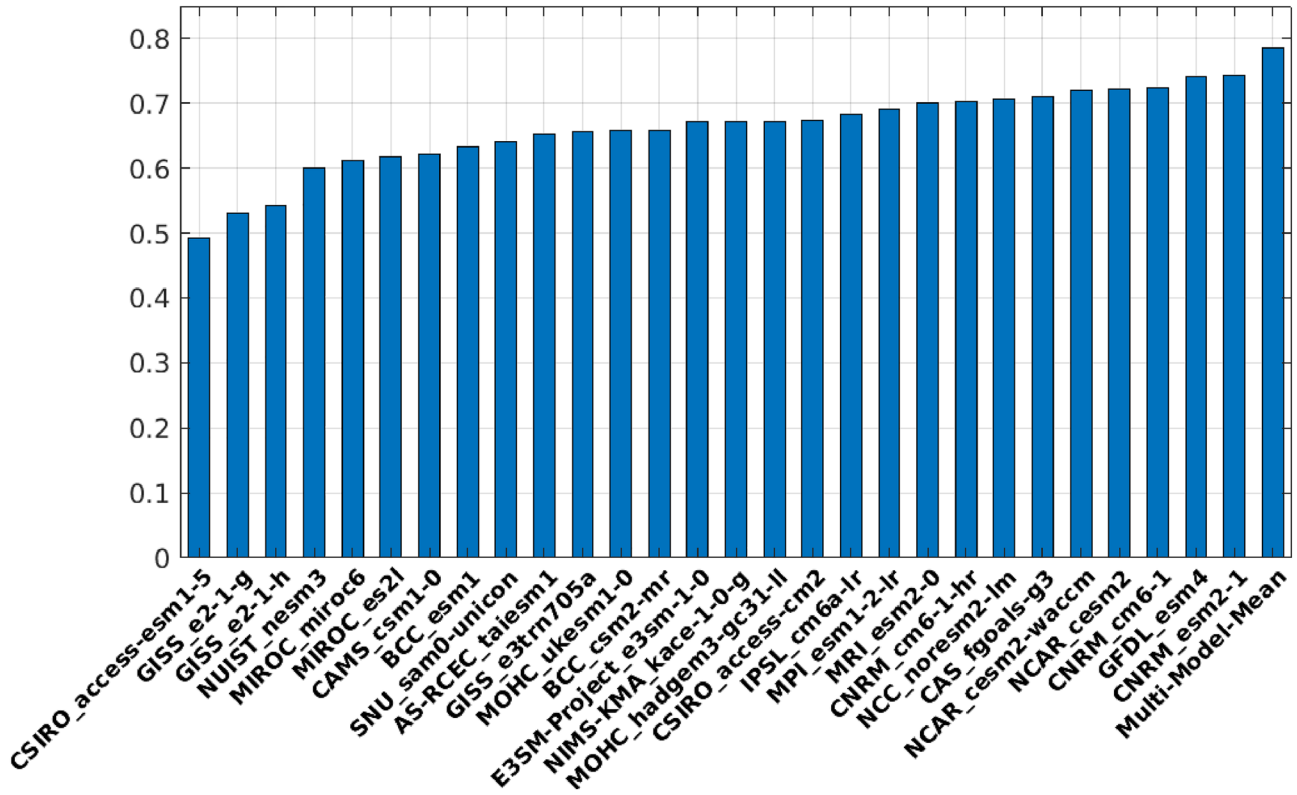

Figure 7. Integrated CMIP6 model performance scores.

NCC-noresm2, MPI-esm1, and CAS-fgoals-g3. There are 13 models scoring between 0.6 and 0.7, including AS-RCEC-taiesm1, the two BCC BCC models, CAMS-csm1-0, the two CRNM-cm6 models, E3SM-1-0, the newest GISS-e3trn705a model, the two MIROC models, MOHC-ukesm1-0-ll, NUIST-nesm3, and SNU-sam0-unicon. The CSIRO-access-esm1 and the two GISS e2 models score between 0.5 and 0.6.

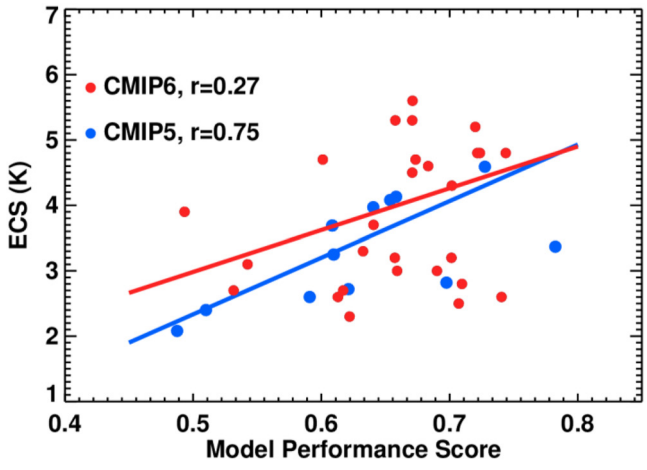

Figure 8. Relationship between the CMIP5 and CMIP6 models performance scores and their ECS values. The CMIP5 performance scores and ECS are from Jiang et al. (2012) and Su et al. (2013); the CMIP6 performance scores are based on this study with ECSs from Meehl et al. (2020). ECS, equilibrium climate sensitivity.

Not surprisingly, the overall score for the MMM turns out to be the best (scored 0.8) among all the models. It is comforting as the use of multi-model ensembles in climate projections is a common practice and the MMM is generally perceived as closer to the "truth" than any single model alone, as found in previous model evaluation studies (e.g., Gleckler et al. 2008).

Lastly, we examine the relationship between the models' performance scores in simulating CWC and H_2O and their ECS values for both CMIP5 and CMIP6. We find that the model performance scores are correlated with ECS with correlation coefficient of 0.27 for CMIP6 and 0.75 for CMIP5 (Figure 8). The positive correlations are primarily contributed by the correlations of the CWC performance scores at all levels with ECS (figure not shown). The models that more closely match the observed cloud and water vapor structures tend to have higher ECS than the models that deviate more from the observations, especially for CMIP5, consistent with many previous studies of the emergent constraints on ECS in CMIP5 (Fasullo & Trenberth, 2012; Sherwood et al., 2014; Su et al., 2014; Tian 2015; Zhai et al., 2015).

The much weaker correlation between the model performance scores and ECS in CMIP6 than in CMIP5 is consistent with the results in Schlund et al. (2020) in that many emergent constraints derived from

CMIP5 models do not correlate with ECS significantly in CMIP6. The reason for the weaker correlation is not clear. One explanation could be that the overall improvement in the model performances (and thus a smaller range in the scores) would tend to explain less variance of the ECS. On the other hand, the increased complexity in CMIP6 model systems may dilute the contribution of certain moist processes related to the water vapor and cloud climatology to the spread of ECS as many more processes are at play. The physical mechanisms responsible for the different correlations are the subject of on-going research.

4. Summary and Conclusions

Using satellite observations (CloudSat, AIRS, and MLS), we assess the simulated multi-year mean cloud and water vapor profiles from 28 CMIP6 models historical simulations. We use the grading scheme in Jiang et al. (2012) to quantitatively evaluate model performance in simulating clouds and water vapor at four pressure levels (from the boundary layer to the tropopause) over the tropical (30°N–30°S) oceans in terms of spatial mean, correlation and standard deviation. Compared to CMIP5, we find that both cloud water content (CWC) and water vapor volume mixing ratio (H_2O) simulations have improved in the lower- (900 hPa) and mid-troposphere (600 hPa) from CMIP5 to CMIP6, where the performance scores are significantly higher than those for the upper troposphere and near the tropopause. CMIP6 shows overall improvement in CWC and H_2O at 100 hPa. However, little improvement is found at 215 hPa. A prevailing moist bias is found at 215 hPa for H_2O over the tropical deep convective regions. The implication of this moist bias for water vapor feedback and cloud radiative effects merits further study.

In addition, we find that the models' equilibrium climate sensitivity (ECS) is positively correlated with the integrated CWC and H_2O performance scores for both CMIP6 and CMIP5 models, but the correlation becomes much weaker in CMIP6. Quantitative evaluation of model simulations using best-available observations has direct relevance to constraining future warming magnitude. Research efforts are underway to disentangle the exact physical processes that underpin the apparent correlation between ECS and model performance scores and why the correlation weakens substantially in CMIP6.

Data Availability Statement

All the climate model data used for this research can be downloaded from the PCMDI website at https://pcmdi.llnl.gov/CMIP6/. The satellite observational datasets can be downloaded from https://airs.jpl.nasa.gov/for AIRS data, http://www.cloudsat.cira.colostate.edu/for CloudSat data and https://mls.jpl.nasa.gov/ for MLS data. For additional questions regarding the data sharing, please contact the corresponding author at Jonathan.H.Jiang@jpl.nasa.gov.

Acknowledgments
The authors conducted the research at the Jet Propulsion Laboratory, California Institute of Technology, under contract by NASA. The research was supported by a NASA funding for CMIP6 climate model analysis and evaluation using NASA data, as well as NASA ROSES MAP and TASNPP Programs. We thank the climate modeling groups, the Program for Climate Model Diagnosis and Inter-comparison (PCMDI), and the WCRP's Working Group on Coupled Modeling for their roles in making the WCRP CMIP6 multi-model datasets available. The authors also thank the support from AIRS, CloudSat and MLS science teams, and the support from NOAA-MAPP (NA20OAR4310394) and DOE-RGMA (253847) program offices.

References

Austin, R. T., & Wood, N. B. (2018). *Level 2B radar-only cloud water content (2B-CWC-RO) process description and interface control document, Product Version P1_R05. NASA JPL CloudSat project document revision 0*. (p. 55). Retrieved from http://www.cloudsat.cira.colostate.edu/sites/default/files/products/files/2B-CWC-RO_PDICD.P1_R05.rev0_.pdf

Cao, J., Wang, B., Yang, Y. M., Ma, L., Li, J., Sun, B., et al. (2018). The NUIST earth system model (NESM) version 3: Description and preliminary evaluation. *Geoscientific Model Development*. https://doi.org/10.5194/gmd-11-2975-2018

Cesana, G., 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*, 2813–2832. https://doi.org/10.5194/acp-19-2813-2019

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*, 2229. https://doi.org/10.1007/s00382-014-2158-9

Elsaesser, G. S., Del Genio, A. D., Jiang, J. H., & van Lier-Walqui, M. (2017). An improved convective ice parameterization for the NASA GISS global climate model and impacts on cloud ice simulation. *Journal of Climate*, *30*, 317. https://doi.org/10.1175/JCLI-D-16-0346.1

Eyring, V., Bony, S., Meehl, 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

Fasullo, J. T., & Trenberth, K. E. (2012). A less cloudy future: The role of subtropical subsidence in climate sensitivity. *Science*, *338*, 792–794. https://doi.org/10.1126/science.1227465

Gleckler, P. J., Taylor, K. E., & Doutriaux, C. (2008). Performance metrics for climate models. *Journal of Geophysical Research*, *113*, D06104. https://doi.org/10.1029/2007JD008972

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.

Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., Miller, S. D., Mitrescu, C., Wood, N. B., & Tanelli, S. (2009). Rainfall retrieval over the ocean with spaceborne W-band radar. *Journal of Geophysical Research*, *114*, D00A22. https://doi.org/10.1029/2008JD009973

Jiang, J. H., Su, H., Zhai, C., Janice Shen, T., Wu, T., Zhang, J., et al. (2015). Evaluating the diurnal cycle of upper-tropospheric ice clouds in climate models using SMILES observations. *Journal of the Atmospheric Sciences*, *72*, 1022–1044. https://doi.org/10.1175/JAS-D-14-0124.1

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*. https://doi.org/10.1029/2011JD017237

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

L'Ecuyer, T. S., & Jiang, J. H. (2010). Touring the atmosphere aboard the A-Train. *Physics Today*, *63*(7), 36–41. https://doi.org/10.1063/1.3463626

Lee, J., Kim, J., Sun, M.-A., Kim, B.-H., Moon, H., Sung, H. M., et al. (2019). Evaluation of the Korea Meteorological Administration Advanced Community Earth-System model (K-ACE). *Asia-Pacific Journal of Atmospheric Sciences*, *56*(3), 381–395. https://doi.org/10.1007/s13143-019-00144-7

Li, J.-L., Waliser, D. E., Jiang, J. H., Wu, D. L., Read, W., Waters, J. W., et al. (2005). Comparisons of EOS MLS cloud ice measurements with ECMWF analyses and GCM simulations: Initial results. *Geophysical Research Letters*, *32*. https://doi.org/10.1029/2005GL023788

Livesey, N. J., Read, W., Wagner, P., Froidevaux, L., Lambert, A., Manney, G., et al. (2020). *Earth observing system (EOS) Aura Microwave Limb Sounder (MLS) version 5.0x level 2 and 3 data quality and description document* (JPL D-105336 Rev. A). Retrieved from https://mls.jpl.nasa.gov/data/v4-2_data_quality_document.pdf

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., et al. (2019). Developments in the MPI-M Earth System Model Version 1.2 (MPI-ESM1. 2) and its response to increasing CO_2. *Journal of Advances in Modeling Earth Systems*, *11*, 998–1038. https://doi.org/10.1029/2018MS001400

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., et al. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*. https://doi.org/10.5194/esd-2020-66-rc3

Park, S., Shin, J., Kim, S., Oh, E., & Kim, Y. (2019). Global climate simulated by the Seoul National University Atmosphere Model Version 0 with a Unified Convection Scheme (SAM0-UNICON). *Journal of Climate*, *32*(10), 2917–2949.

Qu, X., Hall, A., DeAngelis, A. M., Zelinka, M. D., Klein, S. A., Su, H., et al. (2018). On the emergent constraints of climate sensitivity. *Journal of Climate*, *31*(2), 863–875. https://doi.org/10.1175/jcli-d-17-0482.1

Rong, X. Y., Li, J., Chen, H. M., Xin, Y., Su, J., Hua, L., et al. (2019). The CAMS Climate System Model and a basic evaluation of its climatology and climate variability simulation. *Journal of Meteorological Research*, *32*, 839–861. https://doi.org/10.1007/s13351-018-8058-x

Schlund, M., Lauer, A., Gentine, P., Sherwood, S. C., & Eyring, V. (2020). Emergent constraints on equilibrium climate sensitivity in CMIP5: Do they hold for CMIP6? *Earth System Dynamics*, *11*(4), 1233–1258. https://doi.org/10.5194/esd-11-1233-2020

Sherwood, S. C., Bony, S., & Dufresne, J.-L. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, *505*, 37–42.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*, *58*(4), e2019RG000678. https://doi.org/10.1029/2019RG000678

Su, H., Jiang, J. H., Zhai, C., Perun, V. S., Shen, J. T., Del Genio, A., et al. (2013). Diagnosis of regime-dependent cloud simulation errors in CMIP5 models using "A-Train" satellite observations and reanalysis data. *Journal of Geophysical Research: Atmospheres*, *118*, 2762. https://doi.org/10.1029/2012JD018575

Su, H., Jiang, J. H., Zhai, C., Shen, T. J., Neelin, J. D., Stephens, G. L., & Yung, L. Y. (2014). Weakening and strengthening structures in the Hadley circulation change under global warming and implications for cloud response and climate sensitivity. *Journal of Geophysical Research: Atmospheres*, *119*(10), 5787–5805. https://doi.org/10.1002/2014JD021642

Su, H., Waliser, D. E., Jiang, J. H., Li, J.-l., Read, W. G., Waters, J. W., & Tompkins, A. M. (2006). Relationships of upper tropospheric water vapor, clouds and SST: MLS observations, ECMWF analyses and GCM simulations. *Geophysical Research Letters*, *33*, L22802. https://doi.org/10.1029/2006GL027582

Taylor, K. E., Juckes, M., Balaji, V., Cinquini, L., Denvil, S., Durack, P. J., et al. (2018). CMIP6 global attributes, DRS, filenames, directory structure, and CV's. *PCMDI Document*. Retrieved from https://goo.gl/v1drZl

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*, 485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

Tian, B. (2015). Spread of model climate sensitivity linked to double-intertropical convergence zone bias. *Geophysical Research Letters*, *42*, 4133–4141. https://doi.org/10.1002/2015GL064119

Tian, B., Fetzer, E., Kahn, B., Teixeira, J., Manning, E., & Hearty, T. (2013). Evaluating CMIP5 models using AIRS tropospheric air temperature and specific humidity climatology. *Journal of Geophysical Research*, *118*, D50117. https://doi.org/10.1029/2012JD018607

Tian, B., & Hearty, T. J. (2020). Estimating and removing the sampling biases of the AIRS Obs4MIPs V2 data. *Earth Space Science*, *7*, e2020EA001438. https://doi.org/10.1029/2020EA001438

Waliser, D. E., Li, J.-L. F., Woods, C. P., Austin, R. T., Bacmeister, J., Chern, J., et al. (2009). Cloud ice: A climate model challenge with signs and expectations of progress. *Journal of Geophysical Research*, *114*, D00A21. https://doi.org/10.1029/2008JD010015

Wu, L., Li, Q., Ding, Y., Wang, L., Xin, X., & Wei, M. (2018). Preliminary assessment on the hindcast skill of the Arctic Oscillation with decadal experiment by the BCC_CSM1.1 climate model. *Advances in Climate Change Research*, *9*(4), 209–217. https://doi.org/10.1016/j.accre.2018.12.001

Wu, T., Li, W., Ji, J., Xin, X., Li, L., Wang, Z., et al. (2013). Global carbon budgets simulated by the Beijing climate center climate system model for the last century. *Journal of Geophysical Research: Atmosphere*, *118*, 4326–4347. https://doi.org/10.1002/jgrd.50320

Wu, T., Song, L., Li, W., Wang, Z., Zhang, H., Xin, X., et al. (2014). An overview of BCC climate system model development and application for climate change studies. *Journal of Meteorological Research*, *28*(1), 34–56.

Zhai, C., Jiang, J. H., & Su, H. (2015). Long-term cloud change imprinted in seasonal cloud variation: More evidence of high climate sensitivity. *Geophysical Research Letters*, *42*(20), 8729–8737. https://doi.org/10.1002/2015gl065911