Assessment of the cloud liquid water from climate models and reanalysis using satellite observations

Jui-lin F. Li[1,*], Seungwon Lee[1], Hsi-Yen Ma[2], Graeme Stephens[1], and Bin Guan[1,3]

[1] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA*
[2] *Lawrence Livermore National Laboratory, Livermore, California, USA*
[3] *Joint Institute for Regional Earth System Science and Engineering, University of California, Los Angeles, California, USA*

ABSTRACT

We perform a model-observation comparison and report on the state-of-the-art cloud liquid water content (CLWC) and path (CLWP) outputs from the present-day global climate models (GCMs) simulations in CMIP3/CMIP5, two other GCMs (UCLA and GEOS5) and two reanalyses (ECMWF Interim and MERRA) in comparison with two satellites observational datasets (CloudSat and MODIS). We use two different liquid water observation products from CloudSat and MODIS, for CLWP and their combined product for LWC with a method to remove the contribution from precipitating and convective core hydrometeors so that more meaningful model-observation comparisons can be made. Considering the CloudSat's limitations of CLWC retrievals due to contamination from the precipitation and from radar clutter near the surface, an estimate CLWC is synergistically constructed using MODIS CLWP and CloudSat CLWC. The model-observation comparison shows that most of the CMIP3/CMIP5 annual mean CLWP values are overestimated by factors of 2 - 10 compared to observations globally. There are a number of CMIP5 models, including CSIRO, MPI, and the UCLA GCM that perform well compared to the other models. For the vertical structure of CLWC, significant systematic biases are found with many models biased significantly high above the mid-troposphere. In the tropics, systematic high biases occur at all levels above 700 hPa. Based on the Taylor diagram, the ensemble performance of CMIP5 CLWP simulation shows little or no improvement relative to CMIP3.

1. INTRODUCTION

Representing clouds and cloud climate feedbacks in global climate models (GCMs) remains a pressing challenge to reduce and quantify uncertainties associated with climate change projections (IPCC 2007, 2013). Vertical structures of clouds simulated by present-day models have not been extensively examined using vertically-resolved cloud hydrometers such as ice water content (IWC) and liquid water content (LWC). The rudimentary work done thus far suggests significant biases that the Intergovernmental Panel on Climate Change (IPCC) models have in comparison to observational products in ice water path (IWP) (Waliser et al. 2009) and LWP values (Li et al. 2008, 2011). There is a wide disparity in the cloud ice water path (CIWP) (Waliser et al.

2009; Li et al. 2012) and cloud liquid water path (CLWP) (Li et al. 2008) among the Coupled Model Intercomparison Project phase 3 (CMIP3) models and the Coupled Model Intercomparison Project phase 5 (CMIP5) models.

Clouds strongly influence global climate through their effects on the Earth's radiation budget (e.g., Randall and Tjemkes 1991). The importance of low clouds cannot be overstated as "cloud feedbacks remain the largest source of uncertainty" in determining Earth's equilibrium climate sensitivity, specifically to a doubling of carbon dioxide scenario (IPCC 2007, 2013). Some evidence for this uncertainty are given in Fig. A1 which illustrates a considerable model-to-model disagreement in the CMIP3 liquid water path (LWP; g m^{-2}) in the GCM simulations contributed to the 4th Intergovernmental Panel on Climate Change (IPCC) Assessment Report (20c3m scenario) (Li et al. 2011) and

* *Corresponding author*
　E-mail: jli@jpl.nasa.gov

the 5th Assessment Report (AR5) (Boucher et al. 2013). As a matter of fact, LWP/LWC products are derived from observing systems that possess very different characteristics, such as different sensitivities to cloud and precipitation with different physical assumptions employed in the retrieval process (e.g., Stephens and Kummerow 2007). Despite significant efforts to derive LWP measurements from passive and nadir-viewing techniques, the large optical thicknesses, multi-layer structure, and mixed-phase nature, including the presence of precipitating hydrometeors (e.g., drizzle), of many clouds make the estimates from these techniques very uncertain (e.g., Stephens et al. 2008) and therefore cannot be trusted under a precipitating condition and is removed from this study (see section 2). The ramifications of the poor constraints for cloud water mass, even in terms of total water path, are evident in the model-to-model disagreement for globally-averaged cloud LWPs shown in Fig. A1a. As expected, these differences are exacerbated when considering the spatial patterns of the time-mean values shown in Fig. A1b. The significant disagreement between models for such a fundamental quantity, that has important effects in the context of climate change, must be reduced to improve future model projections of climate.

Before CloudSat was launched in 2006, global observations of cloud water, particularly the vertically-resolved cloud liquid water content (CLWC; mg m^{-3}), were not readily available for model development and validation. The CloudSat mission now provides a considerable leap forward in the information gathered regarding tropospheric cloud mass as well as other macrophysical and microphysical properties (e.g., Stephens et al. 2008). CloudSat's cloud profiling radar capabilities provide a new view of the global and vertical structure of clouds, in particular, the vertical structure of cloud condensate. It is worth noting that, for both the passive and active satellite retrievals and for some models, it is understood that "liquid water content" (LWC) should represent all liquid hydrometeors, and should include suspended cloud liquid and liquid mass in precipitating forms such as rain or drizzle. However, these observations are an altogether new resource, albeit with uncertainties and limitation (e.g., Li et al. 2008).

Li et al. (2012) pointed out that, for improving GCM representation of clouds and cloud climate feedback, considerable care and caution have to be taken in order to make judicious comparisons between the GCM representations of (typically only the) clouds and the satellite observations that are an inherent combination of the clouds and falling hydrometeors (e.g., rain or drizzle or cloud ice or falling snow). Such considerations include taking steps to make a sensible comparison for CLWC or CLWP and cloud ice water content (CIWC) or path (CIWP) (e.g., Li et al. 2012) or taking the steps needed for a viable comparison in terms of reflectivity or radiance, which typically means to use satellite "simulators" (e.g., Klein and Jakob 1999; Webb et al.

2001; Delanoë and Hogan 2010; Bodas-Salcedo et al. 2011; Delanoë et al. 2011). However, such distinctions are often not clearly made, and certainly not always made consistently between satellite retrievals, model parameterizations and/or output from models. Because there is no precipitating liquid particle included as a prognostic parameter in most GCMs, the direct comparison with observation may not be appropriate.

In this study, we take the approach in Li et al. (2012), and perform the evaluation in terms of the model representations of CLWC/CLWP utilizing the experience we have gained from cloud ice and liquid (e.g., Li et al. 2005, 2007, 2008, 2011, 2012; Waliser et al. 2009). This includes developing a measure of observational uncertainty (discussed in section 2), and applying an illustrative and quantitative set of evaluation diagnostics. In particular, a new set of CLWC is acquired by the construction of vertical profiles of LWC from MODIS LWP combined with CloudSat LWC. The augmentation with MODIS based LWC with CloudSat LWC is to overcome the limitations of CloudSat-derived LWC/IWC retrievals from radar clutter near the surface (lower two to four bins radar signal) and/or in the present of precipitation. A prominent goal of the study is to examine how the fidelity of the models may have changed from CMIP3 to CMIP5. Moreover, we attempt to discriminate CMIP5 models that achieve a threshold capability of model fidelity by using the Taylor diagram (Taylor 2001) and by taking into account the observational uncertainties (see section 2). In addition, as reanalysis products have become nearly synonymous, in some contexts, with "observations", we also incorporate two recent reanalysis products in our study to provide some assessment of this tenuous perception - particularly for quantities such as CLWP and CLWC that are not strongly constrained by observations.

In sections 2 and 3, we describe the observational resources used in this study, including their retrievals and the methodologies to obtain an observational estimate with some quantitative assessment on uncertainty. In section 4, we briefly describe the models and reanalysis data sets utilized in this evaluation study. In section 5, we illustrate and discuss the results of our model evaluation. Section 6 summarizes the results and draws conclusions.

2. OBSERVED ESTIMATES OF LWC AND LWP

The cloud liquid water observations employed in this study are derived from visible and the near infrared passive observations of the A-Train MODIS and the active space borne radar observations of the A-Train CloudSat (Stephens et al. 2002). Each LWP product considered are derived from observing systems that possess very different characteristics, such as different sensitivities to cloud and precipitation, dynamic range, representativeness and different physical assumptions and other model data employed in the retrieval

process (e.g., Stephens and Kummerow 2007). As a consequence, the liquid water content information inferred from each system is distorted by the retrieval process itself and interpretation of these products and their comparison warrants a note of caution. Despite the differences, under idealized conditions (e.g., warm, homogeneous, single-layered, fully overcast clouds with no precipitation) there is a high degree of agreement among the products used at least over the range of liquid water path between 20 g m^{-2} and about 200 g m^{-2} (e.g., Lin and Rossow 1994, 1997; Borg and Bennartz 2007).

2.1 MODIS Cloud LWP

LWP data used are from the MYD06_L2 product. The MODIS LWP is estimated from cloud optical thickness and droplet effective radius, which are inferred from solar reflectance at 0.64 μm visible band and at one of the water-absorbing near-infrared bands located at 1.6, 2.2, and 3.7 μm (Platnick et al. 2003). The MODIS measurement is based on reflected sun light both over land and ocean and only available for daytime. Details of uncertainties and limitations for MODIS LWP retrieval are described in Appendix A section 1.

The MODIS LWP retrievals are collocated to the CloudSat profiles by finding the nearest neighbor MODIS measurement for each CloudSat footprint location. See Appendix A section 2 for details of the co-location data generation process. The collocation was performed in order to apply the CloudSat-based precipitation and cloud condition flags to the MODIS LWP data to partition these data into the suspended (cloud-only), convective and precipitating portions. The time period of this data used in this study is from May 2008 to April 2010.

2.2 CloudSat CPR LWP and LWC

CloudSat provides vertical profiles of radar reflectivity measured by a 94 GHz cloud profiling radar (CPR) with a minimum sensitivity of ~ -30 dBZ. The profiles extend between the surface and 30 km altitude with a vertical resolution of 240 m and having a footprint of about 2.5 km along track and 1.4 km cross track. To date, two official retrieval products for LWC/LWP are available from the CloudSat data processing center: 2B-CWC-RO and 2B-CWC-RVOD. While 2B-CWC-RO uses only measured radar reflectivity, 2B-CWC-RVOD uses both the radar reflectivity and the visible optical depth retrieved from CloudSat and MODIS measurements together (Austin et al. 2009). For this study, the 2B-CWC-RO4 (Austin et al. 2009) data are used and the detailed sensitivity and uncertainty of this retrieval algorithm are discussed in Austin et al. (2009). The time period of the data used in this study is from January 2007 to December 2010. The interpretation of reflectivity as cloud liquid water is straightforward in the absence of precipitation. But when present larger particle such as precipitation, it grossly distorts the estimate of LWC due to the high sensitivity of radar reflectivity to the presence of large particles. Other shortcomings of the CloudSat data include (1) the effects of ground clutter that mask the lowest kilometer and thus the liquid water content of a significant fraction of low clouds is undetected, (2) the ambiguity of mixed phased clouds and deep convection on radar reflection makes interpretation of reflectivity in terms of liquid water content problematic.

3. SYNERGISTICALLY CONSTRUCTED LWC USING MODIS LWP AND CLOUDSAT LWC

Each of the different cloud liquid water data products described above has particular limitations especially when precipitation is present. For a meaningful comparison between the satellite-estimated and model-simulated LWP, the LWP observation information derived for convective/ precipitating liquid water clouds should be rejected and removed. Convective clouds and precipitation were identified using an approach referred to as the FLAG method (Li et al. 2008, 2012; Waliser et al. 2009) that rejects all the retrievals in any profile that are flagged as precipitating (and drizzle) at the surface (from CloudSat 2C-PRECIP-COLUMN data) and exclude any retrieval within the profile whose cloud type is classified as "deep convection" or "cumulus" (from CloudSat 2B-CLDCLASS data). By excluding these portions of the liquid mass, we obtain an estimate of the cloud-only portion of the LWP/LWC (hereafter, referred to as CLWP/CLWC). This methodology of estimating CLWP/ CLWC and CIWC/CIWP was used in our previous CMIP3 model-data comparisons on LWC and IWC (e.g., Li et al. 2008, 2011, 2012; Waliser et al. 2009).

MODIS and CloudSat CLWP are further filtered out mixed-phase cloud condition because its partitioning of cloud water into liquid and ice is unreliable as described above. For CloudSat, a collocated ECMWF temperature profile (from CloudSat ECMWF-AUX data) is used to determine whether a given cloud water content profile contains mixed-phase clouds. If the temperature profile contains a section with temperature higher than -20°C and lower than 0°C, we use the CloudSat LWC retrieval algorithm to partition the cloud water content into liquid and ice. Since the reliability of the partitioning method is questionable, we filter CloudSat retrieved LWC profiles that contain a section with temperature higher than -20°C and lower than 0°C. For MODIS, the cloud phase is determined by utilizing distinct differences in bulk absorption characteristics between water and ice at infrared wavelengths. This method allows MODIS to distinguish supercooled clouds from mixed-phase clouds or ice containing clouds. Because of this distinction, the MODIS CLWP filtered for no-mixed phase cloud condition captures the supercooled liquid water

path, while the CloudSat CLWP filtered for no-mixed phase cloud condition does not.

The filtered CLWP of MODIS and CloudSat are shown in Figs. 1c and d. Under these conditions of non-convective, non-precipitating clouds, the liquid water path data from MODIS and CloudSat broadly agree with each other except in regions of low stratiform clouds and mixed-phase clouds (mid-latitude storm tracks and high latitudes). Also shown in Fig. 1 is the MODIS LWP retrieved under all conditions (Fig. 1a) and under the conditions of precipitation and convective clouds (Fig. 1b). Since CloudSat does not provide meaningful information in the presence of precipitation, only data from MODIS are shown. Since MODIS is relatively insensitive to precipitation, the difference between the non-precipitating and precipitating LWP shown is interpreted as a measure of the increased cloud LWP associated with precipitating clouds. Precipitating clouds tend to be deeper and contain more cloud liquid than non-precipitating clouds (Stephens et al. 2008). The small difference between MODIS all condition LWP (Fig. 1a) and MODIS filtered LWP (Fig. 1c) confirms that MODIS is relatively insensitive to precipitation.

Figure 2 shows annual, zonally averaged LWP quantities obtained from the three observational estimates: (1) MODIS all condition LWP, (2) MODIS non-convective non-precipitating no-mixed phase cloud LWP, (3) CloudSat non-convective non-precipitating no-mixed phase cloud. The CLWP estimates, in general, agree relatively well between CloudSat and MODIS in the tropical and subtropical regions but they differ significantly in the mid- and high-latitudes due to the differences in the corresponding no-mixed phase cloud conditions as mentioned above.

Another observational reference we prepare for this study is a liquid water content profile that is synergistically constructed using MODIS LWP and CloudSat LWC. As illustrated in Figs. 1 and 2, CloudSat filtered CLWP is a small portion of MODIS filtered CLWP globally (12.7 g m^{-2} versus 35.8 g m^{-2} in global averages) and especially in high latitudes and in stratocumulus regions because it missed supercooled clouds and low topped clouds. Figure 1 also shows that the contribution from precipitating clouds is expected to be about 10% of the cloud liquid water path in all conditions (5.4 g m^{-2} out of 41.1 g m^{-2} in global averages). Therefore, CloudSat LWC by itself is not sufficient enough to provide the vertical profile of CLWC for the comparison with models. We complement the CloudSat LWC with LWC derived from MODIS LWP values in the CloudSat missing conditions. Namely, we construct CLWC for precipitating clouds, LWC for low-topped clouds, and LWC for supercooled clouds. All of the clouds are missed in CloudSat filtered LWC we discussed above.

In order to construct LWC from MODIS LWP, we need to define the cloud top height, cloud base height, and cloud water content vertical structure. We use the cloud top height determined by CALIPSO lidar and CloudSat radar combined retrieval algorithm (from CloudSat GEOPROF-LIDAR data). Details for finding LCL are described in Appendix A section 4.

We select cloud pixels that are not included in the CloudSat CLWC filtering condition. For the cloud pixels, we use MODIS LWP to construct LWC using the method described above. The MODIS derived LWC are decomposed into several conditions so that the relative contributions from the different conditions can be quantified. The MODIS LWC data are first divided into two groups: (1) low-topped clouds with the cloud top height less than 1 km and (2) mid-topped clouds with the cloud top height larger than 1 km. This division is to test how much the low-topped clouds are missed by CloudSat and the contribution of the low-topped clouds to the overall CLWC. The mid-topped clouds are further divided into two groups: (1) precipitating clouds and (2) non-precipitating clouds. This division shows the relative contribution from precipitating clouds, which are less well constrained by both MODIS and CloudSat satellite observations than non-precipitating clouds. The uncertainty of the overall CLWC is largely affected by the amount of water content from the precipitating clouds. Knowing the relative contribution of the precipitating clouds even though the absolute values are not reliable, that is very helpful in the uncertainty estimation.

Note that we use MODIS as complementary data to CloudSat by adding MODIS data when CloudSat data are not available or unreliable. This means that the three MODIS cloud conditions considered are clouds that are not detected or reliably retrieved by CloudSat. Therefore, each of the three MODIS cloud conditions is not the same as all clouds detected by MODIS in that condition. For example, MODIS non-raining mid-level clouds in this study does not represent all MODIS detected non-raining mid-level clouds. It instead represents MODIS detected but CloudSat non-detected/non-retrieved non-raining mid-level clouds.

Figure 3a shows the annual mean and global average of LWC quantities obtained with the method described above. In order to see the relative contribution of the different cloud conditions to the overall CLWC, the CLWC of each cloud condition is separately plotted. Globally, the contribution of CloudSat CLWC is about 25% of the total estimated CLWC (black line, the sum of all the CLWCs from the subgroups). The low cloud contribution missed by CloudSat is significant near 900 hPa. The contribution from precipitating clouds is relatively small but is a measurable amount that can give a systematic negative bias to the observational estimation if neglected. Finally, the non-precipitating clouds that are missed by the CloudSat filtered CLWC contribute significantly. The non-precipitating clouds are mainly composed of the supercooled clouds that CloudSat cannot retrieve well because of its retrieval algorithm of artificially partitioning cloud water contents into liquid and ice when

the temperature is between -20 and 0°C, even though the clouds are supercooled.

Figure 3b shows the annual, zonal mean of relative contributions of LWC quantities by comparing the corresponding LWP values in the subgroups. Figure 3c shows the relative occurrence frequency (ROF) of the cloud subgroups. All the three MODIS cloud subgroups show the increased contribution to liquid water path (therefore content as well) in the higher latitudes. This is a direct consequence of the no mixed-phase condition used to filter the Cloud-

Sat clouds. The largest contribution missed by CloudSat is no-rain mid-level clouds, followed by low clouds, and rain mid-level clouds. Comparing Figs. 3b and c, the order of the relative size of the liquid water path does not always follow the order of its ROF. For example, the CloudSat cloud ROF is lower than the MODIS no-rain mid-level cloud; the CloudSat cloud LWP is larger than the MODIS no-rain mid-level cloud LWP in the tropics. This suggests that the no-rain mid-level clouds missed by CloudSat are mainly tropical low-value LWP clouds.
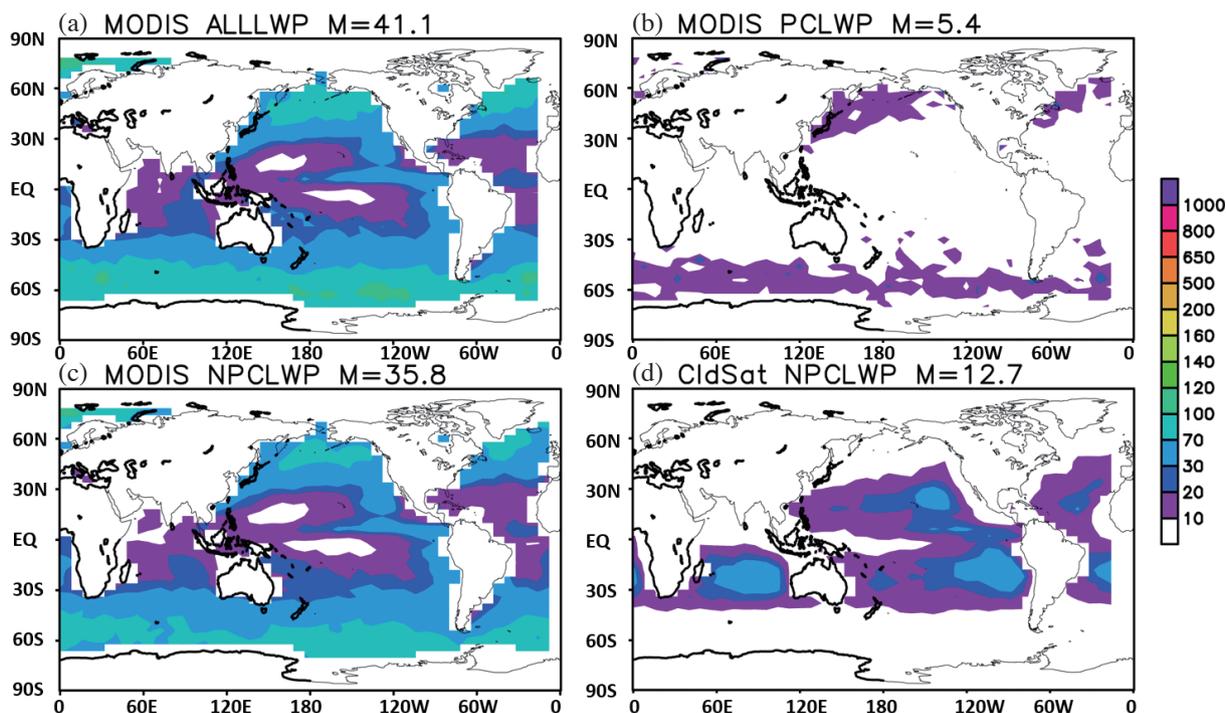
Fig. 1. Annual mean maps of (a) MODIS-based cloud liquid water path (CLWP) under all conditions, (b) MODIS CLWP under convective and precipitating conditions, (c) MODIS CLWP under non-convective and non-precipitating conditions, and (d) CloudSat CLWP under non-convective, non-precipitating, and non-mixed-phase conditions. The unit for these values is g m^{-2}. All the retrievals are used only over the oceans.
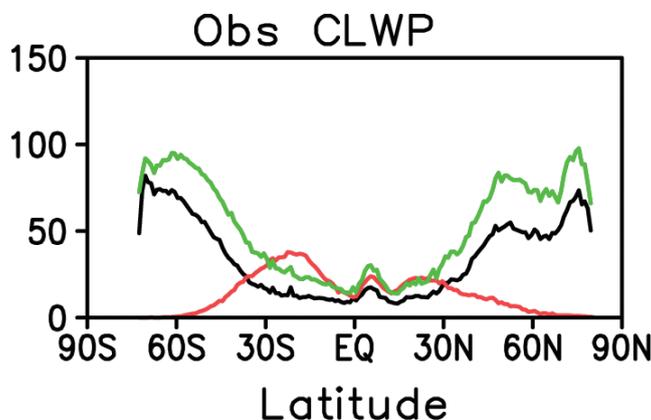
Fig. 2. Zonal average of cloud liquid water path (CLWP) quantities associated with the three observational estimates shown in Fig. 1: the CloudSat warm liquid cloud only with no rain and convection conditions (red), MODIS no rain and no convection conditions CLWP (black), and MODIS with all conditions CLWP (green). The unit is g m^{-2}.
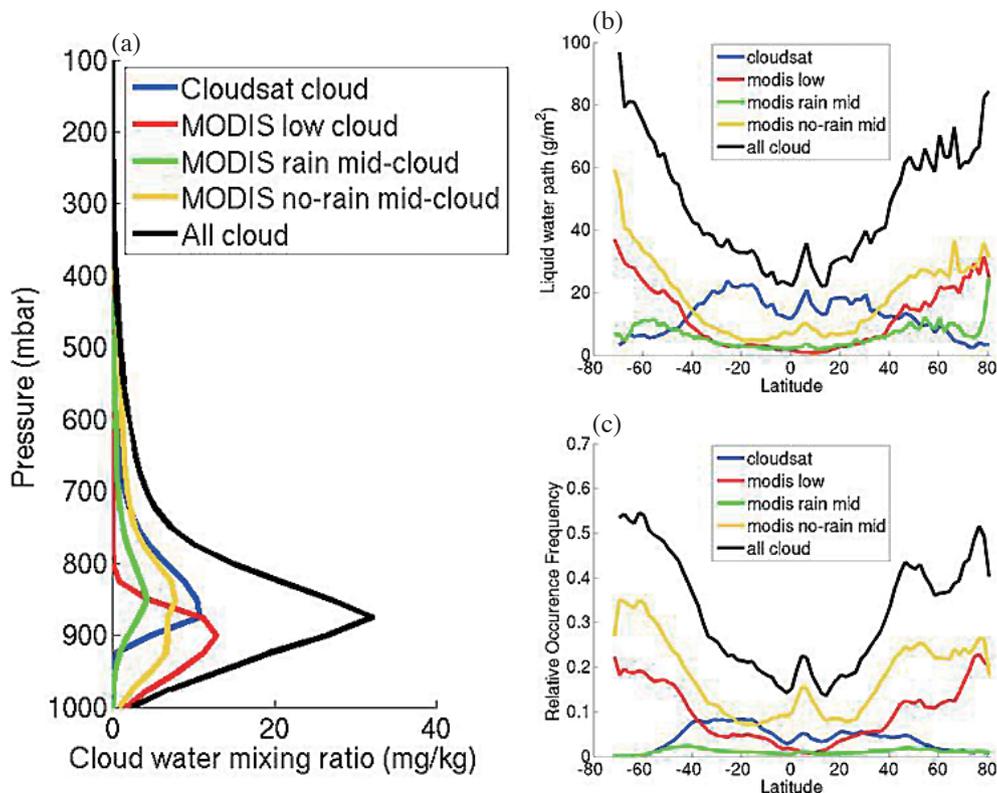
Fig. 3. (a) Liquid water content prepared from CloudSat and MODIS retrievals (black), partitioned into four cloud conditions: CloudSat-based warm non-precipitating, non-convective clouds (blue); MODIS-based non-convective, low-topped clouds (red); MODIS-based precipitating, non-convective, mid-topped clouds (green); and MODIS-based non-precipitating, non-convective, mid-topped clouds (yellow). (b) Liquid water path of each of the four cloud conditions contributing to the liquid water content. (c) Relative occurrence frequency of each of the four cloud conditions partitioned.

Overall, it is very clear that the complementary use of CloudSat and MODIS is critical in estimating the LWC because many subgroups of clouds are not well retrieved by CloudSat alone. The contributions of the clouds missed (meaning either non-detected or not retrieved reliably) by CloudSat are not negligible in all latitudes and especially at high latitudes. The method we used to construct the liquid water content using MODIS is not perfect because the information available in these conditions is limited (i.e., no direct information about vertical distribution of the cloud mass). However, the zero-th order estimation of LWC from all cloud types is a useful observational reference to use for comparisons with models.

Apart from the uncertainty of the retrieval method, an additional uncertainty to consider in light of making model-observation comparisons concerns the differences in the spatial and temporal sampling between the observations and the GCMs, such as those in the CMIP archives.

Li et al. (2012) and Guan et al. (2013) found that the bias introduced by the satellite sampling of cloud water is negligible, which is within 3% of the standard deviation of the unsampled data. It is plausible to compare the observed, satellite-sampled, liquid water estimates to those

from the GCMs without the need to sample the GCMs along the A-Train satellite track (cf. Jiang et al. 2012).

4. MODELED VALUES OF LWC AND LWP

On the modeling side, LWC is usually a prognostic variable based on a balance equation contributed by large-scale advection and parameterizations of subgrid-scale convective cloud, shallow cumulus, and stratocumulus. LWP is obtained as the vertical integral of LWC. The models examined in this study, except for GFDL-CM3, do not include liquid water mass from precipitating rain and/or convective-type clouds in their LWC. Therefore, we consider the model LWP/LWC as CLWP/CLWC. Following Li et al. (2012) and using the observations described in section 2, we evaluate CLWP/CLWC in ECMWF (ERA-Interim, Dee et al. 2011) and NASA MERRA reanalyses, coupled atmosphere-ocean GCMs (CGCMs) from CMIP3 (for CLWP only), CGCMs from CMIP5, and two additional state-of-the-art GCMs: the UCLA GCM (Ma et al. 2013) and the NASA GEOS5 GCM. The CMIP3 simulations are the same as those described in Li et al. (2008, 2011, 2012) - although excluding the two UKMO models which we have learned that the provided

output on CLWP was incompatible with the CMIP3 output specifications (cf. Li et al. 2011). In CMIP3 and CMIP5, the intended meaning of "clwvi" is total water path, i.e., ice plus liquid. Our investigation (Li et al. 2011; referred to LET) has determined that thirteen of the CMIP3 GCMs (labeled in LET as: ccc_ma63, cnrm, csiro, gfdl, iap, ipsl, mirochr, gisseh, gisser, inmcm, mpi, ukmogem, and ukmocm) provided output that was consistent with the intended interpretation [i.e., the total water path (TWP)], while three of the GCMs (labeled in LET as: bccr, csiro, ncar) provided the output with the interpretation that the quantity was just associated with cloud liquid water path (i.e., CLWP). Two GCMs output from ukmocm3 and ukmogem, the total water path (TWP) to all CMIP3 did not include water associated with the convection scheme.

In CMIP5, twelve GCMs (bcc, bccesm, CanESM2, inmcm4, inmcm4_esm, CNRM, GISSE2R, GISSE2H, MRI, NorESM1, GFDL-CM3) provided output that was consistent with the intended interpretation of "clwvi" (i.e., TWP), eight of the CMIP5 GCMs (CCSM4, CSIRO, IPSL, MPI, MIROc4h, MIROC5, MIROC-ESM, MIRO-ESM-CHEM) provided the output with the interpretation that the quantity was just associated with cloud liquid water (i.e., CLWP). Given this situation, we derived unfiltered CLWP for those models in CMIP3 and CMIP5 that provided TWP using the following relationship: LWP = TWP - IWP (Cloud Ice Water Path).

Table 1a lists the CMIP5 simulations included in this study, and Table 1b describes a summary of cloud microphysics parameterizations used in the selected CMIP5 models. As mentioned in Li et al. (2012), the performance of simulated warm cloud properties in CMIP5 models arises from a highly coupled system (land, ocean, atmosphere etc.) and the behavior is not likely to be simply explained by any single component/scheme, but rather by details of the model's specific schemes and the coupling among schemes related with a particular process such as clouds, aerosols and turbulence for boundary layer clouds, and clouds and convection for deeper clouds, as well as the interactions with sea surface temperatures (SSTs). However, we attempt to explain the causes in the behavior of some of the best/worst performing models in section 5.

The specific experimental scenario is the historical 20th century simulation, which used the observed 20th century greenhouse gas, ozone, aerosol, and solar forcing. The time period used for the long-term mean is 1970 - 2005, and if a model provided an ensemble of simulations, only one of them was chosen for this evaluation.

One cautionary remark for the model-observation comparison is regarding the difference between the water mass from cloud particles in precipitating conditions and the water mass from precipitating particles in precipitating conditions. These two water mass sources are different in terms of contributing particle sizes. In the observation side, both water mass sources are removed because any retrievals

from the precipitating condition are filtered out in the FLAG filtering method described in section 2. In the model side, all of the models examined in this study include the water mass from cloud particles in precipitating conditions but only the two GFDL models include the water mass from precipitating particles in addition to that from cloud particles.

Ideally, for the comparison with the non-precipitating CloudSat and MODIS observational estimates, precipitating profiles from the model should also be excluded. However, at this stage it is not feasible because model precipitation is parameterized and "averaged" in time and grid box (contrast to a snapshot CloudSat cloud profile) over a physical model time step (~30 mins to hour) with significantly coarser resolution (~50 - 100 km) than the size of the instantaneous "snapshot" CloudSat footprint (~1 - 2 km) or MODIS pixel (1 km). Thus, in a GCM, the determination of a threshold for non-precipitation and/or convective mass that is equivalent to CloudSat footprint/MODIS pixel resolution is difficult.

Given this situation, we use the filtered MODIS and CloudSat observational estimates of LWP as a lower limit. At the same time, the MODIS unfiltered CLWP can serve as a upper limit since it includes cloud liquid water (i.e., no precipitating particle) in precipitating conditions but MODIS retrievals in the condition are known to be less reliable and overestimate LWP in comparison with in situ measurements (King et al. 2013 and see section 2 for discussion). Additionally, we have combined the MODIS and CloudSat cloud retrievals to construct the CLWC profile for all conditions (see section 2 for discussion), which include cloud water profiles in the precipitating conditions. These estimates should be used with caution because the uncertainty/errors in the retrievals from the precipitating conditions is large (> 100%).

Unlike all the other models examined in this study, which do not include liquid mass from precipitating rain and convective-type clouds in their CLWC, the two GFDL models include the liquid mass by adding grid means over shallow cumulus, deep cumulus cells, and convective mesoscale clouds, weighted by their respective area fractions. Thus, the GFDL models should be considered somewhat carefully with respect to the other models, and their CLWC/ CLWP fields would be more commensurate with the total liquid water content (TLWC)/ and total liquid water path (TLWP), which include water mass from both cloud particles and precipitating particles. Currently, we do not have reliable global observational datasets to estimate TLWP because available observation instruments are either sensitive to only cloud particles (as in MODIS) or are more sensitive to the precipitating particles so that the contribution from cloud particles is not retrievable reliably (as in CloudSat).

For both the GCM and observational data sets, all fields have been re-gridded to 40 levels in vertical (with a constant pressure interval of 25 hPa) and mapped onto common 8° × 4° longitude by latitude grids.

Table 1. (a) Institution, model resolution and label for the CMIP5 GCMs examined in this study. (b) Outline of cloud microphysics parameterizations used in the selected CMIP5 models, as well as UCLA GCM, GEOS5, MERRA, and ERA-Interim model.

(a)

Model Label	Model resolution	Institution/Model Name
BCC	64 × 128 × 26	Beijing Climate Center, China/BCC-CSM1-1
BCC_ESM	64 × 128 × 26	Beijing Climate Center, China/BCC-CSM1-1_esm
CanESM2	64 × 128 × 35	Canadian Centre for Climate Modelling and Analysis, Canada/CanESM2
CCSM4	288 × 192 × 26	National Center for Atmospheric Research, USA/CCSM4
CNRM-CM5	128 × 256 × 17	Centre National de Recherches Meteorologiques, France/CNRM-CM5
CSIRO	96 × 192 × 18	Australian Commonwealth Scientific and Industrial Research Organization, Australia/CSIRO-Mk3-6-0
GISS-E2-H	90 × 144 × 29	NASA/Goddard Institute for Space Studies, USA/GISS-E2-H
GISS-E2-R	90 × 144 × 29	NASA/Goddard Institute for Space Studies, USA/GISS-E2-R
Inmcm4	120 × 180 × 21	Institute for Numerical Mathematics, Russian/Inmcm4
Inmcm4ESM	120 × 180 × 21	Institute for Numerical Mathematics, Russian/Inmcm4_esm
IPSL	96 × 96 × 39	Institute Pierre Simon Laplace, France/IPSL-CM5A-LR
MIROC/MIROC-CHEM	64 × 128 × 80	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC-ESM-CHEM
MIROC4h	320 × 640 × 56	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC4h
MIROC5	128 × 256 × 40	University of Tokyo, NIES, and JAMSTEC, Japan/MIROC5
MRI	160 × 320 × 35	Meteorological Research Institute, Japan/MRI-CGCM3
MPI	96 × 192 × 47	Max Planck Institute for Meteorology/MPI-ESM-LR
NorESM	96 × 144 × 26	Norwegian Climate Centre, Norway/NorESM1-M

(b)

Models	Prognostic cloud variables	Bulk single moment or double moment	Cloud fraction (PDF based or Non-PDF based)	References
UCLA GCM	Mixing ratio of cloud liquid and ice	Bulk single moment	Diagnostic, Non-PDF based, 0 or 1	Köhler (1999); Ma et al. (2012)
ERA-Interim (CY31r1)	Mixing ratio of cloud condensate	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -15°C and liquid at T = 0°C).	Prognostic cloud fraction; PDF-based	Tiedtke (1993); Dee et al. (2011)
GEOS5	Single mixing ratio of cloud condensate	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning; "anvil" cloud, originates in detraining convection. "Large-scale cloud", originates in a probability distribution function (PDF) based condensation calculation.	Diagnostic PDF based	Rienecker et al. (2008)

Table 1. (b) (Continued)

Models	Prognostic cloud variables	Bulk single moment or double moment	Cloud fraction (PDF based or Non-PDF based)	References
MERRA	Single mixing ratio of cloud condensate	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning; "anvil" cloud, originates in detraining convection. "Large-scale cloud", originates in a probability distribution function (PDF) based condensation calculation.	Diagnostic, PDF based	Rienecker et al. (2008)
BCC/BCC_ESM	Single mixing ratio of total water	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -40°C and liquid at T = -10°C).	Diagnostic, non-PDF based	Rasch and Kristjánsson (1998); Sundqvist (1988); Zhang et al. (2003); Collins et al. (2004)
CanESM2	Mixing ratio of cloud liquid and ice	Bulk single moment	Diagnostic, PDF based	Lohmann and Roeckner (1996); Chaboureau and Bechtold (2005)
CCSM4	Single mixing ratio of total water	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -40°C and liquid at T = -10°C).	Diagnostic, non-PDF based	Rasch and Kristjánsson (1998); Zhang et al. (2003); Collins et al. (2004); Boville et al. (2006)
CNRM-CM5	Diagnostic cloud liquid water, PDF based	Statistical formulation	Diagnostic, PDF based	Bougeault (1981, 1982); Ricard and Royer (1993)
GISS-E2-H	Single mixing ratio of total water; Diagnostic falling snow	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -35°C and liquid at T = -4°C over ocean; T = -35°C and liquid at T = -10°C over land).	Diagnostic, non-PDF based	Del Genio et al. (1996)
GISS-E2-R	Single mixing ratio of total water; Diagnostic falling snow	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice T = -35°C and liquid at T = -4°C over ocean; T = -35°C and liquid at T = -10°C over land).	Diagnostic, non-PDF based	Del Genio et al. (1996)
Inmcm4	Mixing ratio of cloud liquid and ice	Bulk single moment; Large scale condensation in the case of relative humidity exceeds 1.	Diagnostic, non-PDF based	Volodin et al. (2010)

Table 1. (b) (Continued)

Models	Prognostic cloud variables	Bulk single moment or double moment	Cloud fraction (PDF based or Non-PDF based)	References
Inmcm4ESM	Mixing ratio of cloud liquid and ice	Bulk single moment; Large scale condensation in the case of relative humidity exceeds 1.	Diagnostic, non-PDF based	Volodin et al. (2010)
IPSL-CM5A-LR	Single mixing ratio of total water	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -15°C and liquid at T = 0°C).	Diagnostic PDF based	Bony and Emanuel (2001)
MIROC (MIROC-ESM; MIROC-ESM-CHEM)	Mixing ratio of cloud liquid and ice	Bulk single moment; different phases determined by temperature	Diagnostic PDF scheme with minor change for calculating anvil cloud	Le Trent and Li (1991); Hourdin et al. (2006); Ogura et al. (2008)
MIROC4h	Mixing ratio of cloud liquid and ice	Bulk single moment; different phases determined by temperature	Diagnostic PDF scheme with minor change for calculating anvil cloud	Le Trent and Li (1991); Ogura et al. (2008)
MIROC5	Mixing ratio of cloud liquid and ice	Bulk single moment; CCN is prognostic	Prognostic PDF scheme	Wilson and Ballard (1999); Watanabe et al. (2009)
MRI-CGCM3	Mixing ratio of cloud liquid and ice	Double moment scheme	Diagnostic PDF based	Tiedtke (1993); Yukimoto et al. (2011)
NorESM1	Single mixing ratio of total water	Bulk single moment; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -40°C and liquid at T = -10°C).	Diagnostic, non-PDF based	Rasch and Kristjánsson (1998); Zhang et al. (2003); Boville et al. (2006)
GFDL-CM3	Mixing ratio of cloud liquid and ice	Bulk single moment	Prognostic cloud fraction; PDF based; The Tiedtke (1993) parameterization includes among its assumptions a simple PDF for humidity in clear air (uniform distribution). The Donner (2011) cumulus parameterization generates a discrete PDF (defined by its ensemble of cumulus elements), each member of which has a joint distribution of thermodynamic and microphysics properties and vertical velocities.	Tiedtke (1993); Rotstayn (1997); Donner (2011)
CSIRO-Mk3.6.0	Mixing ratio of cloud liquid and ice; Diagnostic falling snow	Bulk single moment; ice crystal number concentration is diagnosed; mixing ratio of cloud condensate with temperature dependent partitioning (The bounds are adjustable constants with current settings ice at T = -40°C).	Diagnostic, PDF based	Rotstayn (1997); Rotstayn et al. (2000)

5. RESULTS

Figure 4 shows the long-term annual mean spatial distributions of simulated values of CLWP from eighteen conventional CMIP5 CGCMs (see Table 1; 1 - 18: bcc, bcc-esm, CanESM2, CCSM4, CNRM, CSIRO, GISSE2H, GISSE2R, Inmcm4, Inmcm4esm, IPSL, MIROC4h, MIROC5, MIROCCHEM, MIROC, MPI, MRI, and NorESM1), the multi-model ensemble mean (MMM) from the eighteen CMIP5 models (Fig. 4s), and UCLA GCM (Fig. 4t), GEOS5 (Fig. 4u), two analyses MERRA (Fig. 4v) and ECMWF-Interim (Fig. 4w), as well as observed estimates of (Fig. 4y) CloudSat filtered CLWP, (Fig. 4z) MODIS filtered CLWP, and (Fig. 4aa) MODIS unfiltered CLWP. We use the MODIS filtered CLWP (Fig. 4z) as the reference, the CloudSat filtered CLWP (Fig. 4y) as a "lower" limit, and the MODIS unfiltered CLWP as an upper limit for the model evaluations of CLWP. In addition, Fig. 4 shows the TLWP distributions from GFDL CM3 (Fig. 4x).

Overall, the multi-model mean CMIP5 CLWP values (Fig. 4s) are similar to observations in terms of spatial distribution, but they are biased high globally even when compared against the MODIS unfiltered CLWP. Individually, most models tend to qualitatively capture the global and regional CLWP patterns. This includes the relatively high values of CLWP in the storm tracks from the subtropics to high latitudes. However, they all overestimate CLWP over ITCZ and warm pool. Note that the relative magnitudes between tropical and mid-latitude values can be quite different across models.

None of the CMIP5 models provides a good representation of both the magnitude and spatial pattern of CLWP. The three that perform relatively better are CSIRO, Inmcm4, and Inmcm4esm but even these have significant shortcomings relative to the observation data. Most of the models overestimate (~a factor of 2 or more) tropical CLWP. The MIROC, MIROCCHEM, and CCSM4 GCMs greatly overestimate (~a factor of 5) tropical and storm track CLWP. The IPSL, CSIRO, MIROC5, MIROC4h, and the two GISS GCMs moderately overestimate CLWP in the extra-tropics. For the non-CMIP5 GCMs, the GEOS5 atmospheric GCM (AGCM) overestimates (~a factor of 2) CLWP in the storm tracks while the UCLA GCM does relatively well over most of the globe except over the northeast of South America. The two analyses, ECMWF and MERRA, show relatively good spatial patterns of CLWP patterns, with both being biased high (factor ~1 - 2). The GFDL model simulates and provides output on TLWP. When compared with the observational estimates of CLWP (Figs. 4y - aa), GFDL exhibits in a relatively good agreement in the extra-tropics storm track regions but its TLWP is larger than the observational CLWP in the tropical ITCZ and warm pool. Since the observational CLWP does not include water mass from deep convective cores and precipitating hydrometeors, it is uncertain whether it indicates the model biases of GFDL TLWP estimations.

Figure 5 shows the long-term annual zonal average of CLWP quantities associated with CMIP5 displayed in Fig. 4. Figure 5a represents the multi-model mean (blue line), the one standard deviations for upper and lower bound (red line). The latitudinal distributions clearly illustrates a wide spread of CLWP in the CMIP5 models. Figure 5b shows the observational estimations of CLWP from CloudSat and MODIS. The green/black line is the MODIS unfiltered/filtered CLWP which includes/excludes the cloud water mass from both convective cores and precipitating clouds. The magenta line is the CloudSat filtered CLWP which excludes the cloud water mass from convective cores and precipitating clouds. The multi-model mean (the blue line) and the model standard-deviation added and removed to the mean (the red lines) are also plotted for comparison with the observational estimates. It is evident that the multi-model mean CLWP is one standard deviation larger than the observational MODIS unfiltered CLWP and filtered CLWP. Since the MODIS unfiltered CLWP is the upper limit of the observational CLWP, this model-observation comparison illustrates the model CLWP means are significantly larger than the observational CLWP.

To summarize the multi-model performance of CMIP3 and CMIP5 in representing the time-mean pattern of CLWP, Fig. 6 illustrates the multi-model mean biases against the observed estimate calculated across the model ensembles. The observational estimate used in this calculation is the MODIS CLWP filtered for no rain and no convection conditions. Note that the MODIS filtered CLWP is provided only over oceans so the ensemble mean is obtainable only over oceans. The patterns of CLWP bias in CMIP3 (Fig. 6a) and CMIP5 (Fig. 6b) are systematically similar: CLWP is overestimated globally. In CMIP3, the ITCZ/SPCZ, high latitudes, and storm tracks are significantly biased high (65.2 g m^{-2}), while in CMIP5 the bias (78.3 g m^{-2}) is even higher. Therefore, based on this initial comparison, the fidelity of CMIP5 models in representing cloud liquid mass exhibits no progress relative to CMIP3.

The CMIP3 and CMIP5 multi-model biases against the observational estimates of CLWP have implications on radiative flux calculations in the models. The radiative fluxes are affected by LWP from all types of hydrometeors, but all the models in CMIP3 and CMIP5 (except GFDL) include only the contribution from cloud particles excluding the contribution from precipitating and deep convective hydrometeors. Therefore, the underestimation of water mass by including only cloud-only water mass can lead to systematic biases in radiative fluxes. The present study run shows that the cloud-only water mass in the model is significantly overestimated, which may be induced to compensate the error of excluding the water mass from the precipitating and deep convective hydrometeors since the models are better
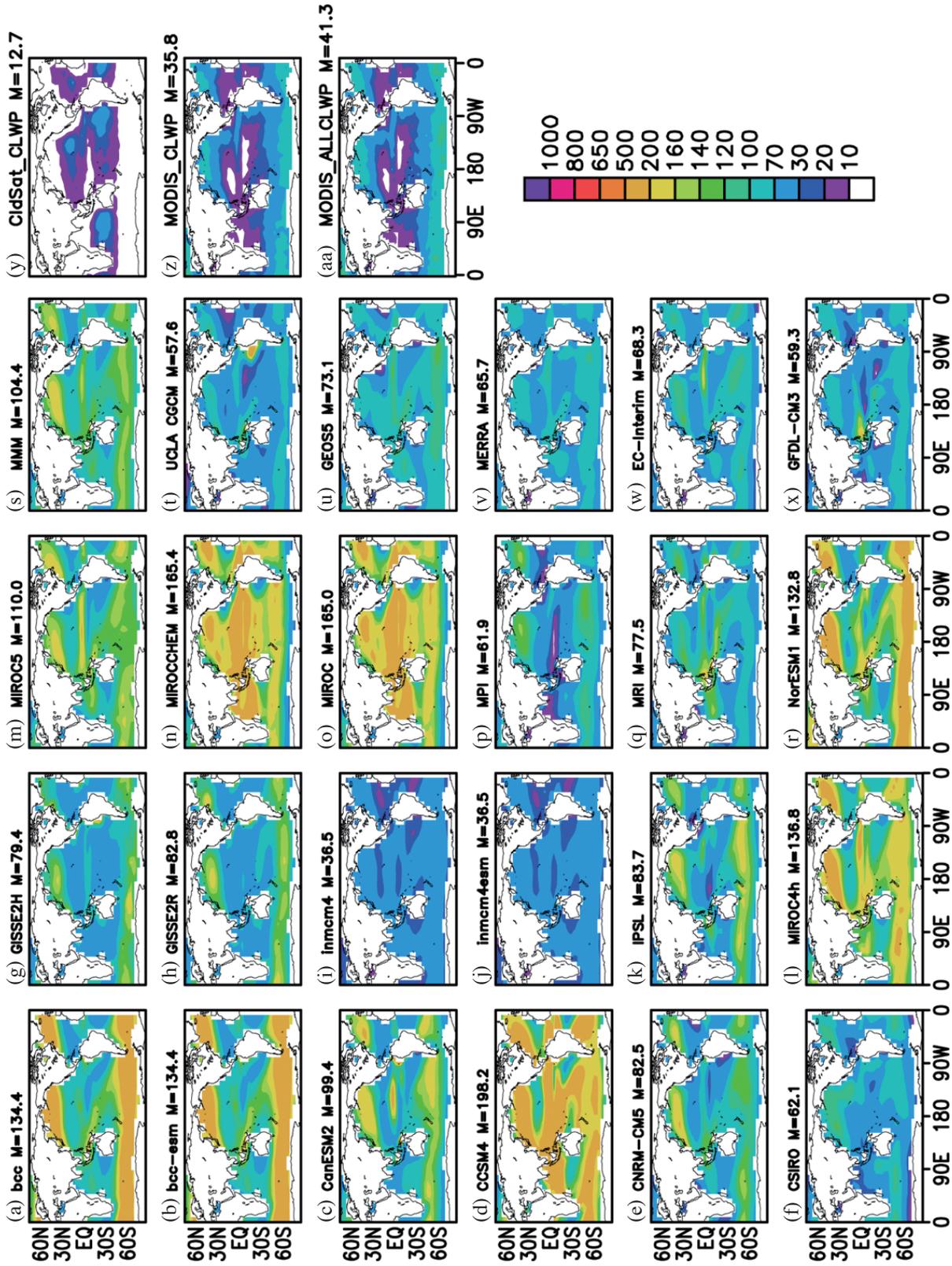
Fig. 4. (a) - (r) Annual mean maps of liquid water path (LWP; g m⁻²) from the 20th century GCM simulations contributed to the CMIP5, with their multi-model mean shown in (s). (t) UCLA GCM, (u) NASA GEOS5 AGCM, (v) MERRA reanalyses, and (w) ECMWF-Interim reanalyses, in (x) GFDL-CM3 TLWP as well as (y) observed estimate CloudSat CLWP under non-precipitating, non-convective, and non-mixed-phase conditions. (z) MODIS CLWP under non-precipitating non-convective conditions and (aa) MODIS CLWP under all conditions. Here the reference used is from (ab) for CLWP and (y) for lower limit and (aa) as upper limit for CLWP (see section 2).
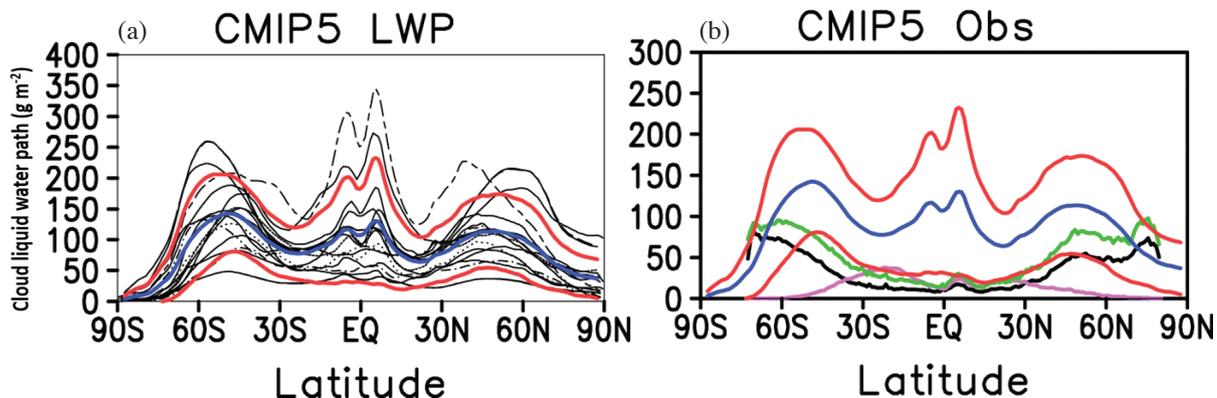
Fig. 5. (a) Zonal average of LWP quantities associated with each CMIP5 model (thin black solid, dotted, and dashed lines), multi-model mean (blue line), one standard deviation subtracted from and added to the mean for lower and upper bound (red line). (b) Zonal average of the three observationally estimated CLWP shown in Fig. 4: CloudSat CLWP (magenta), MODIS CLWP (black), and MODIS CLWP under all conditions (green). The CMIP5 multi-model mean and its upper and lower bounds are also plotted for direct comparison. The units: Cloud liquid water path (g m^{-2}).
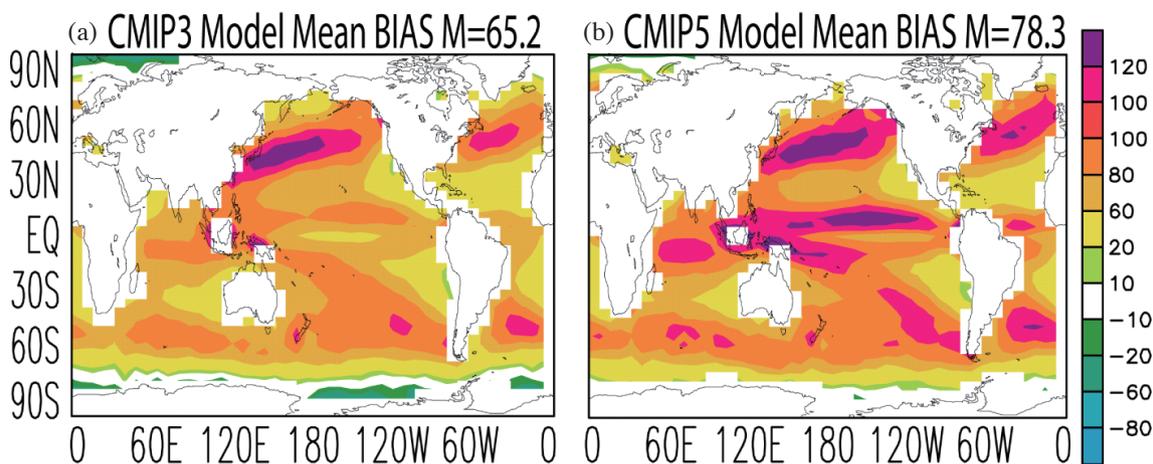
Fig. 6. (a) Multi-model mean CLWP bias (g m^{-2}) for the CMIP3 GCMs shown in Fig. A1; (b) multi-model mean CLWP bias associated with the eighteen CMIP5 GCMs shown in Figs. 4a - r with MODIS no rain and no convection cloud only LWP (Fig. 4z) used as a reference.

constrained to achieve the correct radiative fluxes.

To summarize the multi-model performance of CMIP5 in representing the time-mean pattern of CLWP, all the models are biased high (more than factor of 2 and even 3) in mid- and high-latitudes except for some CMIP5 models like CNRM, CSIRO, two GISS models, MPI and MRI as well as all the uncoupled models (UCLA and GEOS5) and reanalyses (MERRA and ECMWF-Interim) are comparable to the MODIS unfiltered CLWP. CMIP5 models show that mid- and high-latitude regions have stronger biases than tropical regions.

To further quantify and synthesize the comparative information discussed above, we use a Taylor diagram (Taylor 2001) as we did early for the IWC (Li et al. 2012). The Taylor diagram used in this study is a very commonly used statistical metric that relates two statistical measures of model fidelity: the spatial correlation and the spatial standard deviations (Taylor 2001). These statistics are calculated for the

long-term time mean and over the global ocean-only domain (area-weighted). The reference dataset is plotted along the x-axis at the value 1.0. The radial distance from the origin is proportional to the ratio of the standard deviations of the given dataset relative to the reference dataset. The azimuthal angle represents the spatial correlation between the given dataset and the reference dataset. The ratio of the standard deviation exhibits the relative amplitude of the simulated and the "reference" variations, whereas the correlation indicates the degree of similarity of variation between the two.

We use the MODIS filtered CLWP as the reference for the Taylor diagram analysis. The Taylor diagram shown in Fig. 7 summarizes both the degree of agreement in the overall spatial pattern correlations and the standard deviations for the individual CMIP5 CGCMs, their multi-model mean, two analyses, three other GCMs. MODIS unfiltered CLWP is used as another observational estimation to give a range of the observational uncertainties.
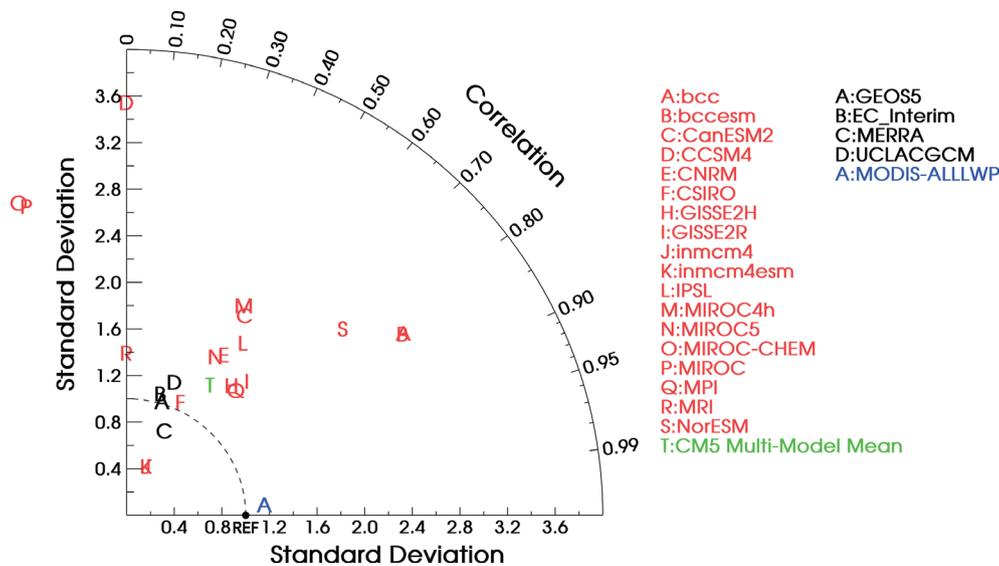
Fig. 7. Taylor diagram for GCM annual mean (80°N - 80°S) cloud liquid water path (CLWP; g m⁻²) where the reference value is the observed estimate MODIS CLWP shown in Fig. 4z (see section 2). The diagram includes the CMIP5 GCMs (red), their multi-model mean (green), GEOS5 GCM, UCLA GCM, MERRA, EC Interim analyses (black), and MODIS CLWP under all conditions (blue).

The two reanalyses (ECMWF and MERRA; see Table 1b) and AGCM simulations (i.e., specified SST; GEOS5, see Table 1b) perform as a group considerably better than the CMIP coupled GCMs in terms of the standard deviation ratio. The former have correlations between 0.2 and 0.4 and standard deviation ratios of between 0.8 and 1.1. The AGCM simulations are expected to perform better in various ways (such as spatial distribution of ITCZ, SPCZ, and SACZ) compared to CGCM simulations due to the use of prescribed SSTs. In a coupled GCM, the horizontal spatial distribution of surface ocean heat fluxes (radiation, latent and sensible as well as surface wind stress etc.) is important in driving model ocean dynamics (in particular in a long climate run) and therefore the distribution/values of the SSTs which determines the locations of the warm pool, ITCZ/SPCZ as well as the distribution/values of LWP/IWP and cloud fraction etc. In an uncoupled AGCM, on the other hand, prescribed SSTs are used, such that its surface heat fluxes, PBL, cloud-topped PBL and convection with cloud rooted from the PBL are strongly constrained by these prescribed SSTs, resulting in more realistic precipitation, deep convection and cloud and their associated LWC/LWP and IWC/IWP.

For the group of CMIP5 values (red), most of them have correlations between from about 0.3 to 0.8 with standard deviation ratios above 1 and up to 4 except for the two Inmcm4 models with values of 0.44. Two bcc models and the two MIROC models (MIROC and MIROC-CHEM) as well as CCSM4 have the standard deviation ratios that are highest among all with values of 2.7 - 3.6. Other poorly represented CLWP fields with extremely weak standard deviation by this metric are exhibited by the CMIP5 Inmcm4 and

Inmcm4ESM models. In terms of the distance to the reference point, the two best performing CMIP5 models by this metric are CSIRO (F) and MPI (Q) with the correlation and standard deviation ratios of about 0.4 - 0.6 and 1.0 and 1.4, respectively. Since the precipitating LWP is not considered here, so we do not include the GFDL model in this comparison. In principle, GFDL should be compared with observational TLWP with contributions from precipitating particles since GFDL LWC includes both precipitating and non-precipitating particles. Currently, the observational TLWP estimations are not available as discussed in section 3.

For the non-CMIP5 models (black in Fig. 7), the EC-Interim (B), MERRA (C), and GEOS 2.5 (A) perform well relative to the others in this group in terms of standard deviation ratio but it has the advantage in this case of being an AGCM-only run, and thus uses specified SSTs and with assimilation for MERRA and ERA-Interim, while all other models examined here are fully coupled. Noteworthy in this regard is the relatively good performance of the UCLA GCM (non-CMIP5) with values of 1.2, with metric values much better than most of the CMIP5 GCMs [except the models of MPI (Q) and CSIRO (F)]. In most cases, the correlation values are poor except NorESM1 (0.74), two BCC models (0.8). Two MIROC models even have negative correlations and CCSM4 and MRI have almost no correlation.

Next, we examine the fidelity of the model CLWC vertical structure using the synergistically constructed observational CLWC shown in Fig. 3. Such a comparison may provide clues as to where the model CLWP values are most awry. A comparison is given in Fig. 8, which shows the CLWC zonal and annual mean values from seventeen CMIP5 CGCMs (Figs. 8a - q; note that the CNRM-CM5 CGCM CLWC is
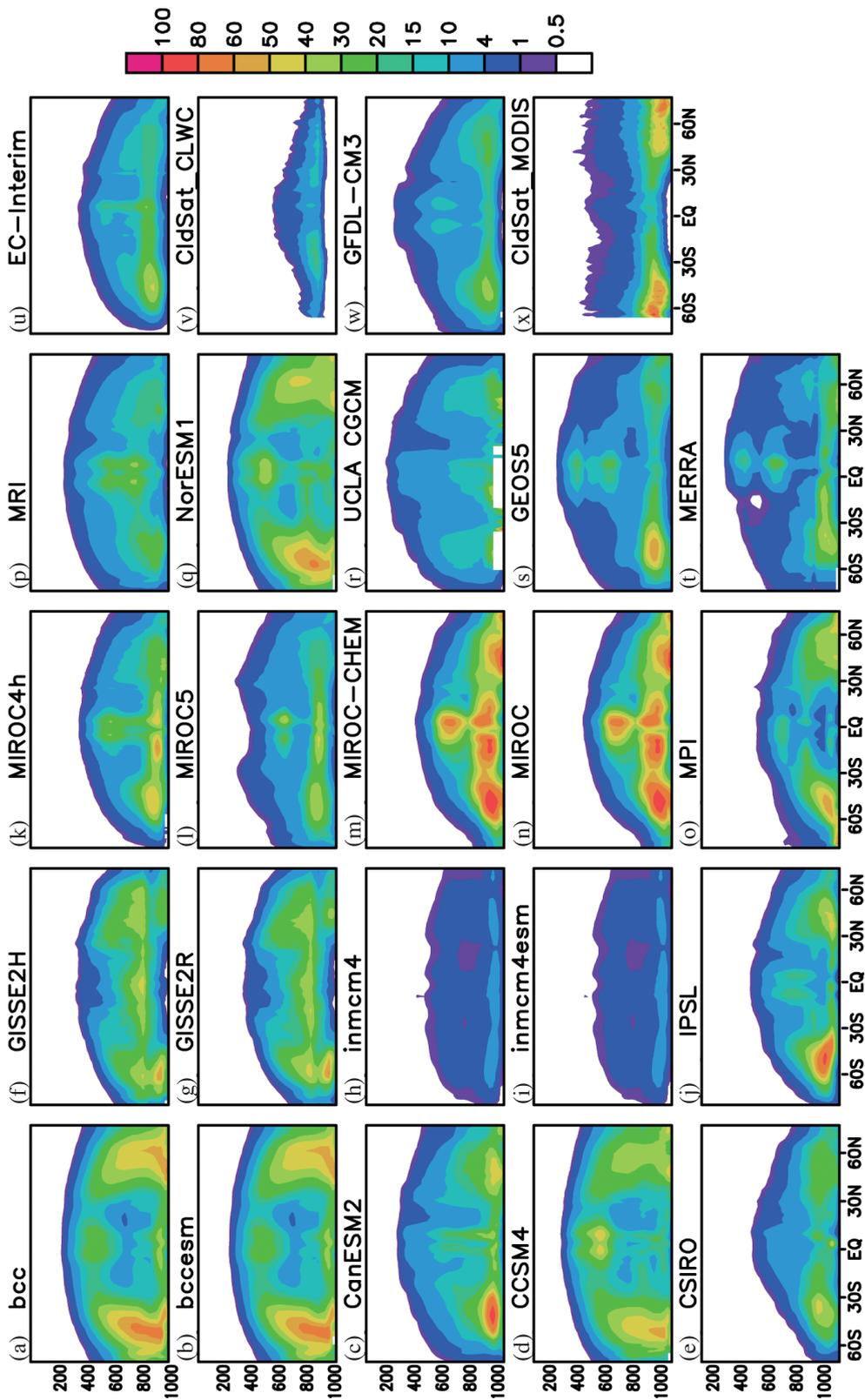
Fig. 8. Zonally-averaged annual mean CLWC (mg kg⁻¹) from the 20ᵗʰ century GCM simulations contributed to the CMIP5 (a) - (q), CLWC from (r) UCLA GCM, (s) NASA GEOS5 AGCM, (t) MERRA reanalyses, as well as (u) ECMWF-Interim reanalysis, and observed estimate CloudSat CLWC (v), (w) TLWC from GFDL CM3, and (x) CloudSat+MODIS CLWC, which is CloudSat CLWC combined with CLWC derived from MODIS CLWP assuming a moist adiabatic condition (see section 2). Note that the CNRM-CM5 CGCM CLWC is not available from the CMIP5 data portal at the time.

not available from the CMIP5 data portal at the time), UCLA GCM (Fig. 8r) and GEOS5 AGCM (Fig. 8s), as well as MER-RA (Fig. 8t), ERA-Interim (Fig. 8u). These models provide output specifically on CLWC only. The GFDL-CM3 shown in (Fig. 8w), on the other hand, provides output for TLWC including precipitating LWC. Overall, there are significant disparities above 800 hPa among the CMIP5 CGCMs against the observational CLWC with overall discrepancies ranging from multiplicative factors of about 0.2 of the observations (e.g., Inmcm4) to factors of 10 (e.g., MIROC-CHEM, MI-ROC GCMs).

Moreover, the general structure of their vertical distributions with respect to pressure levels is considerably different among the models. The large model spread might be due to the mixed phase and supercooled phase uncertainty. Each model has a different way to determine the thermodynamic phases of clouds and partition the cloud water mass into LWC and IWC. The different thermodynamic phase determination method can lead to a large variation in LWC in the low pressure levels in particular. About six of the CMIP5 models do a fair job in representing the vertical structure and magnitude of CLWC [i.e., CSIRO (Fig. 8e), IPSL (Fig. 8j), MIROCH4 (Fig. 8k), MIROC5 (Fig. 8l), MPI (Fig. 8o), MRI (Fig. 8p)]. Some models [bcc (Fig. 8a), bccesm (Fig. 8b), CanESM1 (Fig. 8c), CCSM4 (Fig. 8d), NorESM1 (Fig. 8q)] generally tend to qualitatively capture the patterns of CLWC over mid- and high-latitudes below 800 hPa. The GEOS5 model and two analyses from ECM-WF and MERRA analyses, on the other hand, tend to slightly overestimate CLWC in the tropics above 600 hPa. The UCLA GCM shows relatively underestimated CLWC values vertically compared to the observational CLWC values. However, it is reasonable to exercise caution when considering the fidelity of the observed values in these mid-lower tropospheric regions, or anywhere around the freezing level as the observational data from both CloudSat and MODIS retrievals are filtered for no mixed-phase clouds. Therefore, any liquid contributions from the mixed-phase clouds are not counted in the observational CLWC. Compared to the observed CLWC (Fig. 8x), the GFDL model captures the ITCZ in tropical regions, the extra-tropical storm track and polar regions pretty well but slightly overestimates CLWC above 600 hPa. To summarize, for many CMIP5 models, it is apparent from Figs. 6, 7, and 8 that significant disparities exist not only horizontally in CLWP but also in the vertical structure of CLWC.

Figure 9 summarizes some of the basic features of Fig. 8 by showing the global model mean vertical profiles (80°N - 80°S), against the global observed mean CloudSat CLWC (the thick blue line) and CloudSat+MODIS combined CLWC (the thick red line). Some models, in particular, the MIROC and MIROC-CHEM models, significantly overestimate CLWC with sharp increases below 500 hPa, while the two Inmcm4 models significantly underestimate

CLWC at all levels. Most of the models, in particular, the CCSM4, bcc, NorESM1, MIROCh4, and MIROC-ESM models, significantly overestimate CLWC with sharp increases above 600 hPa, The best simulated CLWC vertical profile is from CSIRO, followed by the MIROC5, MRI, and MPI, CanESM2 models (although keep in mind that CNRM-CM5 CGCM is not available for CLWC). Both MERRA and GEOS5 well capture CLWC in the lower troposphere.

By decomposing the global averages into various belt averages, we show in Figs. 9b, c, and d the average vertical profiles from tropical convectively active regions (30°N - 30°S) and mid- and high-latitudes of both hemispheres (30 - 60°N and 30 - 60°S). Over the tropical convectively active regions (Fig. 9b), the models except MIROC5 generally do not capture the correct CLWC peak, but CLWC values vary greatly from model to model and they all significantly overestimate CLWC at all levels except for GEOS5, two reanalyses and UCLA GCM. From Figs. 8c and d, we find that a large model bias exists in mid- and high-latitudes of both hemispheres and below 700 hPa, especially over the southern high-latitudes, similar to that in IWC reported in Li et al. (2012). Besides the two Inmcm4 models, which are biased low relative to the observed estimate throughout the troposphere, the other models are biased high at all levels in tropics. However, it is important to keep in mind that the observed estimate of CLWC excludes the contribution from mixed-phase clouds. The GFDL-CM3, which includes water mass from precipitating particles, does not have a counter part of observational TLWC, so no comparison can be made.

Finally, in order to determine if there are systematic biases across the models in the vertical structure of their CLWC fields, we examine the spatial correlation and standard deviation at each level, including 900, 850, 800, 700, 600, 500, and 400 hPa, for all the models and the multi-model mean against the observed CLWC values (i.e., Taylor Diagram analysis). A Taylor diagram representing each pressure layer of the annual mean CLWC for the CMIP5 multi-model mean (red), MERRA (blue), and ECMWF-Interim (green) are shown in Fig. 10. The CMIP5 multi-model mean shows poor correlations, even with some having zero and negative values, with observations at all levels considered, while it shows reasonable standard deviation ratios for 900, 850, and 800 hPa levels. The large ratios for 700, 600, and 500 hPa levels are partly due to the exclusion of the mixed-phase cloud mass in the observational estimate of CLWC since the observational partitioning of the mixed-phase cloud into liquid and ice is not reliable.

6. SUMMARY AND DISCUSSION

To assess the fidelity of GCMs in simulating cloud liquid water, liquid water path (LWP) retrieved from one passive satellite sensor, MODIS and the vertically-resolved liquid water content (LWC) estimates from the satellite radar,
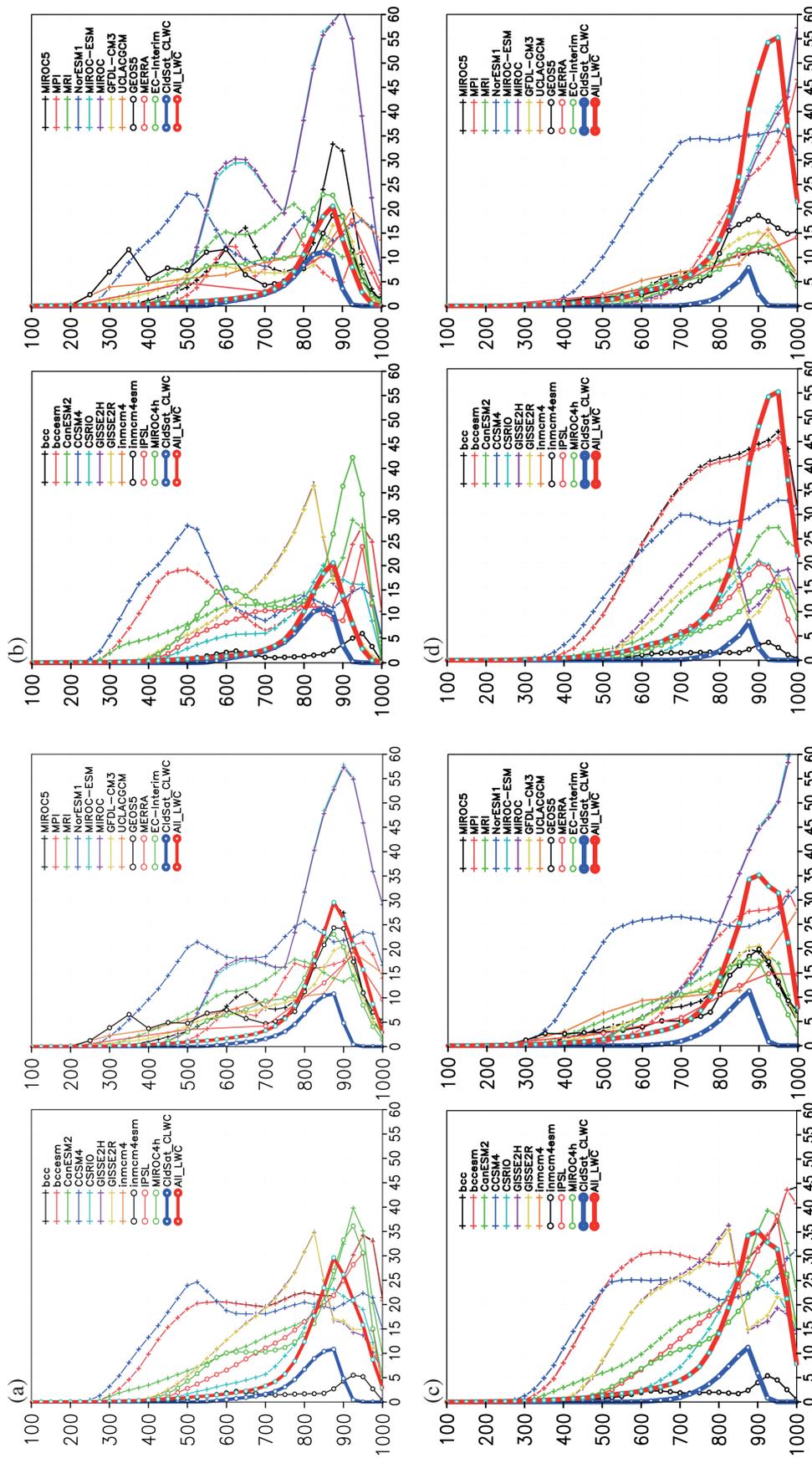
Fig. 9. (a) Globally-averaged (80°N - 80°S), annual mean, vertical CLWC profiles (mg kg⁻¹) from the CMIP5 GCMs, GEOS5 AGCM, EC Interim, and MERRA reanalyses, as well as the the CloudSat CLWC combined with MODIS-derived CLWC (All-LWC (CldSat_CLWC: thick blue profile) and CloudSat-only CLWC (CldSat_CLWC: thick red profile). Note that the CNRM-CM5 CGCM CIWC is not available from the CMIP5 data portal at the time. (b) The same as in Fig. 8a but for Tropical belt average (30°N - 30°S), annual mean, vertical CLWC profiles (mg kg⁻¹). (c) The same as in Fig. 8a but for Northern Mid-latitude belt average (30°N - 60°N), annual mean, vertical CLWC profiles (mg kg⁻¹). (d) The same as in Fig. 8a but for Southern Hemisphere Mid-latitude belt average (30°S - 60°S), annual mean, vertical CLWC profiles (mg kg⁻¹).

CMIP5 Multi-Model Mean CLWC vs Obs CLWC
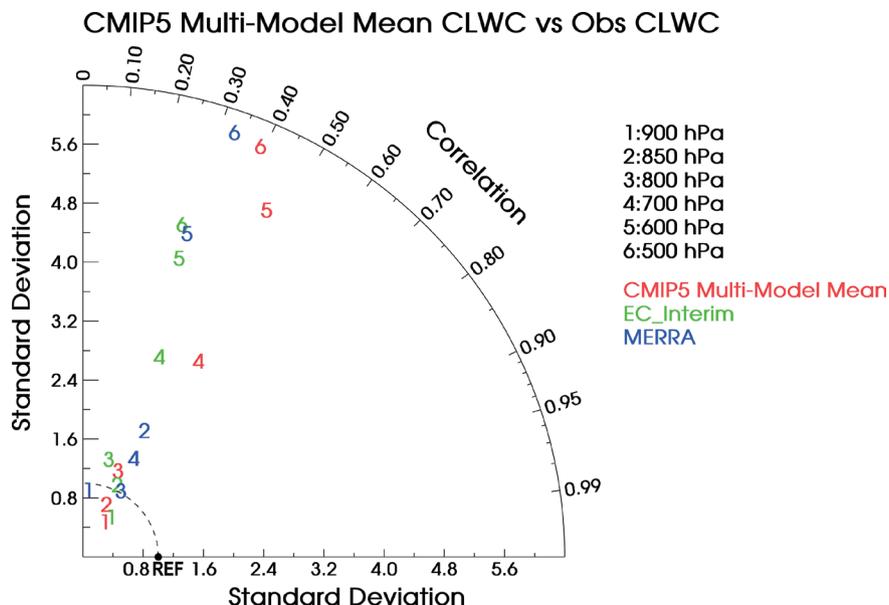
Fig. 10. Taylor diagram for GCM annual mean (80°N - 80°S) cloud liquid water content (CLWC; g kg⁻¹) where the reference value is the observed estimate CloudSat CLWC combined with MODIS-derived CLWC (All-LWC shown in Fig. 8a). The diagram includes the CMIP5 GCM multi-model mean (red), MERRA (blue), and EC Interim analyses (green) at 900, 850, 800, 700, 600, and 500 hPa.

CloudSat, are combined synergistically. We find that the patterns of CLWP bias in CMIP3 and CMIP5 are systematically similar with most models being biased high (Fig. 6). When MODIS filtered CLWP is used as a metric, there is overall a fairly wide disparity in the fidelity of CLWP representations in the CMIP5 models examined. Even for the annual mean maps considered, there are easily factors of 2, and nearly up to 10, for the differences between observations and modeled values for most of the GCMs over a number of regions (Figs. 4 - 7). There is only one model, CSIRO, among the CMIP5 ensemble examined here that performs rather well in regard to the Taylor diagram metrics (i.e., standard deviation ratio and pattern correlation) for CLWP. The following fair performers are MPI and MRI. The models that perform particularly poorly include MIROC, MIROC-CHEM, bcc, bcc-esm, INMCM4, and INMCM4-ESM, with the former (latter) four (two) being biased high (low) significantly in terms of overall CLWP magnitude. The remaining models exhibit intermediate performances.

As expected, the two reanalyses examined perform relatively well compared to the GCM group as a whole due to the incorporation of a wide array of constraining observations. This is still notable though since they do not assimilate cloud liquid observations and thus rely on (parameterized) model physics to represent this quantity. However, even with the assimilation of many other/related quantities, neither MERRA's nor ECMWF-Interim's performance was within the uncertainty of the observations for the pattern correlation. The UCLA GCM is one of the best performing CGCMs along with the three identified above (i.e., CSIRO, MPI, and MRI).

Considering the large disparities between the observa-

tions and modeled values of CLWP, it is evident that while the models may be providing roughly the correct radiative energy budget, many are accomplishing it by means of unrealistic cloud characteristics of cloud liquid mass, which in turn likely indicates unrealistic cloud particle sizes and cloud cover (e.g., Norris and Weaver 2001; Zhang et al. 2003; Lin and Zhang 2004; Schmidt et al. 2006; Cole et al. 2011; Franklin et al. 2012; Kay et al. 2012).

The CloudSat and MODIS combined CLWC (Fig. 3) is generated by complementing CloudSat CLWC by adding the MODIS derived CLWC based on the moist adiabatic assumption with the MODIS CLWP and CALIPSO determined cloud top height. The CloudSat and MODIS combined CLWC is used as a reference for model evaluations. Examination of the vertical structure of CLWC in terms of global, zonal and large-region averages (e.g., high, mid, and tropical latitudes) indicates similar findings in terms of overall performance across the models and reanalyses examined here. Most of the systematic errors in the global-mean vertical profile of CLWC occur below the mid-troposphere where the models tend to significantly overestimate CLWC compared to the observed estimate. No model (except CSIRO in tropics) generates reasonable spatial variability at all levels, in particular, above 800 hPa.

Given that there have been viable observed estimates of CLWP from MODIS and CLWP/CLWC for about 4 - 5 years from CloudSat yet still GCMs exhibit such large biases, the large disparity between the models and observations indicates challenges of utilizing the observations by all the modeling groups. These challenges likely include the uncertainty of CloudSat LWC below 900 hPa and the failure

of retrieval under precipitating/convective conditions that limits the development and improvement of the model PBL cloud and convection parameterizations. The complexity of the present study in using the observational data is an evidence of the difficulty of utilizing the observational data.

Beside the conventional CGCMs, no counterpart of TLWC/TLWP observations that include drizzle and rain is available for evaluating the performance of the GFDL-CM3. The lack of the TLWC/TLWP observational data is because that it is difficult to quantify the drizzle and rain from measurements of remote sensors.

In this study, we mainly focused on CLWC/CLWP comparisons between models/analyses and satellite retrievals. It is beyond the scope of this study to probe the causes of the observation-to-observation, model-to-model differences and model-to-observation biases. However, highlighting a few outstanding questions is instructive to help keep in mind the complexity associated with modeling atmospheric liquid water. This includes representing/parameterizing PBL clouds, shallow and deep convection, evaporation processes, autoconversion, cumulus detrainment, the overall interplay between these different physical parameterizations (and the large scale dynamics).

While the observations exhibit considerably better agreement in CLWP values than the CMIP5 models shown in the Taylor plot, there is a clear disagreement over the tropical western Pacific (i.e., warm pool), ITCZ/SPCZ and mid-latitude storm track. Several factors could contribute to the disagreement between these observational estimates. The deficiencies from passive detection include the presence of multi-level, mixed-phase, and thick clouds as well as surfaces that are bright and/or that have variable emissivity, each of which can represent a significant challenge for passive techniques, and are limited to estimates of CLWP with no/poor profiling capabilities. The main problem of CloudSat CLWC/CLWP retrievals is that they fail whenever there are drizzle and rain conditions and therefore the CLWC retrieval is difficult and often fails below PBL top and convective core under heavy rain. In addition, CloudSat CLWC/CLWP in PBL low clouds (normally lower than 900 - 1000 m) is often not retrieved because of the radar clutter issue near the surface. The PBL stratocumulus clouds, deep convective clouds, and shallow cumulus clouds are critical for the model cloud representations of the regions off the coasts of California and Peru in CGCMs but the CloudSat CLWC/CLWP has a limited use in evaluating the clouds in the regions.

In addition, at the time of this study we used existing flags for precipitation in the CloudSat 2C PRECIP-COLUMN product to distinguish precipitating (including drizzle) from non-precipitating clouds, which have valid flags over only oceans. Recent research has more systematically addressed the identification (Haynes et al. 2009) and quantification (Lebsock and L'Ecuyer 2011) of precipitation from CloudSat data leading to the development of experimental precipitation data products. As of this date, there remain inconsistencies between the various CloudSat products regarding the identification of precipitation, resulting in the existence of highly uncertain cloud water content retrievals in the presence of precipitation. An effort is under way to reconcile these inconsistencies for future product releases; however, determining quantitative information regarding cloud water content in the presence of precipitation will remain a challenge for some time to come.

Finally, the ability to explicitly represent both the cloudy, convective core and precipitating components of the liquid (and ice) mass has important physical considerations apart from just having additional observational constraints. While more work needs to be pursued in this area, there is a strong suggestion that GCMs should strive to explicitly represent a broader range of ice (Waliser et al. 2011; Li et al. 2012) and liquid hydrometeors, namely the larger falling hydrometeors (rain, snow) and include their effects in the radiative heating calculations which for the moment, although indirectly taken into account in many cases, is largely ignored from an explicit point of view (Li et al. 2013). Moreover, along with the evaluation results of this study, this consideration of the radiative impacts provide an additional evidence that the radiation balance in the CMIP class of GCMs which matches observations is still underconstrained and in many cases achieved in unrealistic ways. Taken together, these points indicate the need for (1) much improved retrieval algorithms that include observations from multiple platforms; (2) data-assimilation of space-based observations of cloud properties (something which is currently remarkably absent); and (3) additional observational resources to adequately characterize and constrain cloud-precipitation-radiation interactions. This is likely to include multi-channel radar/lidar information to characterize the profile and spectrum of cloud and precipitation particle sizes as well as Doppler radar capability to provide information on cloud and precipitation dynamics.

REFERENCES

Austin, R. T., A. J. Heymsfield, and G. L. Stephens, 2009: Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature. *J. Geophys. Res.*, **114**, D00A23, doi: 10.1029/2008JD010049. [Link]

Bodas-Salcedo, A., M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John, 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteorol. Soc.*, **92**, 1023-1043, doi: 10.1175/2011BAMS2856.1. [Link]

Bony, S. and K. A. Emanuel, 2001: A parameterization of the cloudiness associated with cumulus convection; Evaluation using TOGA COARE data. *J. Atmos. Sci.*, **58**, 3158-3183, doi: 10.1175/1520-0469(2001)058<3158:APOTCA>2.0.CO;2. [Link]

Borg, L. A. and R. Bennartz, 2007: Vertical structure of stratiform marine boundary layer clouds and its impact on cloud albedo. *Geophys. Res. Lett.*, **34**, L05807, doi: 10.1029/2006GL028713. [Link]

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S. K. Satheesh, S. Sherwood, B. Stevens, and X. Y. Zhang, 2013: Clouds and Aerosols. In: Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Eds.), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Bougeault, P., 1981: Modeling the trade-wind cumulus boundary layer. Part I: Testing the ensemble cloud relations against numerical data. *J. Atmos. Sci.*, **38**, 2414-2428, doi: 10.1175/1520-0469(1981)038<2414:MTTWCB>2.0.CO;2. [Link]

Bougeault, P., 1982: Cloud-ensemble relations based on the gamma probability distribution for the higher-order models of the planetary boundary layer. *J. Atmos. Sci.*, **39**, 2691-2700, doi: 10.1175/1520-0469(1982)039<2691:CERBOT>2.0.CO;2. [Link]

Boville, B. A., P. J. Rasch, J. J. Hack, and J. R. McCaa, 2006: Representation of clouds and precipitation processes in the Community Atmosphere Model Version 3 (CAM3). *J. Climate*, **19**, 2184-2198, doi: 10.1175/jcli3749.1. [Link]

Chaboureau, J.-P. and P. Bechtold, 2005: Statistical representation of clouds in a regional model and the impact on the diurnal cycle of convection during tropical convection, cirrus and nitrogen oxides (TROCCINOX). *J. Geophys. Res.*, **110**, D17103, doi: 10.1029/2004JD005645. [Link]

Cole, J., H. W. Barker, N. G. Loeb, and K. Von Salzen, 2011: Assessing Simulated Clouds and Radiative Fluxes Using Properties of Clouds Whose Tops are Exposed to Space. *J. Climate*, **24**, 2715-2727, doi: 10.1175/2011JCLI3652.1. [Link]

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. Mccaa, D. L. Williamson, J. T. Kiehl, B. Briegleb, C. Bitz, S.-J. Lin, M. Zhang, and Y. Dai, 2004: Description of the NCAR Community Atmosphere Model (CAM 3.0). NCAR Technical Note NCAR/TN-464+STR, National Center for Atmospheric Research, Boulder, Colorado, 210 pp, doi: 10.5065/D63N21CH. [Link]

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J.-N. Thépaut, and F. Vitart, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553-597, doi: 10.1002/qj.828. [Link]

Delanoë, J. and R. J. Hogan, 2010: Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *J. Geophys. Res.*, **115**, D00H29, doi: 10.1029/2009JD012346. [Link]

Delanoë, J., R. J. Hogan, R. M. Forbes, A. Bodas-Salcedo, and T. H. M. Stein, 2011: Evaluation of ice cloud representation in the ECMWF and UK Met Office models using CloudSat and CALIPSO data. *Q. J. R. Meteorol. Soc.*, **137**, 2064-2078, doi: 10.1002/qj.882. [Link]

Del Genio, A. D., M. Yao, W. Kovari, and K. K. Lo, 1996: A prognostic cloud water parameterization for global climate models. *J. Climate*, **9**, 270-304, doi: 10.1175/1520-0442(1996)009<0270:APCWPF>2.0.CO;2. [Link]

Donner, L. J., B. L. Wyman, R. S. Hemler, L. W. Horowitz, Y. Ming, M. Zhao, J.-C. Golaz, P. Ginoux, S.-J. Lin, M. D. Schwarzkopf, J. Austin, G. Alaka, W. F. Cooke, T. L. Delworth, S. M. Freidenreich, C. T. Gordon, S. M. Griffies, I. M. Held, W. J. Hurlin, S. A. Klein, T. R. Knutson, A. R. Langenhorst, H.-C. Lee, Y. Lin, B. I. Magi, S. L. Malyshev, P. C. D. Milly, V. Naik, M. J. Nath, R. Pincus, J. J. Ploshay, V. Ramaswamy, C. J. Seman, E. Shevliakova, J. J. Sirutis, W. F. Stern, R. J. Stouffer, R. J. Wilson, M. Winton, A. T. Wittenberg, and F. Zeng, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3. *J. Climate*, **24**, 3484-3519, doi: 10.1175/2011JCLI3955.1. [Link]

Franklin, C. N., C. Jakob, M. Dix, A. Protat, and G. Roff,

2012: Assessing the performance of a prognostic and a diagnostic cloud scheme using single column model simulations of TWP-ICE. *Q. J. R. Meteorol. Soc.*, **138**, 734-754, doi: 10.1002/qj.954. [Link]

Georgakakos, K. P. and R. L. Bras, 1984: A hydrologically useful station precipitation model: 1. Formulation. *Water Resour. Res.*, **20**, 1585-1596, doi: 10.1029/wr020i011p01585. [Link]

Guan, B., D. E. Waliser, J. F. Li, and A. da Silva, 2013: Evaluating the impact of orbital sampling on satellite-climate model comparisons. *J. Geophys. Res.*, **118**, 355-369, doi: 10.1029/2012jd018590. [Link]

Haynes, J. M., T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli, 2009: Rainfall retrieval over the ocean with spaceborne W-band radar. *J. Geophys. Res.*, **114**, D00A22, doi: 10.1029/2008JD009973. [Link]

Haynes, J. M., C. Jakob, W. B. Rossow, G. Tselioudis, and J. Brown, 2011: Major Characteristics of Southern Ocean Cloud Regimes and Their Effects on the Energy Budget. *J. Climate*, **24**, 5061-5080, doi: 10.1175/2011JCLI4052.1. [Link]

Hourdin, F., I. Musat, S. Bony, P. Braconnot, F. Codron, J. L. Dufresne, L. Fairhead, M. A. Filiberti, P. Friedlingstein, J. Y. Grandpeix, G. Krinner, P. LeVan, Z. X. Li, and F. Lott, 2006: The LMDZ4 general circulation model: Climate performance and sensitivity to parametrized physics with emphasis on tropical convection. *Climate Dyn.*, **27**, 787-813, doi: 10.1007/s00382-006-0158-0. [Link]

IPCC, 2007: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996 pp.

IPCC, 2013: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Jiang, J. H., H. Su, C. Zhai, V. S. Perun, A. Del Genio, L. S. Nazarenko, L. J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. A. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J.-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, T. S. LÉcuyer, E. M. Volodin, T. Iversen, H. Drange, M. D. S. Mesquita, W. G. Read, J. W. Waters, B. Tian, J. Teixeira, and G. L. Stephens, 2012: Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations. *J. Geophys. Res.*, **117**, D1410, doi: 10.1029/2011JD017237. [Link]

Kay, J. E., B. R. Hillman, S. A. Klein, Y. Zhang, B. Medeiros, R. Pincus, A. Gettelman, B. Eaton, J. Boyle, R. Marchand, and T. P. Ackerman, 2012: Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators. *J. Climate*, **25**, 5190-5207, doi: 10.1175/jcli-d-11-00469.1. [Link]

King, N. J., K. N. Bower, J. Crosier, and I. Crawford, 2013: Evaluating MODIS cloud retrievals with in situ observations from VOCALS-REx. *Atmos. Chem. Phys.*, **13**, 191-209, doi: 10.5194/acp-13-191-2013. [Link]

Klein, S. A. and C. Jakob, 1999: Validation and Sensitivities of Frontal Clouds Simulated by the Ecmwf Model. *Mon. Weather Rev.*, **127**, 2514-2531, doi: 10.1175/1520-0493(1999)127<2514:VASOFC>2.0.CO;2. [Link]

Köhler, M., 1999: Explicit prediction of ice clouds in general circulation models. Ph.D. Thesis, Department of Atmospheric Sciences, University of California Los Angeles, Los Angeles, CA, USA, 167 pp.

Kubar, T. L., D. E. Waliser, and J. Li, 2011: Boundary layer and cloud structure controls on tropical low cloud cover using A-Train satellite data and ECMWF analyses. *J. Climate*, **24**, 194-215, doi: 10.1175/2010jcli3702.1. [Link]

Lebsock, M. D. and T. S. L'Ecuyer, 2011: The retrieval of warm rain from CloudSat. *J. Geophys. Res.*, **116**, D20209, doi: 10.1029/2011JD016076. [Link]

Le Trent, H. and Z. X. Li, 1991: Sensitivity of an atmospheric general circulation model to prescribed SST changes: Feedback effects associated with the simulation of cloud optical properties. *Climate Dyn.*, **5**, 175-187, doi: 10.1007/bf00251808. [Link]

Li, J.-L., D. E. Waliser, J. H. Jiang, D. L. Wu, W. Read, J. W. Waters, A. M. Tompkins, L. J. Donner, J.-D. Chern, W.-K. Tao, R. Atlas, Y. Gu, K. N. Liou, A. Del Genio, M. Khairoutdinov, and A. Gettelman, 2005: Comparisons of EOS MLS cloud ice measurements with ECMWF analyses and GCM simulations: Initial results. *Geophys. Res. Lett.*, **32**, L18710, doi: 10.1029/2005GL023788. [Link]

Li, J.-L., J. H. Jiang, D. E. Waliser, and A. M. Tompkins, 2007: Assessing consistency between EOS MLS and ECMWF analyzed and forecast estimates of cloud ice. *Geophys. Res. Lett.*, **34**, L08701, doi: 10.1029/2006GL029022. [Link]

Li, J.-L., D. Waliser, C. Woods, J. Teixeira, J. Bacmeister, J. Chern, B.-W. Shen, A. Tompkins, W.-K. Tao, and M. Köhler, 2008: Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations. *Geophys. Res. Lett.*, **35**, L19710, doi: 10.1029/2008GL035427. [Link]

Li, J.-L., D. E. Waliser, and J. H. Jiang, 2011: Correction to "Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations". *Geophys. Res. Lett.*, **38**, L24807, doi: 10.1029/2011GL049956. [Link]

Li, J.-L., D. E. Waliser, W.-T. Chen, B. Guan, T. Kubar, G. Stephens, H.-Y. Ma, M. Deng, L. Donner, C. Seman, and L. Horowitz, 2012: An observationally based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data. *J. Geophys. Res.*, **117**, D16105, doi: 10.1029/2012JD017640. [Link]

Li, J.-L., D. E. Waliser, G. Stephens, S. Lee, T. L'Ecuyer, S. Kato, N. Loeb, and H.-Y. Ma, 2013: Characterizing and understanding radiation budget biases in CMIP3/CMIP5 GCMs, contemporary GCM, and reanalysis. *J. Geophys. Res.*, **118**, 8166-8184, doi: 10.1002/jgrd.50378. [Link]

Lin, B. and W. B. Rossow, 1994: Observations of cloud liquid water path over oceans: Optical and microwave remote sensing methods. *J. Geophys. Res.*, **99**, 20907-20927, doi: 10.1029/94jd01831. [Link]

Lin, B. and W. B. Rossow, 1997: Precipitation water path and rainfall rate estimates over oceans using special sensor microwave imager and International Satellite Cloud Climatology Project data. *J. Geophys. Res.*, **102**, 9359-9374, doi: 10.1029/96jd03987. [Link]

Lin, W. Y. and M. H. Zhang, 2004: Evaluation of clouds and their radiative effects simulated by the NCAR Community Atmospheric Model against satellite observations. *J. Climate*, **17**, 3302-3318, doi: 10.1175/1520-0442(2004)017<3302:EOCATR>2.0.CO;2. [Link]

Lohmann, U. and E. Roeckner, 1996: Design and performance of a new cloud microphysics scheme developed for the ECHAM general circulation model. *Climate Dyn.*, **12**, 557-572, doi: 10.1007/s003820050128. [Link]

Ma, H.-Y., M. Köhler, J.-L. Li, J. D. Farrara, C. R. Mechoso, R. M. Forbes, and D. E. Waliser, 2012: Evaluation of an ice cloud parameterization based on a dynamical-microphysical lifetime concept using CloudSat observations and the ERA-Interim reanalysis. *J. Geophys. Res.*, **117**, D05210, doi: 10.1029/2011JD016275. [Link]

Ma, H.-Y., H. Xiao, C. R. Mechoso, and Y. Xue, 2013: Sensitivity of global tropical climate to land surface processes: Mean state and interannual variability. *J. Climate*, **26**, 1818-1837, doi: 10.1175/jcli-d-12-00142.1. [Link]

Mace, G. G. and F. J. Wrenn, 2013: Evaluation of the hydrometeor layers in the East and West Pacific within ISCCP cloud-top pressure-optical depth bins using merged *CloudSat* and *CALIPSO* data. *J. Climate*, **26**, 9429-9444, doi: 10.1175/jcli-d-12-00207.1. [Link]

Marchand, R., G. G. Mace, T. Ackerman, and G. Stephens, 2008: Hydrometeor detection using CloudSat—An earth-orbiting 94-GHz cloud radar. *J. Atmos. Ocean. Technol.*, **25**, 519-533, doi: 10.1175/2007jtecha1006.1. [Link]

Norris, J. R. and C. P. Weaver, 2001: Improved techniques for evaluating GCM cloudiness applied to the NCAR CCM3. *J. Climate*, **14**, 2540-2550, doi: 10.1175/1520-0442(2001)014<2540:ITFEGC>2.0.CO;2. [Link]

Ogura, T., S. Emori, M. J. Webb, Y. Tsushima, T. Yokohata, A. A. Ouchi, and M. Kimoto, 2008: Towards understanding cloud response in atmospheric GCMs: The use of tendency diagnostics. *J. Meteorol. Soc. Jpn.*, **86**, 69-79, doi: 10.2151/jmsj.86.69. [Link]

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Trans. Geosci. Remote Sensing*, **41**, 459-473, doi: 10.1109/TGRS.2002.808301. [Link]

Randall, D. A. and S. Tjemkes, 1991: Clouds, the earth's radiation budget, and the hydrologic cycle. *Global Planet. Change*, **4**, 3-9, doi: 10.1016/0921-8181(91)90063-3. [Link]

Rasch, P. J. and J. E. Kristjánsson, 1998: A comparison of the CCM3 model climate using diagnosed and predicted condensate parameterizations. *J. Climate*, **11**, 1587-1614, doi: 10.1175/1520-0442(1998)011<1587:ACOTCM>2.0.CO;2. [Link]

Ricard, J. L. and J. F. Royer, 1993: A statistical cloud scheme for use in an AGCM. *Ann. Geophys.*, **11**, 1095-1115.

Rienecker, M. M., M. J. Suarez, R. Todling, J. Bacmeister, L. Takacs, H.-C. Liu, W. Gu, M. Sienkiewicz, R. D. Koster, R. Gelaro, I. Stajner, and J. E. Nielsen, 2008: The GEOS-5 Data Assimilation System—Documentation of Versions 5.0.1, 5.1.0, and 5.2.0. In: Suarez, M. J. (Ed.), Technical Report Series on Global Modeling and Data Assimilation, Volume 27, NASA/TM-2008-104606, NASA Goddard Space Flight Center; Greenbelt, MD, United States.

Rotstayn, L. D., 1997: A physically based scheme for the treatment of stratiform clouds and precipitation in large-scale models. I: Description and evaluation of the microphysical processes. *Q. J. R. Meteorol. Soc.*, **123**, 1227-1282, doi: 10.1002/qj.49712354106. [Link]

Rotstayn, L. D., B. F. Ryan, and J. J. Katzfey, 2000: A scheme for calculation of the liquid fraction in mixed-phase stratiform clouds in large-scale models. *Mon. Weather Rev.*, **128**, 1070-1088, doi: 10.1175/1520-0493(2000)128<1070:ASFCOT>2.0.CO;2. [Link]

Schmidt, G. A., R. Ruedy, J. E. Hansen, I. Aleinov, N. Bell, M. Bauer, S. Bauer, B. Cairns, V. Canuto, Y. Cheng, A. Del Genio, G. Faluvegi, A. D. Friend, T. M. Hall, Y. Hu, M. Kelley, N. Y. Kiang, D. Koch, A. A. Lacis, J. Lerner, K. K. Lo, R. L. Miller, L. Nazarenko, V. Oinas, J. Perlwitz, J. Perlwitz, D. Rind, A. Romanou, G. L. Russell, M. Sato, D. T. Shindell, P. H. Stone, S. Sun, N. Tausnev, D. Thresher, and M.-S. Yao, 2006: Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *J. Climate*, **19**, 153-192, doi: 10.1175/jcli3612.1. [Link]

Stephens, G. L. and C. D. Kummerow, 2007: The remote sensing of clouds and precipitation from

space: A review. *J. Atmos. Sci.*, **64**, 3742-3765, doi: 10.1175/2006jas2375.1. [Link]

Stephens, G. L., D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, and the CloudSat Science Team, 2002: The CloudSat mission and the A-TRAIN: A new dimension of space-based observations of clouds and precipitation. *Bull. Amer. Meteorol. Soc.*, **83**, 1771-1790, doi: 10.1175/bams-83-12-1771. [Link]

Stephens, G. L., D. G. Vane, S. Tanelli, E. Im, S. Durden, M. Rokey, D. Reinke, P. Partain, G. G. Mace, R. Austin, T. L'Ecuyer, J. Haynes, M. Lebsock, K. Suzuki, D. Waliser, D. Wu, J. Kay, A. Gettelman, Z. Wang, and R. Marchand, 2008: CloudSat mission: Performance and early science after the first year of operation. *J. Geophys. Res.*, **113**, D00A18, doi: 10.1029/2008JD009982. [Link]

Sundqvist, H., 1988: Parameterization of condensation and associated clouds in models for weather prediction and general circulation simulation. In: Schlesinger, M. E. (Ed.), Physically-Based Modelling and Simulation of Climate and Climatic Change, NATO ASI Series (Series C: Mathematical and Physical Sciences), Vol. 243, Springer, Dordrecht, 433-461, doi: 10.1007/978-94-009-3041-4_10. [Link]

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.*, **106**, 7183-7192, doi: 10.1029/2000jd900719. [Link]

Tiedtke, M., 1993: Representation of clouds in large-scale models. *Mon. Weather Rev.*, **121**, 3040-3061, doi: 10.1175/1520-0493(1993)121<3040:ROCILS>2.0.CO;2. [Link]

Volodin, E. M., N. A. Dianskii, and A. V. Gusev, 2010: Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations. *Izv. Atmos. Ocean. Phys.*, **46**, 414-431, doi: 10.1134/S000143381004002X. [Link]

Waliser, D. E., J.-L. F. Li, C. P. Woods, R. T. Austin, J. Bacmeister, J. Chern, A. Del Genio, J. H. Jiang, Z. Kuang, H. Meng, P. Minnis, S. Platnick, W. B. Rossow, G. L. Stephens, S. Sun-Mack, W.-K. Tao, A. M. Tompkins, D. G. Vane, C. Walker, and D. Wu, 2009: Cloud ice: A climate model challenge with signs and expectations of progress. *J. Geophys. Res.*, **114**, D00A21, doi: 10.1029/2008JD010015. [Link]

Waliser, D. E., J.-L. F. Li, T. S. L'Ecuyer, and W.-T. Chen, 2011: The impact of precipitating ice and snow on the radiation balance in global climate models. *Geophys. Res. Lett.*, **38**, L06802, doi: 10.1029/2010GL046478. [Link]

Watanabe, M., S. Emori, M. Satoh, and H. Miura, 2009: A PDF-based hybrid prognostic cloud scheme for general circulation models. *Climate Dyn.*, **33**, 795-816, doi: 10.1007/s00382-008-0489-0. [Link]

Webb, M., C. Senior, S. Bony, and J. J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dyn.*, **17**, 905-922, doi: 10.1007/s003820100157. [Link]

Wilson, D. R. and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK Meteorological Office unified model. *Q. J. R. Meteorol. Soc.*, **125**, 1607-1636, doi: 10.1002/qj.49712555707. [Link]

Yukimoto, S., H. Yoshimura, M. Hosaka, T. Sakami, H. Tsujino, M. Hirabara, T. Y. Tanaka, M. Deushi, A. Obata, H. Nakano, Y. Adachi, E. Shindo, S. Yabu, T. Ose, and A. Kitoh, 2011: Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1) -Model Description-. Technical Reports of the Meteorological Research Institute No.64, Meteorological Research Institute, Japan, 96 pp, doi: 10.11483/mritechrepo.64. [Link]

Zhang, M., W. Lin, C. S. Bretherton, J. J. Hack, and P. J. Rasch, 2003: A modified formulation of fractional stratiform condensation rate in the NCAR Community Atmospheric Model (CAM2). *J. Geophys. Res.*, **108**, 4035, doi: 10.1029/2002JD002523. [Link]

APPENDIX A

(1) The Uncertainties and Limitations of MODIS Cloud LWP

Platnick et al. (2003) summarize most of (although not all) errors associated with the retrieval of cloud optical properties and thus by inference liquid water path inferred in this way. Their analysis indicates that the characteristic optical depth error of low clouds is about 10% over most of the range of optical depths with much larger errors occurring for smaller optical depths especially for clouds over land due to the creeping influence of the surface albedo uncertainties in these cases. The errors on effective radius for low clouds are similar in the range from 10 - 20% but are substantially more uncertain for ice clouds. Platnick et al. (2003) omit contributions from model errors that are discussed in details in Stephens and Kummerow (2007) and for certain retrieval problems [such as broken clouds with three-dimensional (3D) effects] these omitted errors dominate the error budget.

There are also other limitations of these data that are not considered in Platnick et al. (2003). The MODIS optical measurement loses sensitivity for large LWPs since the reflected sunlight saturates to almost constant values at large optical depths. The optical methods cannot retrieve meaningful water paths in mixed-phased clouds or clouds masked by high level ice clouds. Thus, optical methods cannot sample the full distribution of liquid water paths and are limited primarily to warm single layered cloud systems. The method also is subject to uncertainties introduced by the existence of undetected high clouds that occur overlying low clouds. These events are not negligiblele and appear as problematic

biases in current radiance based climatology, (e.g., Haynes et al. 2011; Mace and Wrenn 2013). The third limitation of these optical measurements is that they tend to be weighted high in the cloud and contain restricted information about the liquid water path associated with precipitation.

These limitations are confirmed by a comparison study with in situ observations from VOCALS-Rex (King et al. 2013). The MODIS retrieved effective radius of the droplet size distribution overestimates the in situ measurements on average by 13% with largest overestimation coinciding with the presence of the drizzle sized droplets. The optical depth retrieved from MODIS also overestimate in situ values. These two high biases lead to an overestimation of liquid water path in MODIS. Although these limitations have to be kept in mind in considering the results to follow, we conclude that under the most ideal conditions (no precipitation, no high cloud, fully overcast without any significant 3D effects) the MODIS LWP uncertainty is expected to be in the range of 10 - 30% (King et al. 2013).

(2) The Collocation of CloudSat-MODIS-AMSR-E

The MODIS-collocated-to-CloudSat and AMSR-E-collocated-to-CloudSat datasets are a part of the CloudSat-centric Co-location Data Products for A-Train data and ECMWF analysis outputs available at http://csyotc.cira. colostate.edu/index.php. The data products are generated for the Year of Tropical Convection (YOTC) study (http:// www.ucar.edu/yotc). Specifically, the data products contain convection and cloud related A-Train satellite retrievals that are collocated to CloudSat footprints for the period May 2008 to April 2010. Products include quantities from CALIPSO, MLS, AIRS, AMSR-E, CERES, MODIS, and a number of fields from the specialized YOTCO analysis produced by ECMWF. The co-location data products are produced to facilitate the synergistic use of the multi-source datasets by providing them in common geo-location parameters and a common data format.

The co-location process is done by finding the nearest neighbor footprint in the source data (i.e., A-Train or ECMWF data) for a given footprint in the target data (i.e., CloudSat data). For A-Train co-location, if the distance between the nearest-neighbor footprint and the target footprint is smaller than 1.5x (the sum of the two footprint sizes) and the time difference between the two data is smaller than about 15 minutes (which is the largest time span of the A-Train instruments), we accept the nearest-neighbor footprint and co-locate it onto the target footprint. Neither spatial averaging nor interpolation is applied in the co-location process. The value from the accepted nearest-neighbor footprint is simply copied over to the target footprint. We take this simple nearest-neighbor copying approach because an appropriate averaging and interpolation varies from instrument to instrument and from variable to variable and re-

quires detailed and individual analyses.

The purpose of the collocation was performed in order to apply the CloudSat-based precipitation and cloud condition flags to the MODIS LWP data to partition these data into the suspended (cloud-only), convective and precipitating portions for reasons discussed below. This collocation process reduces the size of temporal and spatial resolutions of the MODIS data significantly. The collocated and partitioned MODIS LWP retrievals are mapped onto 1 × 1 degree grid and averaged for all sky conditions. Note that MODIS MYD06_L2 product includes only LWP values from cloudy conditions and treats the LWP retrievals from a clear sky condition as invalid. In order to make the all sky averaged LWP, we set the LWP retrieval from the clear sky condition to be zero and add it to the averaging. This process is effectively the same as averaging the cloudy-condition LWP values only and multiplying it by the cloud fraction in each grid.

(3) Cloud Detection Capabilities of CloudSat and MODIS

Apart from the retrieval uncertainty, the cloud detection capabilities of CloudSat and MODIS are an important factor in assessing the observational uncertainty. When the CALIPSO lidar measurements are combined with the CloudSat radar measurements, they provide the most accurate measurement of cloud detection and height structure available among all satellite-based measurements today because their complementary features. The radar can probe optically thick layers better than the lidar while the lidar can sense optically thin layers and tenuous cloud tops better than the radar. CloudSat 2B-Radar-Lidar GEOPROF product is a retrieval product, which uses the synergy of the radar and lidar.

Given that the radar and lidar combined detection is superior to other measurements, however, the radar-only measurement misses most low clouds, which is attributed to the radar surface cluttering issue. This is consistent with previous observational studies such as Marchand et al. (2008) and Kubar et al. (2011). While MODIS misses some clouds compared to the lidar and radar combined detection level, MODIS detection capability is overall within 10% error. These detection deficiencies will unavoidably lead to systematic biases in the annual means of CLWC and CLWP retrieved from the CloudSat and MODIS measurements. Quantifying the systematic biases is difficult because there is little information about the cloud water mass (CLWC/CLWP) of the clouds that are missed by MODIS. However, the clouds missed by MODIS are expected to be mainly optically thin clouds and thus have low liquid cloud water mass. Therefore, the effect of missed clouds on the annual means of CLWC/CLWP should be much less than 10%, which is the MODIS cloud detection missing frequency as mentioned above.

The majority of the CMIP5 model outputs of cloud

liquid water path (CLWP) is available from 1850 - 2005. However, the CloudSat is only available starting from June 2006. Due to the battery failure of CloudSat cloud profiling radar, the data after 2011 is daytime only. In this study, we use data from January 2007 to December 2010 to include daytime and night time. The MODIS LWP we use (May 2008 to April 2010) has more profiles collocated path along with CloudSat track.

Since the data period is too short to compute the trend, we compute the annual mean LWP of MODIS data from January 2001 to December 2005 (Fig. A1a) which is closer to the CMIP5 data period, against the annual mean LWP from May 2008 to April 2010 (Fig. A1b), to examine their difference (Fig. A2c) and their changes relative to the annual mean in 2001 - 2005 (Fig. A1d). The changes of LWP is within 10% (Fig. A2d), and their difference (Fig. A2c) is about 4 - 6 (g m^{-2}), which is well below the biases in CMIP3/CMIP5 (values of 20 up to 100 g m^{-2}) shown in Fig. 6. In addition, we compute the annual mean LWP of CMIP5 data from January 1970 to December 2005 against the annual mean LWP from January 2001 to December 2005, and

their changes relative to the annual mean of 1970 - 2005 (figure not shown) is within 2 - 4%, which are also well below the biases in CMIP3/CMIP5.

(4) Determination of Lfting Condensation Level (LCL)

We use the lifting condensation level (LCL) defined as the cloud base height. From Georgakakos and Bras (1984), the LCL is estimated as

$$LCL(p) = \frac{P_{SFC}}{[(T - T_D)/223.15 + 1]^{3.5}}$$
$$T_D = T \left[1 - \frac{T \ln(RH_{SFC}/100)}{L/R_V} \right]^{-1} \quad (A1)$$

where P_{SFC} is the surface pressure, T_D is the dewpoint temperature in Kelvin, RH_{SFC} is the surface relative humidity, and R_V is the gas constant for vapor (= 461 J K^{-1} kg^{-1}). The LCL pressure is converted to height using the CloudSat ECMWF-AUX data.
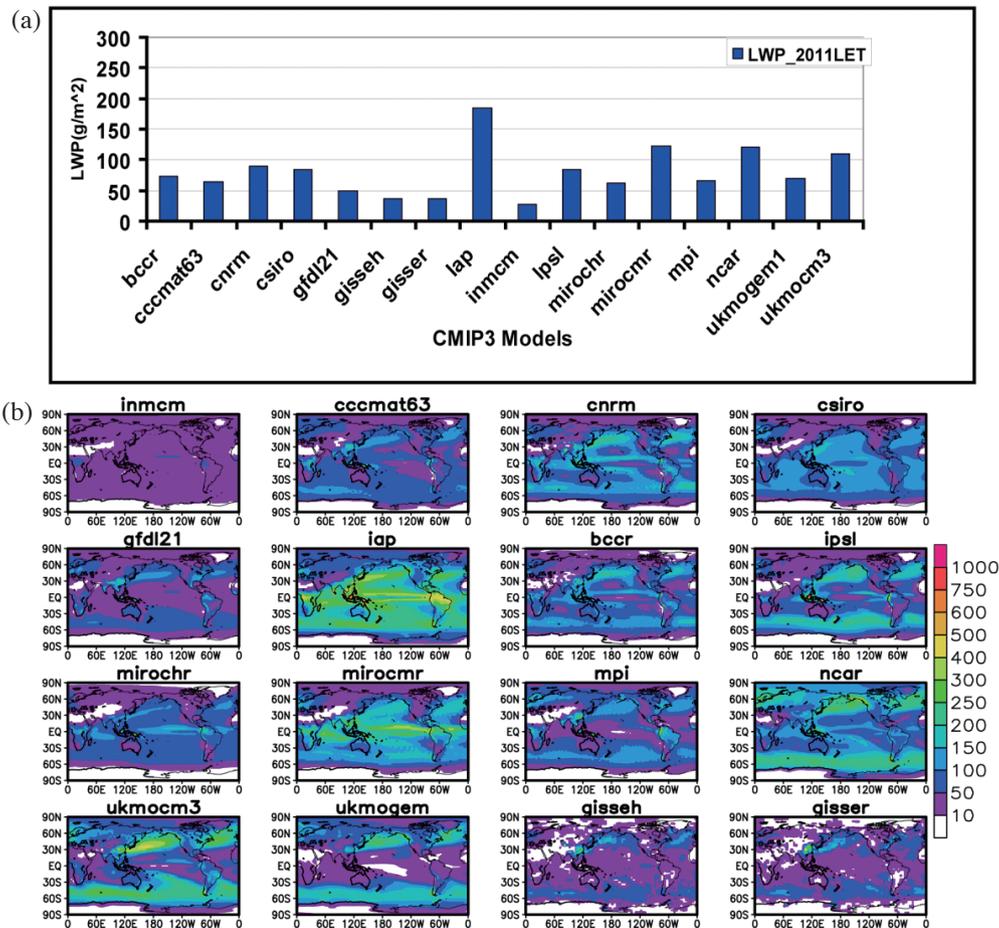
Fig. A1. (a) Globally-averaged, annual means of CLWP from the 1970 - 1994 periods of the 20th century GCM simulations contributed to the IPCC AR4 FAR (20c3m scenario). (b) Annual mean values of CLWP (g m^{-2}) from 1970 - 1994 of the 20th century GCM simulations contributed to the IPCC AR4 FAR (20c3m scenario). The figure is taken from Li et al. (2011).
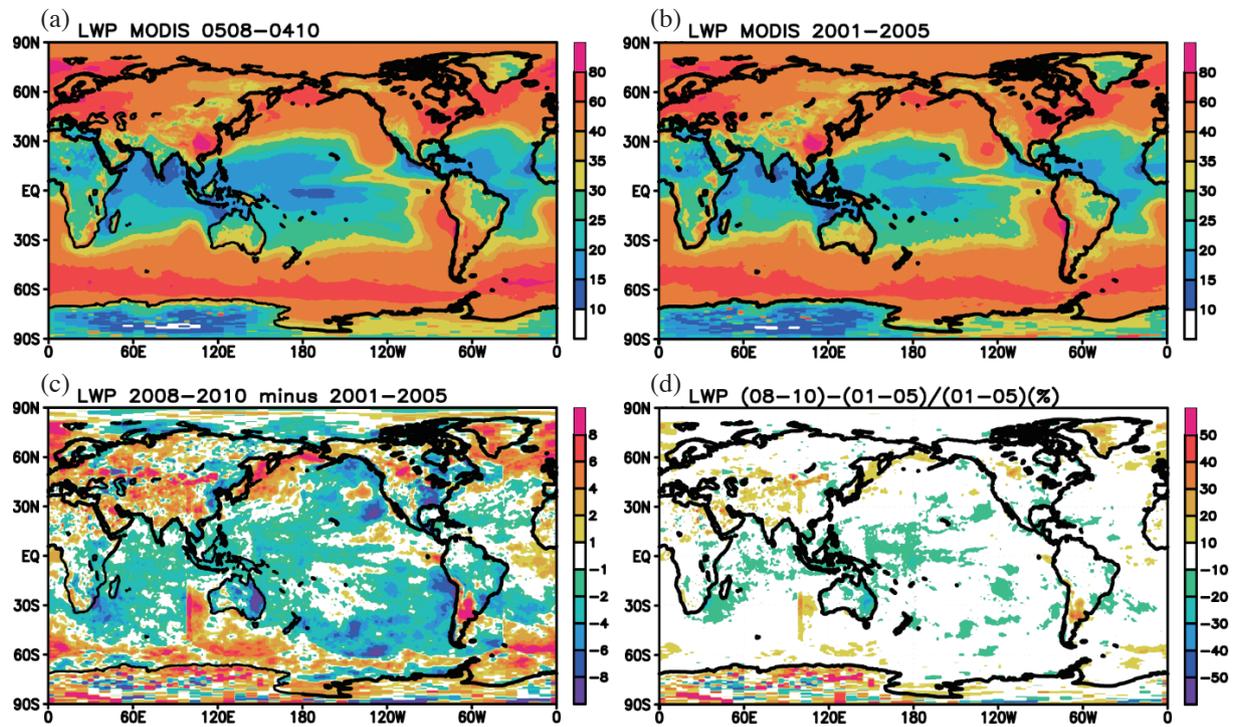
Fig. A2. (a) Annual mean cloud liquid water path (CLWP: g m^{-2}) from MODIS average from May 2008 to April 2010; (b) same as (a) but from January 2001 to December 2005; (c) the difference between (a) and (b); (d) the difference of (c) relative to (a) (unit: %).