

Polar clouds and radiation in satellite observations, reanalyses, and climate models

Jan T. M. Lenaerts^{1,2,3}, Kristof van Tricht², Stef Lhermitte^{2,4}, Tristan

L'Equyer⁵

Corresponding author: J. T. M. Lenaerts, Institute for Marine and Atmospheric Research Utrecht (IMAU), University of Utrecht, Princetonplein 5, Utrecht, 3584 CC, The Netherlands.
(Jan.Lenaerts@colorado.edu)

¹Institute for Marine and Atmospheric

Research Utrecht, Utrecht University,
Utrecht, The Netherlands

²Department of Earth- and

Environmental Sciences, KU Leuven,
Leuven, Belgium

³Department of Atmospheric and Oceanic

Sciences, University of Colorado Boulder,
Boulder CO, USA

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: [10.1002/2016GL072242](https://doi.org/10.1002/2016GL072242)

Clouds play a pivotal role in the surface energy budget of the Polar regions.

Here we use two largely independent datasets of cloud and surface downwelling radiation observations derived by satellite remote sensing (2007-2010) to evaluate simulated clouds and radiation over both Polar ice sheets and oceans in state-of-the-art atmospheric reanalyses (ERA-Interim and MERRA-2) and the CMIP5 climate model ensemble. Firstly, we show that, compared to CERES-EBAF, CloudSat-CALIPSO better represents cloud liquid and ice water path over high latitudes, owing to its recent explicit determination of cloud phase that will be part of its new R05 release. The reanalyses and climate models disagree widely on the amount of cloud liquid and ice in the Polar regions.

Compared to the observations, we find significant but inconsistent biases in the model simulations of cloud liquid and ice water, as well as in the downwelling radiation components. The CMIP5 models display a wide range of cloud characteristics of the Polar regions, especially with regard to cloud liquid water, limiting the representativeness of the multi-model mean. A few

⁴Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, The Netherlands

⁵Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison WI, USA

CMIP5 models (CNRM, GISS, GFDL, and IPSL_CM5b) clearly outperform the others, which enhances credibility in their projected future cloud and radiation changes over high latitudes. Given the rapid changes in Polar regions and global feedbacks involved, future climate model developments should target improved representation of Polar clouds. To that end, remote sensing observations are crucial, in spite of large remaining observational uncertainties, which is evidenced by the substantial differences between the two datasets.

1. Introduction

The Polar regions (poleward of 60° North and South) are fast-changing environments under ongoing climate warming. The Arctic sea ice is melting rapidly, with the September monthly average sea ice cover for the entire Arctic Ocean declining -13.4% per decade relative to the 1981-2010 average [*Blunden and Arndt, 2016*], and progressively smaller ratios of thick, multi-year sea ice relative to <1 year old sea ice [*Maslanik et al., 2007*]. In the meantime, the Greenland Ice Sheet is experiencing dramatic mass loss: it has lost nearly one trillion (10^{15}) kg of ice into the Arctic Ocean over the period 2011-2014 [*McMillan et al., 2016*], to a great extent driven by atmospheric warming [*Van den Broeke et al., 2016*]. In West Antarctica, ice shelves have partly become unstable after several warm summers, sustained surface melt and meltwater-induced hydrofracturing [*Scambos et al., 2000; Van den Broeke, 2005*], while enhanced ice shelf bottom melting and calving rates in the last two decades [*Wouters et al., 2015; Mouginot et al., 2014*] are associated with warming ocean surface waters [*Cook et al., 2016*]. Over the Southern Ocean, and around East Antarctica, ocean waters have not warmed at all yet [*Armour et al., 2016*], and in stark contrast to the Arctic, Antarctic sea ice has slightly expanded since 2000, attributed to natural, decadal variability in the tropical Pacific [*Meehl et al., 2016*].

While the atmosphere has long been recognised to drive some of these rapid changes, the role of clouds has recently come to the fore-front [*Kay et al., 2016a*], for instance in inhibiting meltwater refreezing on Greenland [*Van Tricht et al., 2016a*], enhancing sea ice melting [*Kay et al., 2008*], and controlling Antarctic ice shelf melting [*King et al., 2015*]. The net amount of energy received at the surface (surface energy balance, SEB)

is greatly affected by the presence of clouds. They induce competing effects of decreasing the amount of solar (shortwave) energy and enhancing thermal (longwave) radiation, the ratio of which depends on cloud structure, height and frequency.

Whereas such cloud characteristics are challenging to measure from the ground, especially in the Polar regions, satellite remote sensing has opened up venues to study these cloud characteristics on large scales [Bromwich *et al.*, 2012; Grenier *et al.*, 2009; Cesana *et al.*, 2012; Kay and Gettelman, 2009; Devasthale and Thomas, 2011; McIlhattan *et al.*, 2017]. Using these data, several studies have highlighted large cloud biases in climate models over Greenland [Van Tricht *et al.*, 2016a; McIlhattan *et al.*, 2017], the Arctic Ocean [English *et al.*, 2015; Boeke and Taylor, 2016], the Antarctic Plateau [Lawson and Gettelman, 2014], and the Southern Ocean [Kay *et al.*, 2016b], questioning their performance in representing polar climate and climate change [Bintanja and Krikken, 2016; Notz and Stroeve, 2016]. Detailed evaluation of clouds in climate models have been enabled in recent years by satellite simulators, such as the Observation Simulator Package (COSP, Bodas-Salcedo *et al.* [2011]), a lidar simulator [Chepfer *et al.*, 2008] with snow crystal error correction [English *et al.*, 2014], and a CALIPSO cloud-phase simulator [Cesana and Chepfer, 2013]. These studies have however been limited to small sub-regions, used only a single remote sensing dataset [Komurcu *et al.*, 2014], or evaluated only a single climate model [Chepfer *et al.*, 2008; English *et al.*, 2014; Kay *et al.*, 2016b]. Here we combine two independent remote sensing datasets to evaluate present-day clouds and downward radiation in atmospheric reanalyses and CMIP5 climate models over the entire Polar regions.

First we briefly present the observations (Section 2), we discuss results (Section 3) and end with conclusions (Section 4).

2. Data and methods

2.1. Satellite observations

NASA's CloudSat and CALIPSO (C-C hereafter) satellites currently are the only space-based combination of active cloud radar and lidar observations [*Stephens et al.*, 2002; *L'Ecuyer and Jiang*, 2010], providing the first near-global three-dimensional view on cloud macro-physical and micro-physical properties. Their active emission of radar and lidar pulses allows vertical probing of the atmospheric profile, and overcomes typical issues that passive remote sensing techniques in Polar regions suffer from, such as weak temperature and albedo differences between clouds and the underlying surface [*Kay et al.*, 2008]. In this study, we use a modified version of Release 04 (R04) of the CloudSat's Level 2B Fluxes and Heating Rates (2B-FLXHR-LIDAR) product [*Henderson et al.*, 2013] with specific adaptations for the Polar atmosphere [*Van Tricht et al.*, 2016a], combining cloud properties retrieved by C-C, reanalysis data from ERA-Interim, and surface properties, to drive a broadband radiative transfer model. The modified version of the 2B-FLXHR-LIDAR used here forms the basis of the new Release 05 (R05) of the algorithm, that will soon be made publicly available [*Matus and L'Ecuyer*, 2017]. In each 2B-FLXHR-LIDAR vertical profile, cloud boundaries of the different detected cloud layers are taken from the 2B-GEOPROF-LIDAR product. For each cloud layer, cloud phase (liquid, ice or mixed-phase) is adopted from the 2B-CLDCLASS-LIDAR product, which combines radar and lidar information for a better discrimination of cloud phase. Determination

of LWP and IWP in each cloud layer is performed in order of priority by taking inputs from CloudSat's 2B-CWC-RO radar-only product, CALIPSO lidar-only optical depths, and MODIS optical depths. It should be noted that the final R05 version of 2B-FLXHR-LIDAR will also use the CloudSat-CALIPSO 2C-ICE product which allows for an even more explicit determination of IWP [*Matus and L'Ecuyer, 2017*]. Near-realtime snow and ice cover are taken from passive microwave observations from AMSR-E provided by the National Snow and Ice Data Center (NSIDC). This information is used by C-C to calculate surface albedo for snow- and ice-covered regions. The resulting observationally constrained broadband SW and LW radiative fluxes have been shown to agree well with other observations on a global scale in general [*Henderson et al., 2013*], and in the polar regions in particular [*Kay and L'Ecuyer, 2013; Van Tricht et al., 2016a*]. The cloud properties reported in this study represent the combined C-C retrievals as they are input into the radiative flux calculations. Since CloudSat and CALIPSO observations were fully collocated only from 2007 to 2010, we focus our C-C analyses on this period.

C-C is compared with version 2.8 of the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product (C-E hereafter), which retrieve the SW and LW fluxes based on top of atmosphere (TOA) radiance measurements from passive sun-synchronous Terra, Aqua, and Suomi-National Polar-Orbiting Partnership satellites. In this retrieval algorithm, CERES footprint radiances are converted into instantaneous SW and LW fluxes using scene-dependent angular directional models (ADMs) based on MODIS cloud property retrievals and ancillary meteorological data from the the Goddard Earth Observing System 4 (GEOS-4) reanalysis product.

Over the Polar regions the cloud properties are derived from MODIS Terra and Aqua reflectances using two different retrieval algorithm depending on the presence of snow/ice derived from NISE1 or day/nighttime observations [Minnis *et al.*, 2011a]. The CERES instantaneous fluxes are subsequently converted into daily mean fluxes using diurnal interpolation approaches. The CERES processing incorporates two SW and three LW methods to derive the surface-only flux, based either on simple surface flux algorithms or on detailed radiative transfer computations [Kratz *et al.*, 2010]. Finally, the derived fluxes are energy balanced (EBAF) by filling gaps in the data and by adjusting the derived SW and LW fluxes to the observed CERES TOA fluxes [Kato *et al.*, 2013]. Validation of both the MODIS cloud property retrieval methods [Minnis *et al.*, 2011b] and CERES surface fluxes [Kratz *et al.*, 2010; Corbett *et al.*, 2012] has shown the agreement with other observations, but also has highlighted its uncertainties, especially for cloud properties such as LWP and IWP over Greenland, Antarctica and the Southern Hemisphere's sea ice and for LW fluxes over Polar regions.

2.2. Reanalyses and climate models

Atmospheric reanalyses are climate models that assimilate observations in order to represent the observed atmospheric state with the highest possible accuracy, making them ideal tools to study Polar climate variability and change. Here we include two state-of-the-art reanalyses, namely ERA-Interim [Dee and Uppala, 2009] and MERRA-2 [Molod *et al.*, 2015]. For both datasets, we have extracted cloud liquid and ice water paths, and downward longwave and shortwave radiation. It should be noted that ERA-Interim reanalysis data also provide the temperature and humidity profiles that are input to the

radiative transfer calculations in the 2B-FLXHR-LIDAR product, since this information cannot be retrieved by C-C observations. Consequently, only cloud microphysical properties and their impact on radiative fluxes can be considered independent between these two data sources.

To retrieve output from the Coupled Model Intercomparison Project (CMIP) of the International Governmental Panel on Climate Change (IPCC) Fifth Assessment Report (CMIP5 hereafter), cloud liquid (clwvi minus clivi , following CMIP5 standard output procedures) and ice water path (clivi) and downwelling radiation components at the surface (rlds : longwave downwelling; rsds = shortwave downwelling) were downloaded from the CMIP5 Data Portal. In total, data from 28 CMIP5 historical model simulations could be retrieved, but some of these use the same parent model with slightly differing setup or complexity; for the remainder of the CMIP5 ensemble, no clwvi data were available.

The data from the simulations with the same model were averaged, leaving 14 main models (i.e., CCSM4, CESM1-CAM4, CESM1-CAM5, CNRM, GFDL, GISS, HadGEM2, IPSL_CM5A, IPSL_CM5B, MPI-ESM, NorESM1, bcc-csm1 and inmcm4) for this study.

Note that we solely focus on vertically integrated cloud parameters, which are defined consistently the same in all models. To analyze vertical distributions and governing physical processes in more detail, we would require climate model satellite simulators that are not available for many of the models used in this study.

2.3. Data treatment

To compare the individual datasets over the Polar regions, we have subsetted them poleward of 60 degrees for the overlapping time period 2007-2010. The representativeness

of this short period for a longer period has been analysed by comparing the CERES-EBAF downward fluxes in the period 2007-2010 relative to 2002-2015. For all regions, the 2007-2010 area integrated downwelling fluxes deviate with less than 1 W m^{-2} from the 2002-2015 (Table S1), which leads us to conclude that the four years under study are representative for the period 2002-2015. To enable a fair spatial comparison, and because we do not use satellite simulators, we have regridded all datasets on a shared rectangular 2 degrees longitudinal and 1 degree latitudinal grid, which results from the trade-off between data quality (i.e. sufficient cross-overs in a single cell) and resolution. C-C has data gaps between 1 December 2009 and 21 January 2010; therefore, December 2009 was omitted in the analysis, and the last nine days of January 2010 were assumed representative for the entire month. Moreover, as CloudSat-CALIPSO does not cover the area poleward of 82°N and S, this area was not considered in calculating spatial averages. For CMIP5, we did not select the 2007-2010 period, because the historical CMIP5 simulations end before (in 2005), and CMIP5 climate models are not forced by observed meteorology, limiting the direct comparison with observations; instead, we selected the last 20 years of the CMIP5 historical simulations (1986-2005). This choice is justified by the much larger differences between models than the inter-annual variability produced by each individual model [*Van Tricht et al., 2016a*].

3. Results

3.1. Satellite intercomparison

First we compare the 2007-2010 mean cloud liquid and ice water path over the Polar regions as retrieved from C-C and C-E observations. Figure 1 (first column) shows the

annual mean LWP and IWP as retrieved from C-C. Both LWP and IWP are very small over the interior of the ice sheets ($<0.01 \text{ kg m}^{-2}$), increasing towards the coast. Water paths are much higher over the ocean, especially in regions with frequent storm passages, abundant cloud cover, and precipitation: southeast of Greenland (both LWP and IWP $>0.1 \text{ kg m}^{-2}$) and north of the Amundsen coast in West Antarctica (LWP $> 0.07 \text{ kg m}^{-2}$ and IWP $> 0.1 \text{ kg m}^{-2}$).

Figure 2a shows the seasonal cycle of water path over each of the four regions, and illustrates that water paths are lower over the ice sheets than over the ocean. Lowest LWP and IWP values are observed over Antarctica, while the Arctic and Southern Ocean have comparable cloud liquid water. The amount of ice is clearly higher over the Southern Ocean. LWP exhibits a strong seasonal cycle in all regions, with a maximum in summer. In contrast, IWP is relatively constant over the year, with subtle maxima in Spring and Fall seasons over the oceans, which are concurrent with maxima in precipitation over the ice sheets [Lenaerts *et al.*, 2012].

Compared to C-C, C-E shows higher values of LWP and IWP over most of the Polar regions (Figure 1), and across all seasons (Figure 2). Over the ice sheets, mean C-E ice and water paths are a factor of 2 to 4 higher than those from C-C (Table S1). In contrast, C-E LWP is lower over the North Atlantic region (Figure 1). LWP values over the Arctic and Southern Oceans are comparable in both datasets (Figure 2a, Table S1).

The contrast of cloud characteristics between the ice sheets and the Polar oceans are reflected in the downwelling fluxes (Figure 3). The annual mean downwelling longwave radiation (LW_d) as retrieved from C-C decreases from the coastal regions ($>250 \text{ W m}^{-2}$)

to the interior of Greenland and Antarctica ($<150 \text{ W m}^{-2}$), where the atmosphere is thin and stable, atmospheric temperatures and humidities are low, and clouds are thin, all limiting longwave energy emission. In contrast, downwelling shortwave radiation (SW_d) is higher ($>150 \text{ W m}^{-2}$) over the ice sheets, owing to fewer, thinner clouds and higher surface albedo, enhancing multiple SW scattering between surface and atmosphere. Both longwave and shortwave fluxes exhibit strong seasonal variability (Figure 2b).

Downwelling longwave radiation estimates of C-C and C-E agree within their uncertainties ($\sim 8 \text{ W m}^{-2}$ for C-C [*L'Ecuyer et al.*, 2015] and 13.3 W m^{-2} in C-E [*Boeke and Taylor*, 2016]) over Greenland (Table S1). However, the seasonal cycle is much less pronounced in C-C relative to C-E. Downwelling longwave radiation is much higher in C-E over interior East Antarctica ($>20 \text{ W m}^{-2}$), especially in winter (Figure 2b) invoking higher annual mean LW_d values in C-E relative to C-C over Antarctica. For downwelling shortwave radiation SW_d , C-E diverges considerably from C-C over Greenland, Southern Ocean and especially the Arctic Ocean, but is more consistent with C-C over Antarctica (Table S1). To further assess the absolute quality of these satellite-based downwelling fluxes, and extending earlier C-C [*Van Tricht et al.*, 2016b] and C-E [*Zhang et al.*, 2016; *Christensen et al.*, 2016] evaluation studies over Greenland, we compared longwave and shortwave fluxes with five Baseline Surface Radiation Network (BSRN, *Ohmura et al.* [1998]; *Roesch et al.* [2011]) ground stations (Figures S2 and S3), two in the Arctic and two over Antarctica. All stations (except for Dome-C) are located near the coast. The spatial sampling of the satellite products unavoidably leads to the grid cells containing mixed retrievals from ocean and land, complicating direct comparisons to the ground stations.

Therefore, we have considered the nine grid points surrounding the station to estimate the spatial variability in the satellite products. For longwave radiation, we find comparable biases (0 to 15 W m⁻²) in C-C and C-E. However, the spatial variability, driven by land-ocean contrasts, is generally better represented in C-C, with lower standard deviation of the biases in the grid points surrounding the station (Figure S2). Spatial variability is very large in both remote sensing products at Syowa station, which is located in an area with a very strong meridional LW_d gradient from the Antarctic interior to the ocean. At Dome C, C-C clearly outperforms C-E, which signals that C-E overestimates LW_d over interior Antarctica (Figure 3). Figure S3 illustrates that C-C tends to overestimate SW_d at the stations, whereas C-E shows a comparable, but negative bias. C-C shows clearly more realistic SW_d at Neumayer and Dome-C stations.

Apart from the better representation of surface downwelling fluxes, we believe that the representation of clouds is superior in C-C in comparison to C-E. Not only are the active radar and lidar sensors in C-C insensitive to day/night differences, changes in surface characteristics, and temperature inversions (when infrared channels are used), but C-C uses MODIS observations as well. Hence, the C-C based 2B-FLXHR-LIDAR algorithm has the benefit of 3 sensors as opposed to only one used in C-E. Especially the explicit inclusion of cloud phase determination by the CLDCLASS-LIDAR combined CloudSat-CALIPSO product [Sassen and Wang, 2012] results in an important advantage with regard to cloud treatment in the radiative flux calculations over the passive-only cloud detection by MODIS in the CERES-EBAF algorithm, especially with regard to the ability to detect the frequently-occurring low-level supercooled liquid-bearing clouds over the

Polar regions [Bennartz *et al.*, 2013; Van Tricht *et al.*, 2016a; Matus and L'Ecuyer, 2017].

Mace and Wrenn [2013] showed that the same TOA radiance signatures can originate from a multitude of distinct cloud scenes that can only be distinguished through the addition of active sensors. That is in addition to C-C's ability to perform retrievals in low sunlight or over sea ice that are known to be extremely difficult for passive sensors.

Additionally, C-E finds unrealistically high IWP values over ice sheets, mainly over interior Antarctica. An issue with the C-E approach is that in these areas, where the contrast between surface and cloud albedo is small, a large correction of cloud water path is necessary to match the TOA fluxes since they are insensitive to small changes. Furthermore, since LWP has limited observational constraints in these regions, the algorithm likely has to resort to increasing the IWP dramatically to compensate for any lack of brightness owing to missing liquid. Besides the issues in cloud representation, we can attribute differences between C-C and C-E to differences in vertical atmospheric temperature and humidity profiles, the radiative transfer model, and possibly skin temperature and surface albedo, although these are likely small compared to the cloud issues.

We conclude that the cloud water path algorithm based exclusively on passive satellite observations in C-E is only partly suitable to accurately detect cloud water over high latitudes, and especially fails over the ice sheets. Since C-C outperforms C-E in several aspects of its representation of high latitude clouds and radiation, we use this dataset for model evaluation, although the reader should bear in mind that satellite remote sensing retrievals of cloud properties and radiative fluxes are unavoidably prone to substantial uncertainties.

3.2. Model biases: clouds

Although the models agree with C-C in the general spatial patterns of cloud water and ice, the amount of LWP and IWP deviates substantially from the C-C observations (Figure 1). ERA-Interim underestimates the amount of cloud water over the Polar regions, both in the form of ice and water, and especially over the oceans. MERRA-2 produces generally too many liquid clouds, at the expense of ice. This is especially evident over the Southern Ocean, where LWP bias exceeds 0.04kg m^{-2} throughout the entire year, compensated by a consistent, severe ($<-0.06\text{ kg m}^{-2}$) underestimation of cloud IWP (Figure 2a). The CMIP5 multi-model mean is consistent with MERRA-2 and suggests excessive LWP values, and too little IWP, albeit with smaller discrepancies with respect to C-C (Figure 1). The comparable magnitude of the spread in the CMIP5 multi-model ensemble for LWP and the actual observed C-C LWP indicates that the simulated uncertainty of LWP is very substantial, and caution is warranted in drawing conclusions from looking at CMIP5 mean only. Especially over the ice sheets, the differences in LWP across models is extremely large (Figure S4), with some models (CESM-CAM5, HadGEM2) simulating almost no LWP, and some models (bcc-csm1, NorESM1) simulating LWP values that are up to 10 times larger than in the observations. Excluding one of these outliers will affect the CMIP5 mean considerably, limiting use of the multi-model mean as representation of the average model behaviour. The same holds true, albeit to a lesser extent, for cloud IWP.

Instead, we have used the 10-90% probability in the 14 available models along with the CMIP5 mean in Figure 2a. The large difference between the 10 and 90% percentiles for LWP confirm that CMIP5 models disagree widely on the amounts of LWP in the Polar

regions. Moreover, given that the 10 and 90% percentiles indicate a bias of opposite sign, no conclusions can be drawn regarding a consistent excess or lack of simulated LWP in CMIP5 (Figure S4). For IWP, the CMIP5 spread appears to be much smaller, although the Southern Ocean stands out with considerable disagreement in CMIP5.

3.3. Model biases: radiation

Compared to C-C, the models generally show too little LW_d (Figure 3, Table S1), although regional and seasonal variations are substantial (Figure 2b). In ERA-Interim, the bias is particularly present on the Antarctic ice shelves and sea ice. In MERRA-2, LW_d is mainly underestimated over parts of the ice sheets, and over the North Atlantic, while LW_d is excessive over the northern Arctic Ocean. In CMIP5 we find the highest biases over the Arctic Ocean and coastal Antarctica. Similar to what we found with LWP and IWP, the CMIP5 ensemble is not equivocal in the simulation of LW_d in the Polar regions, and a large spread exists between the CMIP5 models (Figure 2b). For downwelling shortwave radiation, a clearer picture emerges, as all models underestimate SW_d consistently over the seasons and regions. The highest biases are found in ERA-Interim over the Arctic Ocean, with a SW_d deficit exceeding -40 W m^{-2} . However, we should be cautious in interpreting this signal, since C-C tends to overestimate SW_d in coastal areas (Figure S4), a bias that is probably enhanced over the Arctic Ocean where surface albedos are poorly constrained [Van Tricht *et al.*, 2016b].

4. Discussion and conclusions

We used state-of-the-art, gridded remote sensing observations to evaluate simulated cloud water paths and surface downwelling radiation components in the Polar regions by

climate models and reanalyses. First, we compared the independent observational estimates of clouds and downwelling radiation from CERES-EBAF with those from CloudSat-CALIPSO. While cloud water paths as retrieved by C-C and C-E are relatively similar over the oceans, our results indicate large differences in LWP and IWP values over the ice sheets. For downwelling fluxes, a comparison to ground stations suggests that both satellite remote sensing datasets exhibit significant uncertainties. The spatial gridding of satellite tracks enhances errors in coastal regions, where most stations are located [*Van Tricht et al.*, 2016b]. Moreover, the crude representation of surface albedo in the satellite product algorithms, in combination with fixed overpass times, leads to significant uncertainties in downwelling shortwave radiation retrievals, especially over the ice sheets and over sea ice.

Since C-C includes a more sophisticated representation of cloud properties, overall lower bias than C-E compared to ground observations, and lower spatial inconsistencies in C-C, justify its application to evaluate simulated cloud water path in two state-of-the-art reanalyses (ERA-Interim and MERRA-2) and the CMIP5 model ensemble. Overall, it appears that total cloud water path (liquid + ice) is underestimated in nearly all models. However, the cloud water path (LWP and IWP) biases are inconsistent among models, precluding any conclusions on shared model deficiencies. For example, ERA-Interim reanalysis shows too little cloud water over the Arctic, while MERRA-2 reanalysis simulates realistic cloud LWP but severely underestimates the total amount of cloud ice (IWP). Additionally, the CMIP5 model ensemble gives a wide range of cloud water path simulations. For downwelling radiation, the climate models tend to underestimate longwave downwelling

radiation, especially over the ice sheets. Although observational uncertainties are considerable for SW_d , models also underestimate shortwave downwelling radiation, especially over the oceans.

Our results are mostly in line with previous studies that evaluate clouds and radiation in climate models over the Arctic. We confirm existing work of *Liu and Key* [2016] which showed that atmospheric reanalysis products exhibit significant differences in their simulation of Arctic clouds, and differences with observations of IWP and LWP are largest in summer. The large underestimation of summer SW_d in CMIP5 over the Arctic is consistent with *English et al.* [2015] and *Karlsson and Svensson* [2013], who attribute these deficiencies to underestimated cloud amounts and sea ice albedo biases. Consistently with *Boeke and Taylor* [2016], we find small summer LW_d biases over the Arctic Ocean that are much larger in winter, and very large summer SW_d biases. Here we further show that, while the SW_d bias is shared between the Arctic and Antarctic regions, the LW_d bias is much smaller over Antarctica, and especially over the Southern Ocean.

Our results indicate that differences across CMIP5 models are substantial enough to preclude any definite conclusion on a shared bias. As a consequence, a direct comparison with earlier studies that suggest that cloud LWP [*Komurcu et al.*, 2014] and IWP [*Liu et al.*, 2012] are overestimated in CMIP5 is challenged by the choice of other models in the comparison. Some CMIP5 models show a substantial bias in cloud water and ice path, simulating either far too little or clearly excessive values compared to C-C. A few models are more consistent with the observations, including CNRM, GISS, GFDL and the IPSL_CM5b models. Although a correct simulation of clouds does not solely contribute

to a realistic Arctic climate, clouds exert a strong control on modeled climate sensitivity [Vavrus, 2004] and changes in surface radiation [Notz and Stroeve, 2016]. Our results suggest therefore that Arctic and Antarctic climate change simulated by these models can be considered more credible.

Our work shows that climate models exhibit strong, but partly inconsistent, biases in their simulation of clouds and radiation over the Polar regions. We suggest that future model development should focus on improving polar cloud micro-physical properties and assessing their radiative impact on the surface. At the same time, efforts should be made in reducing the observational uncertainties of clouds and their radiative impact.

To achieve those goals, observations from combined active and passive remote sensing, such as CloudSat-CALIPSO and the upcoming EarthCARE Satellite Mission [*Illingworth et al., 2015*] are crucial.

Acknowledgments. Jan Lenaerts acknowledges support from NWO ALW through its Innovational Research Incentives Scheme Veni. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. All data used in this study are available online. Satellite observations from CloudSat-CALIPSO are available at <http://www.cloudsat.cira.colostate.edu/>, and from CERES-EBAF via <https://ceres.larc.nasa.gov/products.php?product=EBAF>.

Surface. ERA-Interim data are downloaded from the ECWMF web server (<http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/>), and MERRA-2 via the Modeling and Assimilation Data and Information Services Center (MDISC, <http://disc.sci.gsfc.nasa.gov/mdi/>). CMIP5 data are available on the Earth System Grid - Center for Enabling Technologies (ESG-CET) CMIP5 data portal (<https://pcmdi.llnl.gov>). BSRN data used in this study are available at <http://bsrn.awi.de/en/data/>.

References

- Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom (2016), Southern Ocean warming delayed by circumpolar upwelling and equatorward transport, *Nat. Geosci.*, 9(7), 549–554, doi:10.1038/ngeo2731.
- Bennartz, R., M. D. Shupe, D. D. Turner, V. P. Walden, K. Steffen, C. J. Cox, M. S. Kulie, N. B. Miller, and C. Pettersen (2013), July 2012 Greenland melt extent enhanced by low-level liquid clouds., *Nature*, 496(7443), 83–6, doi:10.1038/nature12002.
- Bintanja, R., and F. Krikken (2016), Magnitude and pattern of Arctic warming governed by the seasonality of radiative forcing, *Sci. Rep.*, 6, 38,287.
- Blunden, J., and D. S. Arndt (2016), State of the Climate in 2015, *Bull. Am. Meteorol. Soc.*, 97(8), Si–S275, doi:10.1175/2016BAMSStateoftheClimate.1.
- Bodas-Salcedo, A., et al. (2011), COSP: Satellite simulation software for model assessment, *Bull. Am. Meteorol. Soc.*, 92(8), 1023–1043, doi:10.1175/2011BAMS2856.1.
- Boeke, R. C., and P. C. Taylor (2016), Evaluation of the Arctic surface radiation budget in CMIP5 models, *J. Geophys. Res. Atmos.*, 121(14), 8525–8548, doi:

10.1002/2016JD025099.

Bromwich, D. H., et al. (2012), Tropospheric clouds in Antarctica, *Rev. Geophys.*, 50(1), RG1004, doi:10.1029/2011RG000363.

Cesana, G., and H. Chepfer (2013), Evaluation of the cloud thermodynamic phase in a climate model using CALIPSO-GOCCP, *J. Geophys. Res. Atmos.*, 118(14), 7922–7937, doi:10.1002/jgrd.50376.

Cesana, G., J. E. Kay, H. Chepfer, J. M. English, and G. de Boer (2012), Ubiquitous low-level liquid-containing Arctic clouds: New observations and climate model constraints from CALIPSO-GOCCP, *Geophys. Res. Lett.*, 39(20), doi:10.1029/2012GL053385.

Chepfer, H., S. Bony, D. Winker, M. Chiriaco, J. L. Dufresne, and G. Sèze (2008), Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model, *Geophys. Res. Lett.*, 35(15), doi:10.1029/2008GL034207.

Christensen, M. W., A. Behrangi, T. S. L'Ecuyer, N. B. Wood, M. D. Lebsack, and G. L. Stephens (2016), Arctic observation and reanalysis integrated system: A new data product for validation and climate study, *Bull. Am. Meteorol. Soc.*, 97(6), 907–915, doi:10.1175/BAMS-D-14-00273.1.

Cook, A. J., P. R. Holland, M. P. Meredith, T. Murray, A. Luckman, and D. G. Vaughan (2016), Ocean forcing of glacier retreat in the western Antarctic Peninsula, *Science.*, 353(6296), 1261–1273, doi:10.1126/science.aae0017.

Corbett, J. G., W. Su, and N. G. Loeb (2012), Observed effects of sastrugi on CERES top-of-atmosphere clear-sky reflected shortwave flux over Antarctica, *J. Geophys. Res. Atmos.*, 117(17), doi:10.1029/2012JD017529.

Dee, D. P., and S. Uppala (2009), Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, *Quart. J. Roy. Meteorol. Soc.*, 135, 1830–1841, doi: 10.1002/qj.493.

Devasthale, A., and M. A. Thomas (2011), A global survey of aerosol-liquid water cloud overlap based on four years of CALIPSO-CALIOP data, *Atmos. Chem. Phys.*, 11(3), 1143–1154, doi:10.5194/acp-11-1143-2011.

English, J. M., J. E. Kay, A. Gettelman, X. Liu, Y. Wang, Y. Zhang, and H. Chepfer (2014), Contributions of Clouds, Surface Albedos, and Mixed-Phase Ice Nucleation Schemes to Arctic Radiation Biases in CAM5, *J. Clim.*, 27(13), 5174–5197, doi: 10.1175/JCLI-D-13-00608.1.

English, J. M., A. Gettelman, and G. R. Henderson (2015), Arctic Radiative Fluxes: Present-Day Biases and Future Projections in CMIP5 Models, *J. Clim.*, 28(15), 6019–6038, doi:10.1175/JCLI-D-14-00801.1.

Grenier, P., J. P. Blanchet, and R. Muñoz-Alpizar (2009), Study of polar thin ice clouds and aerosols seen by CloudSat and CALIPSO during midwinter 2007, *J. Geophys. Res. Atmos.*, 114(9), D09201, doi:10.1029/2008JD010927.

Henderson, D. S., T. L'Ecuyer, G. Stephens, P. Partain, and M. Sekiguchi (2013), A multisensor perspective on the radiative impacts of clouds and aerosols, *J. Appl. Meteorol. Climatol.*, 52(4), 853–871, doi:10.1175/JAMC-D-12-025.1.

Illingworth, A. J., et al. (2015), The EarthCARE Satellite: The Next Step Forward in Global Measurements of Clouds, Aerosols, Precipitation, and Radiation, *Bull. Am. Meteorol. Soc.*, 96(8), 1311–1332, doi:10.1175/BAMS-D-12-00227.1.

Karlsson, J., and G. Svensson (2013), Consequences of poor representation of Arctic sea-ice albedo and cloud-radiation interactions in the CMIP5 model ensemble, *Geophys. Res. Lett.*, 40(16), 4374–4379, doi:10.1002/grl.50768.

Kato, S., N. G. Loeb, F. G. Rose, D. R. Doelling, D. A. Rutan, T. E. Caldwell, L. Yu, and R. A. Weller (2013), Surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances, *J. Clim.*, 26(9), 2719–2740, doi:10.1175/JCLI-D-12-00436.1.

Kay, J. E., and A. Gettelman (2009), Cloud influence on and response to seasonal Arctic sea ice loss, *J. Geophys. Res.*, 114(D18), D18,204, doi:10.1029/2009JD011773.

Kay, J. E., and T. L'Ecuyer (2013), Observational constraints on Arctic Ocean clouds and radiative fluxes during the early 21st century, *J. Geophys. Res. Atmos.*, 118(13), 7219–7236, doi:10.1002/jgrd.50489.

Kay, J. E., T. L'Ecuyer, A. Gettelman, G. Stephens, and C. O'Dell (2008), The contribution of cloud and radiation anomalies to the 2007 Arctic sea ice extent minimum, *Geophys. Res. Lett.*, 35(8), 1–5, doi:10.1029/2008GL033451.

Kay, J. E., T. L'Ecuyer, H. Chepfer, N. Loeb, A. Morrison, and G. Cesana (2016a), Recent Advances in Arctic Cloud and Climate Research, *Curr. Clim. Chang. Reports*, pp. 1–11, doi:10.1007/s40641-016-0051-9.

Kay, J. E., C. Wall, V. Yettella, B. Medeiros, C. Hannay, P. Caldwell, and C. Bitz (2016b), Global Climate Impacts of Fixing the Southern Ocean Shortwave Radiation Bias in the Community Earth System Model (CESM), *J. Clim.*, 29(12), 4617–4636, doi:10.1175/JCLI-D-15-0358.1.

King, J. C., et al. (2015), Validation of the summertime surface energy budget of Larsen

C Ice Shelf (Antarctica) as represented in three high-resolution atmospheric models, *J.*

Geophys. Res. Atmos., 120(4), 1335–1347, doi:10.1002/2014JD022604.

Komurcu, M., T. Storelvmo, I. Tan, U. Lohmann, Y. Yun, J. E. Penner, Y. Wang, X. Liu, and T. Takemura (2014), Intercomparison of the cloud water phase among global climate models, *J. Geophys. Res. Atmos.*, 119(6), 3372–3400, doi:10.1002/2013JD021119.

Kratz, D. P., S. K. Gupta, A. C. Wilber, and V. E. Sothcott (2010), Validation of the CERES edition 2B surface-only flux algorithms, *J. Appl. Meteorol. Climatol.*, 49(1), 164–180, doi:10.1175/2009JAMC2246.1.

Lawson, R. P., and A. Gettelman (2014), Impact of Antarctic mixed-phase clouds on climate., *Proc. Natl. Acad. Sci. U. S. A.*, 111(51), 18,156–61, doi: 10.1073/pnas.1418197111.

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

L'Ecuyer, T. S., et al. (2015), The observed state of the energy budget in the early twenty-first century, *J. Clim.*, 28(21), 8319–8346, doi:10.1175/JCLI-D-14-00556.1.

Lenaarts, J. T. M., M. R. van den Broeke, W. J. van de Berg, E. Van Meijgaard, and P. Kuipers Munneke (2012), A new, high-resolution surface mass balance map of Antarctica (1979–2010) based on regional atmospheric climate modeling, *Geophys. Res. Lett.*, 39(L04501), doi:10.1029/2011GL050713.

Liu, X., et al. (2012), Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5, *Geosci. Model*

Dev., 5(3), 709–739, doi:10.5194/gmd-5-709-2012.

Liu, Y., and J. R. Key (2016), Assessment of Arctic Cloud Cover Anomalies in Atmospheric Reanalysis Products Using Satellite Data, *J. Clim.*, 29(17), 6065–6083, doi:10.1175/JCLI-D-15-0861.1.

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. Clim.*, 26(23), 9429–9444, doi:10.1175/JCLI-D-12-00207.1.

Maslanik, J. A., C. Fowler, J. Stroeve, S. Drobot, J. Zwally, D. Yi, and W. Emery (2007), A younger, thinner Arctic ice cover: Increased potential for rapid, extensive sea-ice loss, *Geophys. Res. Lett.*, 34(24), L24,501, doi:10.1029/2007GL032043.

Matus, A. V., and T. S. L'Ecuyer (2017), The role of cloud phase in Earth's radiation budget, *J. Geophys. Res. Atmos.*, doi:10.1002/2016JD025951.

McIlhattan, E. A., T. S. L'Ecuyer, and N. B. Miller (2017), Observational Evidence Linking Arctic Supercooled Liquid Cloud Biases in CESM to Snowfall Processes, *J. Clim.*, doi:10.1175/JCLI-D-16-0666.1.

McMillan, M., et al. (2016), A high-resolution record of Greenland mass balance, *Geophys. Res. Lett.*, 43(13), 7002–7010, doi:10.1002/2016GL069666.

Meehl, G. A., J. M. Arblaster, C. M. Bitz, C. T. Y. Chung, and H. Teng (2016), Antarctic sea-ice expansion between 2000 and 2014 driven by tropical Pacific decadal climatevariability, *Nat. Geosci.*, 9(8), 590–595, doi:10.1038/ngeo2751.

Minnis, P., et al. (2011a), CERES Edition-2 Cloud Property Retrievals Using TRMM VIRS and Terra and Aqua MODIS Data-Part I: Algorithms, *IEEE Trans. Geosc. Remote Sens.*, 49(11), 4374–4400, doi:10.1109/TGRS.2011.2144601.

Minnis, P., et al. (2011b), CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data-Part II: Examples of average results and comparisons with other data, *IEEE Trans. Geosci. Remote Sens.*, 49(11), 4401–4430, doi:10.1109/TGRS.2011.2144602.

Molod, A., L. Takacs, M. Suarez, and J. Bacmeister (2015), Development of the GEOS-5 atmospheric general circulation model: evolution from MERRA to MERRA2, *Geosci. Model Dev.*, 8(5), 1339–1356, doi:10.5194/gmd-8-1339-2015.

Mouginot, J., E. Rignot, and B. Scheuchl (2014), Sustained increase in ice discharge from the Amundsen Sea Embayment, West Antarctica, from 1973 to 2013, *Geophys. Res. Lett.*, 41(5), 1576–1584, doi:10.1002/2013GL059069.

Notz, D., and J. Stroeve (2016), Observed Arctic sea-ice loss directly follows anthropogenic CO₂ emission, *Science.*, pp. 1–9, doi:10.1126/science.aag2345.

Ohmura, A., et al. (1998), Baseline Surface Radiation Network (BSRN/WCRP): New Precision Radiometry for Climate Research, *Bull. Am. Meteorol. Soc.*, 79(10), 2115–2136, doi:10.1175/1520-0477(1998)079;2115:BSRNBW;2.0.CO;2.

Roesch, A., M. Wild, A. Ohmura, E. G. Dutton, C. N. Long, and T. Zhang (2011), Assessment of BSRN radiation records for the computation of monthly means, *Atmos. Meas. Tech.*, 4(2), 339–354, doi:10.5194/amt-4-339-2011.

Sassen, K., and Z. Wang (2012), The Clouds of the Middle Troposphere: Composition, Radiative Impact, and Global Distribution, *Surv. Geophys.*, 33(3), 677–691, doi: 10.1007/s10712-011-9163-x.

Scambos, T. a., C. L. Hulbe, M. Fahnestock, and J. Bohlander (2000), The Link between Climate Warming and Ice Shelf Breakups in the Antarctic Peninsula, *J. Glaciol.*, 46(154)(154), 516–530.

Stephens, G. L., et al. (2002), The CloudSat mission and the A-train: A new dimension of space-based observations of clouds and precipitation, *Bull. Am. Meteorol. Soc.*, 83(12), 1771–1790, doi:10.1175/BAMS-83-12-1771.

Van den Broeke, M. (2005), Strong surface melting preceded collapse of Antarctic Peninsula ice shelf, *Geophys. Res. Lett.*, 32(12), 1–4, doi:10.1029/2005GL023247.

Van den Broeke, M. R., E. M. Enderlin, I. M. Howat, P. Kuipers Munneke, B. P. Y. Noël, W. J. van de Berg, E. van Meijgaard, and B. Wouters (2016), On the recent contribution of the Greenland ice sheet to sea level change, *Cryosph.*, 10(5), 1933–1946, doi:10.5194/tc-10-1933-2016.

Van Tricht, K., S. Lhermitte, J. T. M. Lenaerts, I. V. Gorodetskaya, T. S. L'Ecuyer, B. Noël, M. R. van den Broeke, D. D. Turner, and N. P. M. van Lipzig (2016a), Clouds enhance Greenland ice sheet meltwater runoff, *Nat. Commun.*, 7(May), 10,266, doi: 10.1038/ncomms10266.

Van Tricht, K., S. Lhermitte, I. V. Gorodetskaya, and N. P. M. van Lipzig (2016b), Improving satellite-retrieved surface radiative fluxes in polar regions using a smart sampling approach, *Cryosph.*, 10(5), 2379–2397, doi:10.5194/tc-10-2379-2016.

Vavrus, S. (2004), The impact of cloud feedbacks on Arctic climate under Greenhouse forcing, *J. Clim.*, 17(3), 603–615, doi:10.1175/1520-0442(2004)017j0603:TIOCFO;2.0.CO;2.

Wouters, B., A. Martin-Espanol, V. Helm, T. Flament, J. M. van Wessem, S. R. M. Ligtenberg, M. R. van den Broeke, and J. L. Bamber (2015), Dynamic thinning of glaciers on the Southern Antarctic Peninsula, *Science.*, 348(6237), 899–903, doi: 10.1126/science.aaa5727.

Zhang, X., S. Liang, G. Wang, Y. Yao, B. Jiang, and J. Cheng (2016), Evaluation of the reanalysis surface incident shortwave radiation products from NCEP, ECMWF, GSFC, and JMA using satellite and surface observations, *Remote Sens.*, 8(3), doi: 10.3390/rs8030225.

Accepted

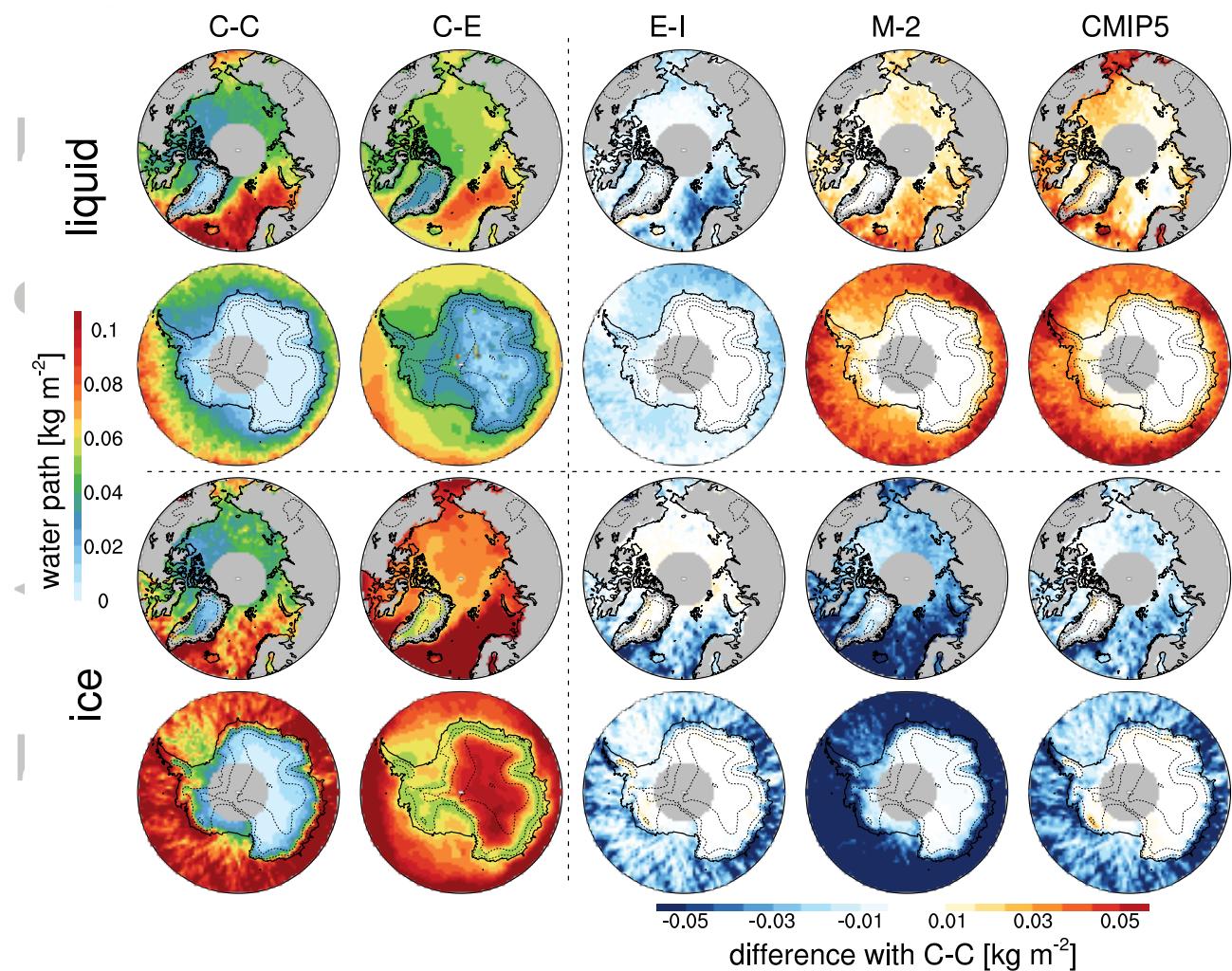


Figure 1. Cloud liquid (top two rows) and ice (bottom two rows) for the Arctic Ocean and Greenland (Arctic land regions are shown in grey and are not considered in this study) and the Southern Ocean and Antarctica according the CloudSat-CALIPSO (C-C) (first column). The ocean regions include both ocean and sea ice covered pixels. The simulated cloud LWP and IWP in CERES-EBAF (C-E) ERA-Interim (E-I), MERRA-2 (M-2) and CMIP5 (avg) are plotted as absolute differences with C-C in column two to five, respectively.

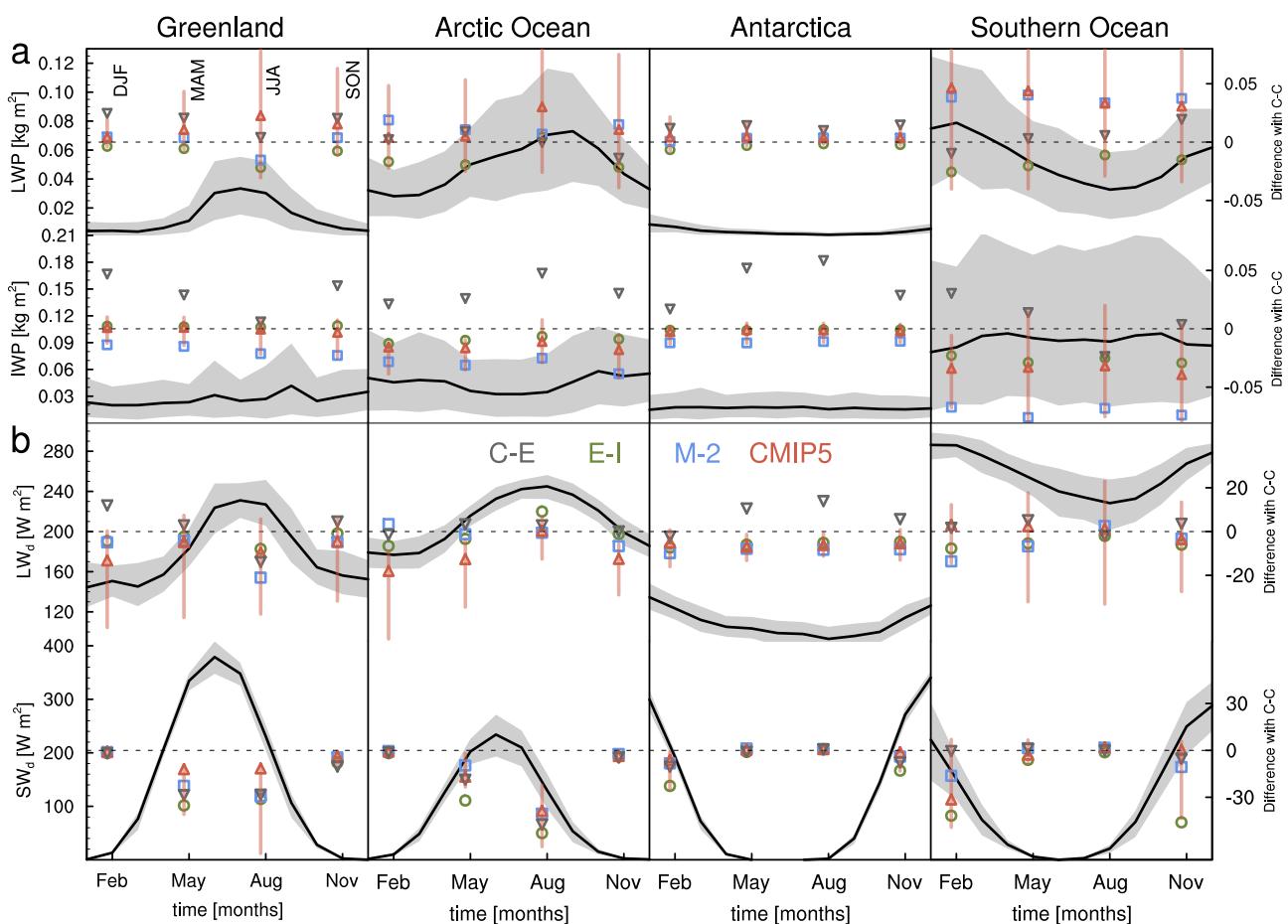


Figure 2. Seasonal cycle of area integrated cloud LWP and IWP (a) and LW_d and SW_d (b) according to C-C. The 2007-2010 mean is given in the black line, with the grey shading representing the difference between minimum and maximum observed value in each month. The model results are plotted as differences with respect to C-C (right axes) for each season. E-I is shown in green circles, MERRA-2 in blue squares, CMIP5 mean in red triangles, and CERES-EBAF in grey triangles. The vertical red lines indicate the difference between the 10 and 90 percentiles of CMIP5. Greenland and Antarctic ice sheets only contain grid cells with 100% ice cover, while Arctic and Southern Oceans are defined as all grid cells poleward of 60 degrees containing 100% water.

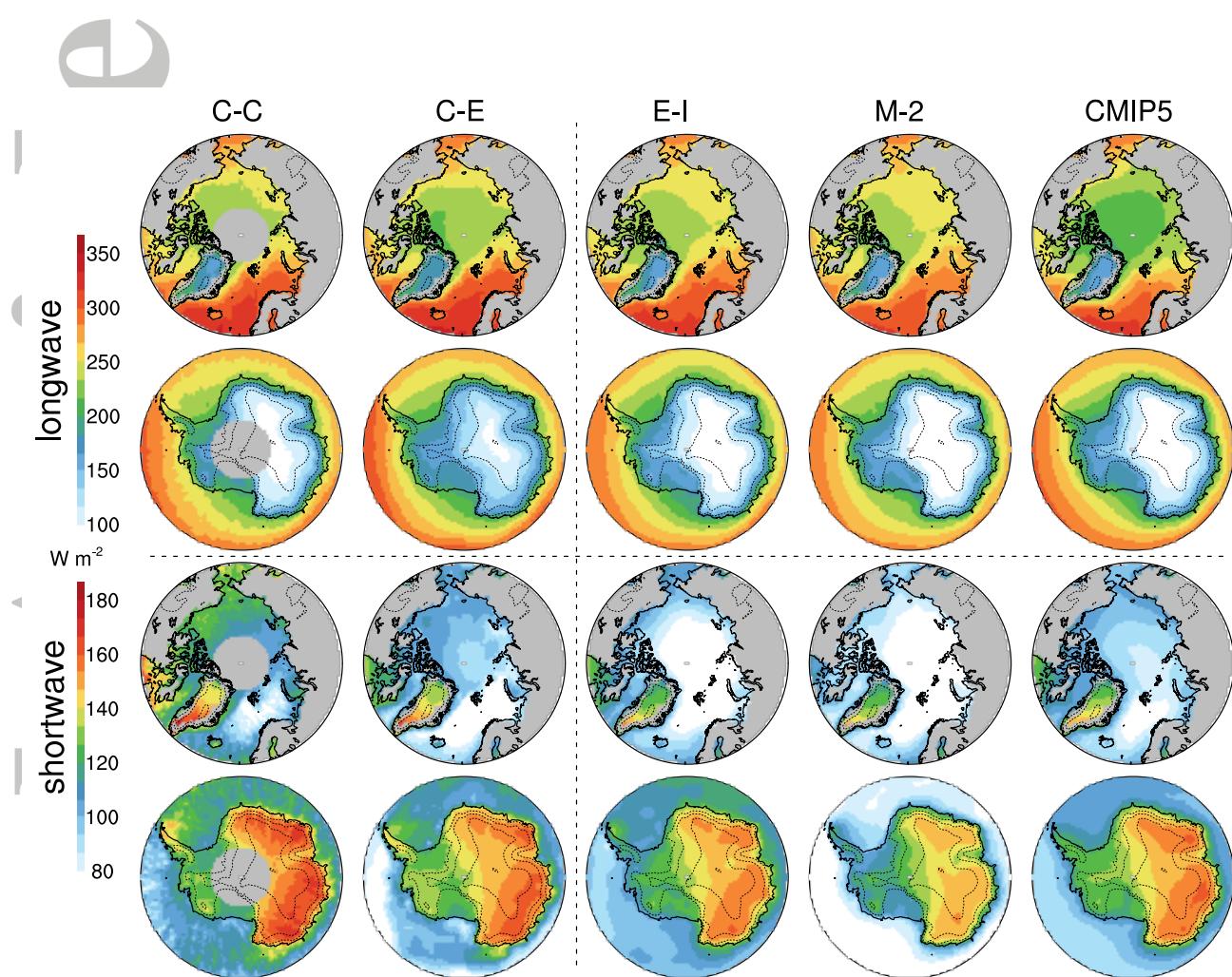


Figure 3. Downwelling longwave (top two rows) and shortwave (bottom two rows) surface fluxes for the Arctic Ocean and Greenland (Arctic land regions are shown in grey and are not considered in this study) and the Southern Ocean and Antarctica according the CloudSat-CALIPSO (C-C), from CERES-EBAF (C-E), and simulated LW_d and SW_d in ERA-Interim (E-I), MERRA-2 (M-2) and the ensemble mean of CMIP5.