Systematic and random errors between collocated satellite ice water path observations

Systematic and random errors between collocated satellite ice water path observations

S. Eliasson,[1] G. Holl,[1] S. A. Buehler,[1] T. Kuhn,[1] M. Stengel,[2] F. Iturbide-Sanchez,[3] and M. Johnston[4]

Received 29 June 2012; revised 5 October 2012; accepted 15 November 2012; published 25 March 2013.

[1] There remains large disagreement between ice-water path (IWP) in observational data sets, largely because the sensors observe different parts of the ice particle size distribution. A detailed comparison of retrieved IWP from satellite observations in the Tropics (±30° latitude) in 2007 was made using collocated measurements. The radio detection and ranging(radar)/light detection and ranging (lidar) (DARDAR) IWP data set, based on combined radar/lidar measurements, is used as a reference because it provides arguably the best estimate of the total column IWP. For each data set, usable IWP dynamic ranges are inferred from this comparison. IWP retrievals based on solar reflectance measurements, in the moderate resolution imaging spectroradiometer (MODIS), advanced very high resolution radiometer–based Climate Monitoring Satellite Applications Facility (CMSAF), and Pathfinder Atmospheres-Extended (PATMOS-x) datasets, were found to be correlated with DARDAR over a large IWP range (~20–7000 g m^{-2}). The random errors of the collocated data sets have a close to lognormal distribution, and the combined random error of MODIS and DARDAR is less than a factor of 2, which also sets the upper limit for MODIS alone. In the same way, the upper limit for the random error of all considered data sets is determined. Data sets based on passive microwave measurements, microwave surface and precipitation products system (MSPPS), microwave integrated retrieval system (MiRS), and collocated microwave only (CMO), are largely correlated with DARDAR for IWP values larger than approximately 700 g m^{-2}. The combined uncertainty between these data sets and DARDAR in this range is slightly less MODIS-DARDAR, but the systematic bias is nearly an order of magnitude.

Citation: Eliasson, S., G. Holl, S. A. Buehler, T. Kuhn, M. Stengel, F. Iturbide-Sanchez, and M. Johnston (2013), Systematic and random errors between collocated satellite ice water path observations, *J. Geophys. Res. Atmos.*, *118*, 2629–2642, doi:10.1029/2012JD018381.

1. Introduction

[2] Clouds have a dominant effect on the radiation entering and leaving the atmosphere [*Hartmann et al.*, 1992]. Better understanding of the impact of ice clouds on the radiation budget and the hydrological cycle is paramount to improving climate models [e.g., *Stephens et al.*, 1990]. Climate models are the most important tools for understanding long-term atmospheric processes and simulating climate scenarios. However, fundamental ice-cloud properties such as ice-water path (IWP) are difficult to measure accurately, making them poorly constrained. This leads to large differences of globally averaged IWP between models [*Waliser et al.*, 2009]. Depending on their, e.g., microphysical properties, ice clouds may either cool or warm the atmosphere. The average radiative impact of all ice clouds is thought to be a net cooling effect, although semitransparent ice clouds, which may cover large areas, have a mostly warming effect on the atmosphere [*Khvorostyanov and Sassen*, 2002].

[3] In situ techniques used on aircraft and balloon campaigns provide the most detailed measurements of ice-water content (IWC), which, integrated by height, becomes IWP. However, the global coverage of such campaigns is sparse and limited. Ice-cloud retrievals based on satellite measurements are, in contrast, abundant. They provide macrophysical information on ice clouds such as their temporal variation and their spatial distribution, and are the most important source of validation of clouds in climate models. However, the uncertainties in satellite ice-cloud retrievals are still considerable, depending on the cloud characteristics and

[1]Department of Computer Science, Electrical and Space Engineering, Division of Space Technology, Luleå University of Technology, Kiruna, Sweden.
[2]Satellite Application Facility on Climate Monitoring, Deutscher Wetterdienst, Offenbach, Germany.
[3]I. M. Systems Group, Inc., National Oceanic and Atmospheric Administration/National Environment Satellite, Data and Information Service/Center for Satellite Applications and Research, Camp Springs, Maryland, USA.
[4]Department of Earth and Space Sciences, Chalmers University of Technology, Göteborg, Sweden.

Corresponding author: S. Eliasson, Department of Computer Science, Electrical and Space Engineering, Division of Space Technology, Luleå University of Technology, Kiruna, Sweden. (salomon.eliasson@gmail.com)

2169-897X/13/2012JD018381

the instrument sensitivities [e.g., *Zhao and Weng*, 2002; *Cooper et al*., 2003; *Austin et al*., 2009]).

[4] IWP [g m^{-2}] is one of the most important ice-cloud properties [*Buehler et al*., 2007, 2012, and references therein]. Previous studies have shown that the climate models show a very large spread in their reported cloud ice amounts [*Waliser et al*., 2009; *John and Soden*, 2006; *Wyant et al*., 2006; *Eliasson et al*., 2011]. However, the models are in good agreement with observations in terms of one of the most fundamental quantities, top of atmosphere net radiative flux [*Smith et al*., 1994]. This implies that the models may be adjusting poorly constrained parameters, such as clouds, to achieve the correct top of atmosphere radiative flux, thus leading to unrealistic cloud characteristics in the models [*Stephens et al*., 2002; *Li et al*., 2012]

[5] Atmospheric column integrated quantities, such as IWP, can be retrieved from radiances from passive down-looking sensors. Retrievals from passive sensors use two or more channels at microwave (MW), infra-red (IR), near-infra-red (NIR), or visible (VIS) wavelengths [e.g., *Heymsfield et al*., 2003]. Since June 2006, measurements from active instruments, a cloud profiling radar (CPR) on the CloudSat satellite, and a Cloud-Aerosol Light Detection and Ranging (Lidar) with Orthogonal Polarization (CALIOP) on the Cloud-Aeroso Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, have greatly increased our knowledge of ice clouds [*Heymsfield et al*., 2008]. These instruments measure detailed information on the vertical structure of clouds and can detect multilayered clouds. This can, for example, provide valuable information on how the radiation is distributed in the atmospheric column, because it depends strongly on the vertical structure of clouds [*L'Ecuyer et al*., 2008; *Mace and Benson*, 2008].

[6] In earlier studies [*Waliser et al*., 2009; *Eliasson et al*., 2011], a subset of climate models in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change were intercompared and compared with observational data sets. The models were in large disagreement on IWP (cloud ice). Not only were there large differences in magnitude, their spatial distribution were in poor agreement as well. The observational data sets were in good spatial agreement, but the differences in IWP magnitude was as large as the intramodel difference. This was also expected mainly because the observational data sets were based on measurements that (1) are made at different wavelengths, and thus sensitive to different parts of the particle size distribution (PSD) [*Comstock et al*., 2007; *Waliser et al*., 2009]; and (2) have different resolutions.

[7] Retrievals from passive instruments (nadir viewing) are much more limited in terms of information on the vertical structure of clouds, and most have a much coarser measurement resolution (footprint). Despite their limitations, records from passive instruments cover 10 (moderate resolution imaging spectroradiometer [MODIS]) to 30 years (AVHRR), meaning that only these records can be used for actual climate studies. Data from CloudSat and CALIPSO are limited to approximately 6 years (at the time of writing) and are therefore more suitable for process studies. Furthermore, because passive instruments also have large swath widths, passive instruments provide a much better spatial coverage than active instruments, which measure only at nadir. These advantages motivate the inclusion of passive instrument retrievals in this study. However, passive instruments provide only IWP; therefore, vertical information from active instruments goes unused.

[8] In practice, the longer the wavelength of the measurements, the deeper into a cloud one can measure, but this is at the cost of losing sensitivity to small particles, which are generally present at the top of the ice clouds [*Wu et. al*., 2009]. Therefore, ice-cloud retrievals based on, e.g., passive microwave measurements can be made in deep clouds, but generally not in shallow clouds, which are mostly made up of small ice particles. In other words, microwave retrievals are only sensitive to the precipitation-sized ice particles (>~0.25 mm) in the IWP total column [*Zhao and Weng*, 2002]. Also, although CloudSat CPR measurements can be used to retrieve information deep into most clouds, it is insensitive to small particles. CALIOP, in contrast, is even sensitive to very small ice particles. Because the satellites are always in close proximity to one another, CALIOP can be complementary to CPR. CALIOP by itself is limited for retrieving IWP because the measurements are attenuated for moderately thick clouds [e.g., *Delanoë and Hogan*, 2010]. In summary, satellite retrievals based on a single measurement technique cannot measure the whole depth of the ice cloud [*Wu et. al*., 2009], which would be desirable from a climate model perspective.

[9] The models participating in the upcoming Fifth Assessment Report (AR5) have undergone some improvements in terms of cloud amount and distribution, but there still remains large differences in IWP [*Jiang et al*., 2012]. A similar study by *Li et al*. [2012] reports that climate models do not perform particularly well despite generally simulating cloud ice better in Coupled Model Intercomparison Project Phase 5 (CMIP5) than they did in CMIP3 (included in AR4). This is an added incentive to further improve the understanding of IWP.

[10] This article can be seen as continuation of the work presented in *Eliasson et al*. [2011] and *Waliser et al*. [2009]. It builds on their findings but evaluates the discrepancies in the retrieval schemes rather than of averaged products. The latter, called *level 3 monthly mean* (L3), are representations that have already undergone a conversion from level 2 (L2) to L3, which, because different approaches and filters are applied, can contribute significantly to the differences between data sets. For instance, one important decision is the scan angle and solar zenith angle cutoffs one applies when creating the L3 data set, because there might be quality/bias dependencies of the IWP retrievals on these angles. The decision made in each of these approaches significantly alters the L3 product. Thus, this new study will once again survey satellite-observed IWP, but instead of comparing monthly mean values, we compared collocated measurements directly. By doing so, the same "cloud" is measured from different instruments, without the added complexity of different L2 to L3 conversions. This is not a validation study of any particular data set. The IWP retrievals assessed here, being from different sensors, are inherently different and difficult to compare in the rigorous manner a validation study requires [e.g., *Stein et al*., 2011; *Zhang et al*., 2009].

[11] As reference for the satellite IWP data sets, the radar/lidar (DARDAR) data set, which is based on CALIOP and CPR, was chosen [*Delanoë and Hogan*, 2010]. Data sets compared with this reference are from active sensors, two products from CloudSat's L2 radio detection and ranging

Table 1. List of Data Sets Used in This Study[a]

Short Name	Technique	Long Name	Satellite
DARDAR	Active: combined radar and lidar	DARDAR-cloud	CloudSat and CALIPSO
IORO	Active: radar	Ice only radar only ice water path	CloudSat
RO	Active: radar	Radar only ice water path	CloudSat
MODIS	Passive: SRBS	MODIS	Aqua
CMSAF	Passive: SRBS	CMSAF	NOAA-18
PATMOS-x	Passive: SRBS	PATMOS-x	NOAA-18
MSPPS	Passive: Microwave	MSPPS	NOAA-18
MiRS	Passive: Microwave	MiRS	NOAA-18
CMO	Passive: Microwave	CMO	NOAA-18

[a]Data sets are further referred by the "short names" given here.

(radar) only (RO) data set (called 2B-CWC-RO), in this article called RO and ice only radar only (IORO) [*Austin et al*., 2009], and from passive sensors; three data sets retrieved from solar reflectivities in the VIS and NIR spectral regions (MODIS [*King et al*., 2003], the advanced very high resolution radiometer [AVHRR] climate monitoring satellite applications facility [CMSAF] data set [*Roebeling et al*., 2006], and Pathfinder Atmospheres-Extended [PATMOS-x] [*Walther and Heidinger*, 2012]); and three data sets that are based on MW channels (microwave surface and precipitation products system [MSPPS] [*Ferraro et al*., 2005], microwave integrated retrieval system [MiRS] [*Boukabara et al*., 2011], and collocated microwave only [CMO] *Holl et al*. [2010]). An overview of all the data sets used in this study is given in Table 1.

[12] Because of the combination of radar and lidar, DARDAR has a very large IWP value range that covers the range of all IWP data sets in this study. Except for DARDAR, it is assumed that each data set contributes information for certain IWP value ranges. These IWP retrievals from different techniques are compared "out of the box." We do not attempt to make the best possible comparisons between data sets such as by limiting comparisons with optical depths that both data sets are sensitive to, as done in *Stein et al*. [2011] for DARDAR and MODIS. Instead, we acknowledge that all retrieval techniques have their merits and limitations, and compare them as they are, restricted only by the collocation rules described in section 3.

[13] In this article, IWP measurements are compared at the smallest possible temporal and spatial scales by stringently collocating measurements using a flexible collocations toolbox based on work first described in *Holl et al*. [2010] and later expanded in *John et al*. [2012]. The main aim is to quantify which measurement techniques work for which IWP value ranges and to separate the systematic and random errors of the collocated data sets. The data sets are described in more detail in section 2. The collocation and comparison methodologies, including the systematic and random errors, are described in section 3. In section 4, first the uncertainty of DARDAR IWP is investigated before identifying the valid IWP ranges of the collocated data sets and finding their systematic and random errors. Section 5 gives a recollection and states our main conclusions.

2. Description of Data Sets

[14] The data sets chosen for this study cover a wide range of measurement and retrieval techniques that are used to determine IWP. They are expected to report varying IWP magnitudes because they are sensitive to different ice particle sizes and shapes, and therefore are generally sensitive to different altitudes of the cloud. The data sets are based on instruments found on satellites flying in or near the afternoon train (A-train) [*Stephens et al*., 2002], so that a large number of collocated measurements can be found. For data sets that rely on visible reflectance measurements, only data from the daytime exist. Common to all data sets are the large uncertainties because of the assumptions on cloud microphysical properties that must be made to complete the cloud retrieval. Although these assumptions are best guesses mostly based on knowledge gained from in situ campaigns, many cloud properties, such as particle habit, are very variable [e.g., *Heymsfield and McFarquahar*, 2002]. Several studies [e.g., *Zhang et al*., 2009] have shown that different but widely accepted assumptions on the microphysical properties of ice particles lead to very large deviations in the retrieved IWP. In addition, the uncertainty because of the cloud particle phase, as the distinction between liquid and ice clouds, has a significant impact on the retrieved IWP as described later.

2.1. IWP From Active Instruments

[15] Three data sets in this comparison originate from active measurements and are all based on CPR measurements from the CloudSat satellite.

2.1.1. DARDAR

[16] The DARDAR data set is based on a combination of CPR, CALIOP, and MODIS measurements [*Delanoë and Hogan*, 2010]. By combining these techniques using the variational method described therein, DARDAR uses the very different particle size sensitivities of radar and lidar measurements. The retrieval is seamless and works as long as at least one of CPR or CALIOP detects a cloud. If the cloud is detected by both instruments, the cloud properties are retrieved using both measurements. *Delanoë and Hogan* [2010] showed that this combined retrieval approach is less sensitive to changes to assumed microphysical properties than retrievals based on CPR or CALIOP alone. The footprint size is the same as CPR because the CALIOP measurements are averaged horizontally in the CPR footprint. Further details on the retrieval technique are given in *Delanoë and Hogan* [2008], and a comparison of the DARDAR ice-cloud retrieval algorithms was performed in *Stein et al*. [2011]. The specific product used in this study is called *DARDAR-cloud* and is derived from CPR and CALIOP only.

[17] DARDAR was chosen as the reference data set because it currently provides the best estimate of the total column IWP.

It is expected to cover a larger IWP range than any other data set assessed here. For details on the uncertainty associated with DARDAR IWP, see section 4.1.

2.1.2. CPR (RO and IORO)

[18] The CloudSat satellite carries a CPR that measures at 94 GHz [*Stephens et al*., 2002]. Measurements are made at a vertical resolution of 250 m, and the measurement footprint size is about 1.4 × 3.5 km, with the long axis along the satellite flight path. CloudSat is less sensitive to small ice particles than lidar, but it is not saturated unless the clouds are very thick and there is heavy precipitation.

[19] One of the two CloudSat products assessed in the study allows IWC and liquid-water content (LWC) to coexist between -20 °C and 0 °C (but only the IWC part is included in the integration), and in this study, this data set is called RO (field name: RO_ice_water_path) [*Austin et al*., 2009]. RO was also the reference data set in *Eliasson et al*. [2011].

[20] CloudSat also provides a combined radar and visible optical depth (RVOD) product with presumably better constrained retrievals. However, *Protat et al*. [2010] state that RO and RVOD are statistically virtually identical, and therefore only the RO data set is used here. The uncertainty of IWC retrievals using simulated CPR was determined to be about 40% [*Austin et al*., 2009]. If this were the only uncertainties in the retrieval, 40% would be the upper limit of the uncertainty for a column integrated RO IWP product. For RO data set, there is a substantial additional uncertainty because of the assumed cloud phase. A linear combination of IWC and LWC is used between the temperature range 0°C to –20°C, where the fractional ice phase increases from 0 to 1 (liquid-water path [LWP]: 1 to 0). *Devasthale and Thomas* [2012] showed that other realistic ice to liquid phase relationships in this temperature range lead to very different IWP retrievals.

[21] The other CloudSat product used in this study does not attempt to separate IWP and LWP, and instead reports the whole column in the above temperature range as IWP; we have called this product *IORO* (field name: IO_RO_ice_water_path). As mentioned earlier, the IORO data set does not suffer from cloud-phase uncertainties caused by the above rigid approach, but by assuming the whole layer contains only ice, the IWP is overestimated. The a priori temperature information that both data sets rely on is model auxiliary data from European Centre for Medium-range Weather Forecasting. Further uncertainties for both data sets are inherited from the a priori input used in the retrieval [*Austin et al*., 2009].

2.2. IWP From Solar Reflectances

[22] The data sets in this section are derived from solar reflectance and share the same retrieval technique. This method is called *solar reflectance bispectral* (SRBS) and is described in *Nakajima and King* [1990]. The method uses passive measurements of reflected solar radiation to retrieve visible cloud optical depth (τ_v) and $\bar{r}_e$ simultaneously, where $\bar{r}_e$ is the mean effective radius. The solar reflectance of the nearly nonabsorbing wavelengths is used to retrieve τ_v, and moderately absorbing solar reflectance measurements are used to retrieve $\bar{r}_e$. This is done in conjunction with a lookup table of simulated reflectances.

[23] In SRBS retrievals, assumptions have to be made about the horizontal and vertical structure of the cloud, i.e., constant IWC and constant particle effective radius (r_e) throughout the cloud [*Stein et al*., 2011, and references therein]. These assumptions lead to uncertainties, because inhomogeneities in the vertical structure of clouds have a strong influence on the retrieved cloud properties [*Zhang et al*., 2010]. Also, the retrieved $\bar{r}_e$ is biased toward the top of the thick clouds, because only the top four or five optical depths contribute to the reflectance in the moderately absorbing channel [*McFarquhar and Heymsfield*, 1998]. Because the smallest ice particles are generally at the top of the cloud, this may create systematic errors by underestimating $\bar{r}_e$ for thick clouds.

[24] Nonetheless, from the retrieved τ_v and $\bar{r}_e$, IWP can be derived using the following relationship:

$$\mathrm{IWP} = \frac{4\tau_v \bar{r}_e \rho_{ice}}{3Q_e} \tag{1}$$

where ρ_{ice} is the density of ice, and Q_e is the average extinction efficiency for ice at a wavelength of 0.66 μm ($Q_e \approx 2$) [e.g., *Meyer et al*., 2006]. We have used ρ_{ice} = 930 kg m^{-3}, because this is the value used in the MODIS retrievals [*King et al*., 2006].

2.2.1. MODIS

[25] The MODIS cloud retrieval products are based on daytime measurements from 36 channels in the VIS, NIR, and IR spectral ranges from the MODIS instrument on board two polar-orbiting satellites, Terra and Aqua. However, only a subset of four channels is used to ultimately retrieve IWP, namely, either channel 1 (0.645 μm), 2 (0.858 μm), or 5 (1.24 μm), depending on the underlying surface, and channel 7 (2.13 μm) [*King et al*., 1997]. The lookup table used in the MODIS retrieval algorithm is based on the ice particle model by *Baum et al*. [2005] (Baum05).

[26] The MODIS L3 monthly mean IWP data set (called MYD08_M3) was assessed in *Eliasson et al*. [2011]. This article assesses the L2 data set (called MYD06_L2) collection 005 [*King et al*., 2003], which is from only the Aqua satellite. The IWP is extracted using the fields *cloud_water_path, Cloud_Phase_Optical_Properties*, and *Quality_Assurance_1km*. The resulting IWP data set is a 1 × 1 km pixel (footprint) product and, therefore, is of a similar size as the DARDAR retrieved footprint. MODIS data used in this study are described in documentation available through the MODIS Web site (http://modis-atmos.gsfc.nasa.gov/).

2.2.2. AVHRR PATMOS-x

[27] PATMOS-x IWP is derived from the PATMOS Daytime Cloud Optical Microphysical Properties (DCOMP), based on passive measurements in the VIS and NIR spectral ranges from AVHRR global area coverage (GAC) data [*Walther and Heidinger*, 2012]. The lookup table is based on Baum05. GAC data are a reduced resolution data set based on AVHRR. Four adjacent footprints in every scan are averaged together; then the next three scan lines are skipped. The AVHRR instrument is installed on the polar-orbiting operational environmental satellites from National Oceanic and Atmospheric Administration (NOAA). In this study, the L2b cloud parameter product (version 5) is further sampled onto a 0.1° x 0.1° grid. The gridded data points are treated as "pseudo footprints" that are largest at the equator

(roughly 10×10 km), shrinking laterally toward the poles. The product does not contain an IWP product, but contains τ_v and $\bar{r}_e$. Therefore, IWP is extracted using equation 1 from these quantities from cloud types considered ice phase (called opaque_ice, cirrus, overlapping, and overshooting in this data set, but slightly different terms are used in *Walther and Heidinger* [2012].

[28] PATMOS-x has only two solar reflectance channels for the τ_v and particle effective radius (r_e) retrievals: channels 1 (0.6 μm) and 3b (3.7 μm). The PATMOS-x L3 IWP data set (version 4) was assessed in *Eliasson et al*. [2011]. In this study, the L2b data set is based on measurements from NOAA-18, which flies close to the A-train; hence, the PATMOS-x measurements are often collocated with DARDAR. The PATMOS-x data used in this study are described and made available online at http://cimss.ssec.wisc.edu/patmosx/.

2.2.3. AVHRR CMSAF

[29] CMSAF IWP retrieval data are based on the cloud physical properties (CPP) algorithm developed at Koninklijk Nederlands Meteorologisch Instituut (KNMI) [*Roebeling et al*., 2006], which is used to retrieve cloud thermodynamic phase, cloud optical thickness, cloud particle effective radius, and LWP/IWP from AVHRR GAC data. Therefore, the CMSAF cloud products have a pseudo footprint of about 4×1 km. The CPP scheme, based on the SRBS method, uses lookup tables of water and ice-cloud reflectance simulated by the Doubling Adding KNMI radiative transfer model [*Stammes*, 2001]. Although MODIS and PATMOS-x both use the Baum05 ice particle model, for CMSAF IWP, ice particles as given in *Hess et al*. [1998] were used. Cloud properties are retrieved by iteratively matching observed and simulated reflectance. The IWP is derived using equation 1 for all cloudy pixels, which were identified using the Polar Platform System cloud processing package developed by Swedish Meteorological and Hydrological Institute [*Dybbroe et al*., 2005a, 2005b], and which were diagnosed to contain ice clouds (internally done in CPP). Also, notably, the AVHRR reflectances are recalibrated following *Heidinger et al*. [2010]. As PATMOS-x, and CMSAF IWP retrievals are based on the 3.7 μm channel of AVHRR, the availability of different NIR channels might significantly affect the IWP retrieval, because of different penetration depth of, e.g., 3.7 and 1.6 μm. This might explain some of the differences seen between MODIS and the AVHRR data sets later in this article.

2.3. IWP From Passive Microwave

[30] Retrieving IWP from passive microwave sensors is analogous to retrieving the IWP of the particles with a radius between approximately 250 and 1500 μm. These are precipitation-sized particles, which should be considered when comparing with other data sets that have sensitivity to small particles. In general, comparisons of microwave data sets where DARDAR IWP is relatively low will not be made, because such "clouds" are likely made up of small particles that are beyond the sensitivity of passive microwave measurements. The three data sets used in this study use measurements from the microwave humidity sounder (MHS) instrument, which has a footprint diameter of around 15 km at nadir on the ground.

2.3.1. MSPPS

[31] MSPPS provides a data set of IWP derived from the ratio of the amount of radiation at 89 and 150 GHz that is scattered out of the line of sight by large ice particles [*Ferraro et al*., 2005]. These two frequencies are measured by MHS channels 1 and 2. The particle-induced scattering intensifies with increasing frequency and is detected as a depression in brightness temperature [*Vivekanandan et al*., 1991]. Assumptions are made on the PSD and bulk volume density, and accurate knowledge on the surface temperature and its emissivity must be available [*Zhao and Weng*, 2002]. As described therein, errors in assumptions can lead to large errors in the IWP retrieval.

2.3.2. MiRS

[32] MiRS is a one-dimensional variational satellite data assimilation retrieval system that uses microwave observations for the retrievals of several atmospheric and surface geophysical parameters simultaneously, among them IWP [*Boukabara et al*., 2011]. Like MSPPS, MiRS is based on MHS sensor observations, but the MiRS retrievals are additionally constrained by measurements from the Advanced Microwave Sounding Unit A (AMSU-A). Because AMSU-A has a nadir footprint of 48 km, the MHS measurements made inside the larger footprint are averaged, so the footprint of MiRS is that of AMSU-A. MiRS is described in detail in *Boukabara et al*. [2011]. The MiRS version used in this study provides an IWP product called *graupel water path*, which indicates by name alone that only precipitating sized particles are retrieved. In MiRS, the IWP product is used as a predictor for the estimation of rainfall rate intensities in mm h^{-1}. As described in *Iturbide-Sanchez et al*. [2011], the validation/assessment of the MiRS rainfall rate is an indirect method to assess the quality of the retrieved MiRS hydrometeors, including the IWP.

2.3.3. CMO

[33] CMO is a data set based on a technique first introduced in *Holl et al*. [2010]. It is currently under development and will be described in detail in a paper that is in preparation. High-frequency measurements from MHS channels 3 to 5 located at 183(1) GHz, 183(3) GHz, and 190 GHz, which are traditionally used for water vapor retrievals, form the basis of this data set. This sets it apart from the other microwave-based data sets that use lower frequencies. Therefore, the CMO data set should be able to detect smaller ice particles and retrieve smaller IWP, a hypothesis that can be tested by including it in this study.

[34] Artificial neural networks are used for the retrieval of IWP rather than the depression of brightness temperature directly. The training database is obtained by collocating CloudSat with NOAA-18 MHS, using the collocation toolbox described in *Holl et al*. [2010] and later in section 3. *Holl et al*. [2010] showed that the MHS channels 3 to 5 are not sensitive to clouds with RO IWP less than 100 g m^{-1}, and the data set is planned to include radiances from IR measurements to potentially increase sensitivity to lower IWP values. However, this study uses CMO version 0.4, which does not yet include IR radiances.

3. Methodology

[35] To compare simultaneous measurements of the same cloud situation using several instruments, the measurements

need to be stringently collocated in time and space. To find such collocated measurements, we have used the collocation toolbox, a highly flexible toolkit that allows for easy collocation between different data sets, first presented in *Holl et al*. [2010]. The software, some precalculated collocated data sets, and detailed documentation can be accessed online at http://www.sat.ltu.se/docs/data/collocations. The toolbox has been used in several studies and continues to be developed further (V. O. John, G. Holl, N. Atkinson, and S. A. Buehler, Monitoring scan asymmetry of microwave humidity sounding channels using simultaneous all angle collocations [SAACs], *Journal of Geophysical Research*, in press 2013).

[36] One very important consideration to take into account when collocating measurements that use different techniques is their different horizontal resolutions given by their "footprints." Figure 1 illustrates footprint sizes for the instruments that the data sets in this study are based on. Technically, to collocate data sets, usually the one with the smallest footprints (secondary) is collocated to the other (primary). All the small-footprint measurements from the secondary data set that are close enough to the center of a larger footprint from the primary data set are found and are considered collocated. The user defines what "close enough" means in space and time in the presets (see later). Each collocated data set contains information such as distances, number of secondary measurements, and the statistics of each subset of small-footprint measurements such as standard deviation, mean, number of elements, or other user-defined functions such as "cloud amount."

[37] Comparing measurements with large footprints with measurements with small footprints introduces a sampling problem. For instance, a single CPR measurement (and DARDAR retrieval) covers around 0.65% of the area of an MHS footprint, and even the maximum possible number of DARDAR retrievals collocated with one MHS measurement cover less than 10% of its footprint [*Holl et al*., 2010]. In this article, large-footprint measurements are considered collocated with a set of small-footprint measurements only if the following two conditions are met, to mitigate this sampling problem: First, the number of smaller footprints inside the larger footprint must be at least 70% of the maximum possible collocations of small footprints to one large footprint within the set collocation criteria. Second, because partly cloudy atmospheric conditions can lead to large inhomogeneities in the measured IWP, only collocations where all measurements from the small-footprint data set are ice cloud filled (i.e., not cloud free and not liquid cloud) are assessed. The collocated data sets are compared using the retrieved IWP from the large-footprint data set and mean retrieved IWP from the small-footprint data set. For collocated data sets with similar-sized footprints, the comparison is done directly; that is, only the closest collocated measurements are considered.

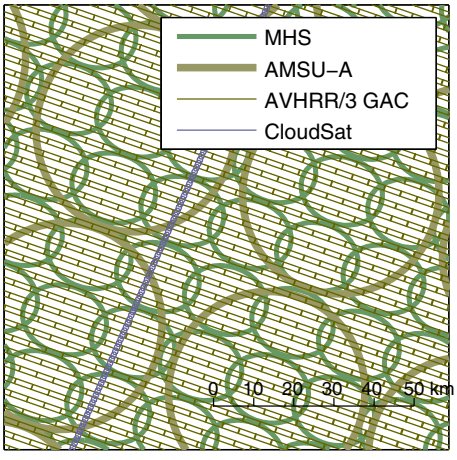

Figure 1. Illustration of footprint sizes for a selection of instruments.

[38] The data sets assessed in this study are from sun-synchronous polar-orbiting satellites on a similar orbit, being either in the A-train or close to it. As a consequence, all collocation pairs contain collocations across all latitudes (global collocations). The collocations where both data sets are from satellite platforms in the A-train (MODIS from Aqua, RO and IORO from CloudSat, and DARDAR from CloudSat and CALIPSO) contain only measurements made at nadir. However, collocations including data sets from NOAA-18 (PATMOS-x, CMSAF, MiRS, MSPPS, and CMO) may include measurements made off-nadir. Deviations may occur because of comparing off-nadir with nadir measurements, but should be limited assuming that all data sets take angular dependencies into account. In addition, all problematic measurements are assumed to be flagged, and hence are not included. All collocations in this study are limited to the Tropics defined as within $\pm 30^{\circ}$ latitude, where the atmospheric conditions are assumed to be more homogeneous than if higher latitudes were included. For practical reasons such as data storage, the study is limited to 1 year. The period 1 January 2007 to 31 December 2007 was chosen because it is the first full year of CloudSat data.

[39] In this article, measurements are considered collocated if their footprints are overlapping each other. The specific limitations applied to each collocated data set depend on the size of the larger footprint. Each data set's measurement resolution is described in section 2. For simplicity, footprints are assumed circular, and the random errors introduced by this assumption are assumed to be minor by comparison with the other sources of random error, such as a rapidly changing atmospheric state, or because of the varying distances between individual collocated measurements. The maximum distance that the center of a small footprint is allowed to be away from the center of the large footprint is the radius of that footprint. For example, for data sets based on MHS measurements, the footprint size is approximately 15 km in diameter at nadir. Only DARDAR footprints that are a maximum distance of 7.5 km away from the center of the large footprint are collocated.

[40] The temporal limit chosen is the time for advection to move by the radius of the footprint, so that at most half the footprint's condition is changing. In this study, it is estimated that the air mass will be replaced at roughly a rate of 1 km min^{-1} (16.7 m s^{-1}), which is a fairly conservative estimate for the Tropics [*Koh et al*., 2011]. Changes caused by other atmospheric processes such as convection, which, in most cases, will take much more than 10 minutes to change conditions completely, are also taken into account. Therefore, the largest permitted temporal difference between data sets is either the footprint radius [km] in minutes or 10 min, whichever value is smallest. The exact limitations used for each collocated data set are given in Table 2.

Table 2. Collocation Rules (from left to right): Data Set, Maximum Allowed Distance, Maximum Allowed Time Difference, Minimum Required Number of Collocated DARDAR to Each (larger) Footprint, Maximum Number of Small Footprints per Large Footprints, and Some Additional Comments[a]

Data Set	Distance [km]	Time [min]	N Required	N Maximum	Notes
RO	—	—	—	—	Already collocated
IORO	—	—	—	—	Already collocated
MODIS	2	2	—	—	Pick closest only
CMSAF	2.5	5	4	5	CA = 100%
PATMOS-x	4	8	6	8	CA = 100%
MSPPS	7.5	7.5	10	14	CA = 100%
MiRS	25	10	33	46	CA = 100%
CMO	7.5	7.5	10	14	CA = 100%

[a]CA refers to cloud amount, and this means, in practice, that only collocations that have at least "CA" number of cloudy DARDAR footprints are considered.

3.1. Systematic and Random Errors of Collocated IWP

[41] The errors are split into systematic and random errors. What constitutes a random error or a systematic one depends on context. In this study, the systematic error is defined as the mean difference between the measurements (e.g., the bias, if one is taken as a reference) and the random error as the residual after subtracting the mean (e.g., the variance). Both may be a function of IWP or of other quantities. As an example of an expected systematic error, a product based exclusively on radar should retrieve systematically less IWP than a product based on combined radar and lidar, if all other factors (such as microphysical assumptions) are the same. Random errors between collocated data sets originate from a number of different sources. These include, but are not limited to, the representation error (poor collocation in space and time) and the random errors within each data set (such as noise). If the random errors of two data sets errors are Gaussian and uncorrelated, variances (square of the standard deviation) add up to the total variance for the collocated data set,

$$\sigma^2(Y-X) = \sigma^2(X) + \sigma^2(Y) + \sigma_r^2, \tag{2}$$

where Y and X are collocated measurements with uncorrelated random errors, and σ_r is the representation error caused by imperfect collocation. In this study, Y is $log_{10}(\mathrm{IWP_{dataset}})$ and X is $log_{10}(\mathrm{IWP_{DARDAR}})$. Thus, the observed random error in the comparison puts an upper bound on the random error of the individual data sets. All active data sets and the CMO data set are based on the same CloudSat data, so that it would be wrong to assume the data sets have noncorrelated random errors. Therefore, approximating random errors for these collocated data sets (RO, IORO, CO) -DARDAR cannot be done in the above manner but can be considered applicable for the comparison of the other passive instruments with DARDAR. Further investigation on whether the IWP distributions are, in fact, Gaussian in logarithmic space (i.e., lognormal) is made in section 4.2.3.

4. Results and Discussion

4.1. DARDAR Uncertainties in IWP

[42] DARDAR retrievals use the optimal estimation method; hence, the errors are retrieved for each IWC value [*Delanoë and Hogan*, 2008]. 1-sigma errors of lognormal quantities mean that the errors are also given in log space; therefore, the uncertainty is a fractional uncertainty (where, e.g., a 50% uncertainty for the value 100 g m^{-2} means a value range between 100/1.5 and 100*1.5). However, it is not easy to assign an error to the column integrated IWP using only the fractional errors of each IWC, because the vertical autocorrelation of IWC errors is unknown. The error on the logarithm of IWP can be estimated using the provided fractional error (1-sigma random error in the natural logarithm of IWC called "ln_iwc_error") and assuming that the IWC uncertainties are correlated throughout the column. The ±1-sigma values of IWC can be used,

$$\mathrm{IWC}^{+} = \mathrm{IWC}_0 e^{+\ln_i \mathrm{wc_error}}$$
$$\mathrm{IWC}^{-} = \mathrm{IWC}_0 e^{-\ln_i \mathrm{wc_error}},$$

where IWC_0 is the retrieved IWC, to get the corresponding column integrated quantities IWP^{+} and IWP^{-}. From IWP^{+}, IWP^{-}, and IWP_0 (integrated from IWC_0), we can find the fractional errors, $ln(\mathrm{IWP}^{+}/\mathrm{IWP}_0)$ and $ln(\mathrm{IWP}_0/\mathrm{IWP}^{-})$, and their average,

$$\sigma_{\mathrm{IWP}} = \frac{ln\left(\frac{\mathrm{IWP}^{+}}{\mathrm{IWP}^{-}}\right)}{2}.$$

[43] This error σ_{IWP} is assigned to every IWP_0 value such that the $\pm 1 - \sigma_{\mathrm{IWP}}$ limits (in log space) can be estimated as

$$\mathrm{IWP}_0 \times e^{\pm\sigma_{\mathrm{IWP}}}.$$

[44] Hence, a σ_{IWP} of, e.g., 0.4, constitutes a fractional error of 50% ($e^{0.4} = 1.5$). Figure 2 shows the IWP uncertainty given by σ_{IWP} expressed as fractional error ($100 \times [e^{\sigma_{IWP}} - 1]$ [%]). In a similar way as *Austin et al*. [2009] assigned an uncertainty of 40% to IWC retrievals from CloudSat's RO product, the uncertainty of DARDAR IWP is assigned to approximately 20 to 50%, with the uncertainty generally increasing with decreasing IWP.

4.2. Collocated Data Set Comparisons

[45] In this section, collocated data sets are compared in log space in the IWP range from around 0.6 to 16,000 g m^{-2} ($10^{-0.2} - 10^{4.2}$). The measurements are binned into 50×50 equally spaced $\log_{10}$(IWP) bins using DARDAR as the reference.

[46] The comparison is made by first identifying the IWP value ranges where both the compared data sets and DARDAR are sensitive to clouds, and second by comparing similar data sets with one another within the determined valid

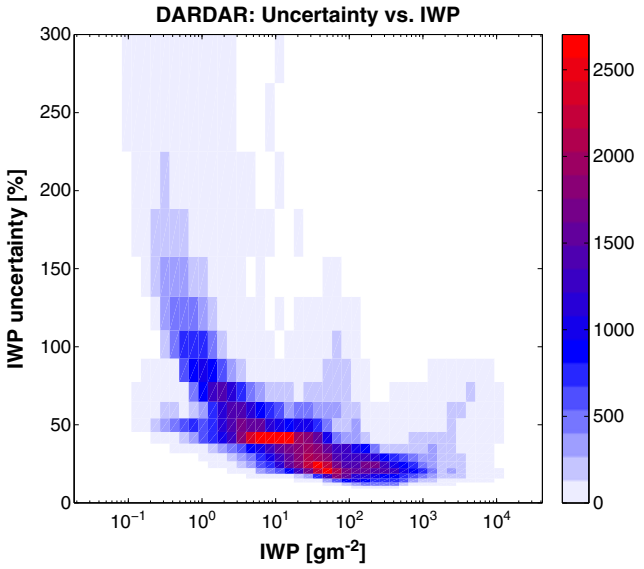

Figure 2. The fractional uncertainty of IWP (%) as a function of IWP for DARDAR.

ranges from which statements about their systematic and random errors are made. The median and the 16th/84th percentiles are used to describe the distribution of IWP.

[47] The median is used instead of log-mean because the IWP distributions are not quite lognormal, albeit nearly (see section 4.2.3). The choice of 16th/84th percentiles ("pseudo 1-sigma") corresponds to ±1 standard deviation in log space if the IWP distribution would be lognormal. Furthermore, because the comparisons are made in log space, all measurements where the compared data sets report IWP = 0 are removed. Cases where DARDAR has cloud-free measurements have already been discarded to minimize the number of partly cloudy footprints (see section 3). The comparisons of collocated data sets reported in the following sections are grouped in active data sets, SRBS data sets, and MW data sets.

4.2.1. IWP Valid Ranges

[48] By comparing collocated measurements directly against each other, one can approximately read out the actual overlap in IWP sensitivities for the different techniques against the reference data set, DARDAR. Figures 4, 6, and 8 (described in sections 4.3, 4.4, and 4.5) show the two-dimensional histogram distributions of IWP for each collocated data set. The median and spread (16th/84th percentiles) of the compared data sets in each DARDAR bin are shown in gray, and the median and spread of DARDAR IWP inside the bins of the compared data set are shown in black. To clarify, in each of the bins (along the *x* axis) there is a distribution of IWP values from the compared data set. For example, for all collocated measurements with DARDAR IWP in the bin between 50 and 70 g m^{-2}, the compared data set may have values ranging from 1 to 1000 g m^{-2}, but with most values centered in a fairly lognormal way around 40 g m^{-2}. This distribution can be viewed as a probability distribution function (PDF) of likely values that this observational data set may report when DARDAR values are between 50 and 70 g m^{-2}. By the same reasoning, the bins of the *y* axis also have a PDF of DARDAR IWP values.

[49] The color scale common to Figures 4, 6, and 8 shows the number of collocations per bin normalized by the largest bin value, and the total number of collocations is reported in each figure. For each group of comparisons (active, SRBS, and MW), the number of collocations that matches the data set with the fewest collocations are randomly selected to avoid data sets with fewer collocations appearing noisier.

[50] Because the data sets are based on instruments with different sensitivities, it only makes sense to discuss comparisons in IWP ranges where both the compared data set and DARDAR are sensitive to clouds. In IWP value ranges where either the compared data set or DARDAR are not sensitive to or cannot retrieve the cloud, the data sets are expected to be uncorrelated. Using Figure 6 (described in section 4.4), we show that the median IWP of the observed three data sets and the median of DARDAR values (shown as thick gray and black lines, respectively) appear to rapidly diverge for IWP values less than approximately 30 g m^{-2} and diverge slightly again for values greater than approximately 2000 g m^{-2}. On one extreme, if the median lines are near perpendicular to one another, the measurements are uncorrelated within that IWP range, whereas if the median lines are parallel to each other, the measurements are correlated. For very low IWP values (<20 g m^{-2}), all data set comparisons are uncorrelated. However, the median lines always converge at some point for increasing IWP values. Therefore, for each data set, the decision on the valid IWP ranges is based on where the median lines converge to being very close to each other and end at high IWP values if the median lines strongly diverge from each other again. The results of this IWP dynamic range analysis are summarized in Table 3.

4.2.2. Systematic and Random Errors

[51] For easier comparison among the various data sets, the median and percentiles of the log ratio of IWP and DARDAR IWP are displayed together in Figures 5, 7, and 9 (described in sections 4.3, 4.4, and 4.5) as a function of DARDAR IWP. Thick lines show the median, and the same colored thin lines depict the random spread using the 16th/84th percentiles (see section 3.1 for description on random errors). Solid lines are drawn in the IWP ranges deemed valid to compare the two data sets (Table 3), and the lines become dashed outside this range. Horizontal thin gray lines showing factor of 2 steps in systematic error and a thick line showing equality are added for clarity.

[52] The 16th/84th percentiles shown in Figures 5, 7, and 9 represent the random error in each comparison. As mentioned earlier, the chosen percentiles match the standard deviation if the distribution is Gaussian.

4.2.3. Are IWP Distributions Lognormal?

[53] In fact, the log-ratio distributions are often found to be close to Gaussian, as illustrated for the PATMOS-x versus DARDAR comparison in Figure 3. It shows the median and percentiles, as well as the mean and standard deviation in terms of log ratio between PATMOS-x and DARDAR. As Figure 3 shows, the log distribution of IWP is nearly Gaussian across the whole IWP range. This applies to most data sets in this study (data not shown). The fact that the log distributions are nearly Gaussian justifies using the

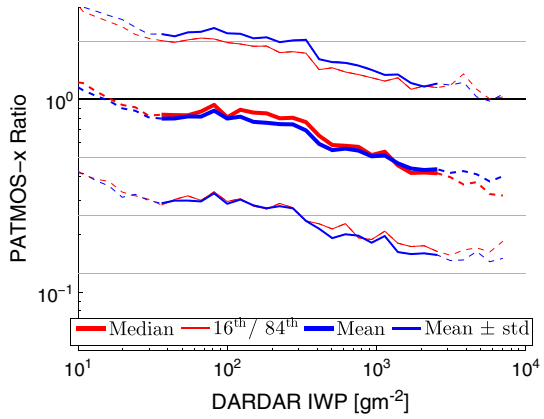

Figure 3. The median and 16th/84th percentiles shown together with the mean and standard deviations of $log_{10}(\frac{IWP_{PATMOS-x}}{IWP_{DARDAR}})$ using the PATMOS-x versus DARDAR comparison. Where the mean and median lines are close to each other and the standard deviation and percentiles are close to each other, the IWP distribution is close to lognormal.

standard deviation (or percentiles) of the log ratio as a random error estimate (see section 3.1). Nonetheless, because the distributions are not exactly Gaussian, the median/percentile approach is chosen for better robustness. It means that the errors generally scale with the IWP value, and relative errors are more appropriate than absolute errors to characterize IWP.

4.3. Active Sensors

[54] Using the median lines in the two-dimensional histogram of IWP in Figure 4, DARDAR is determined comparable with the active data sets in the IWP range of approximately 10 to 9000 g m^{-1}, i.e., a large part of the range retrieved by DARDAR. However, the expected close level of agreement in terms of IWP of RO and IORO to DARDAR is due to large ice particles dominating the IWP (column integrated IWC). In terms of IWC, especially where there is low IWC, the CloudSat data sets diverge from DARDAR [*Stein et al.*, 2011].

[55] When comparing active data sets, one needs to note that DARDAR is based on the same radar measurements as the CloudSat IWP products, so the random errors of these data sets are likely correlated; therefore, nothing can be said about the random errors of the individual data sets. However, because DARDAR uses lidar measurements in the retrieval, and the CloudSat data sets assume different particle phase assumptions, they are appreciably different from each other, as shown in Figure 5. This figure also shows that for thin clouds in the range of 30 to 40 g m^{-2}, RO and IORO both report more than a factor of 2 less IWP than DARDAR. This demonstrates the strong impact of lidar measurements on DARDAR in this range. This was corroborated using the instrument flags of DARDAR. For instance, at 10 g m^{-2}, on average, the fraction of IWC values in each vertical profile retrieved solely from lidar, radar, or from a mixture of both measurements is approximately 55%, 20%, and 25%, respectively.

[56] The IORO and RO data sets converge with DARDAR for clouds around 500 g m^{-2} as the influence of lidar measurements becomes less dominating. The IORO data set reports more IWP than DARDAR between approximately 100 and 800 g m^{-2}, which likely is due to an overestimation of IWP because all cloud water content where the temperature is colder than 0°C is assumed to be ice phase (see section 2). For values larger than 800 g m^{-2}, RO and IORO report lower IWP than DARDAR, but RO decreases more than IORO.

[57] DARDAR detects the cloud phase directly using the intensive backscatter by supercooled water compared with ice in the lidar signal. When the lidar signal is attenuated before the liquid phase level, the radar echo is assumed to be dominated by ice. Thus, only the IORO product was

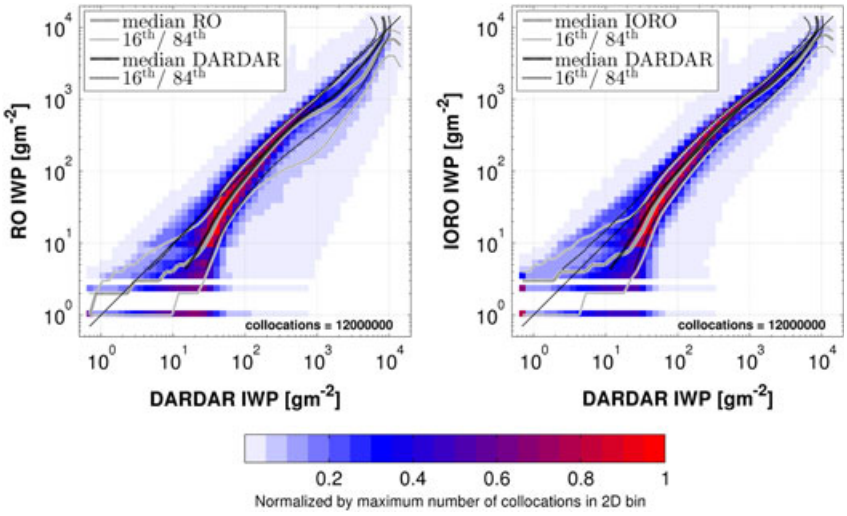

Figure 4. Tropical IWP comparison 2007 for active techniques: CloudSat RO IWP (left) and CloudSat IORO (right) are compared against DARDAR IWP. The measurements are collocated because they are based on the same measurements from CPR. The color scale shows the normalized number of collocations per bin against the maximum bin value. Median lines and percentiles are described in the text.

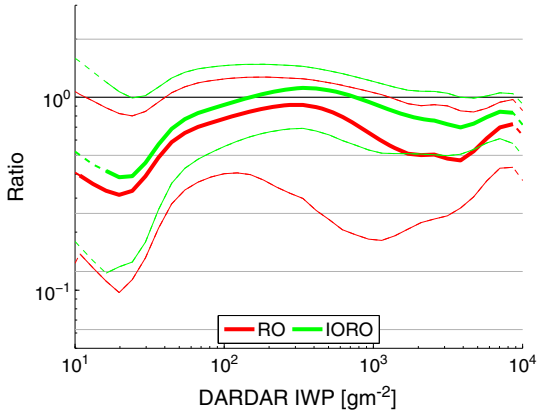

Figure 5. WP from active data sets. The median and "pseudo 1-sigma" of the ratios of RO versus DARDAR and IORO versus DARDAR are shown together. Thick lines indicate medians; thin lines indicate "pseudo 1-sigma"; dashed lines indicate IWP ranges where the data sets are uncorrelated.

considered in the *Stein et al*. [2011] study because it is a closer comparison with DARDAR.

4.4. Passive VIS/NIR Sensors

[58] When comparing passive measurements of IWP with measurements from active instruments, one must bear in mind the limitations of IWP retrievals from passive sensors. In addition to the uncertainties introduced in assuming homogeneous clouds (see section 2.2), only measurements determined to be ice phase are used to retrieve ice-cloud properties, and the whole column is then assumed to be ice. These problems are known to cause large uncertainties in SRBS retrievals [e.g., *Stein et al*., 2011]. By comparison, as mentioned earlier, DARDAR profiles can contain both liquid and ice particles, and only the vertical bins that contain ice (IWC) are considered in the integration.

[59] Three data sets using reflected solar measurements to derive IWP are shown in Figure 6. Using the median lines to find the IWP ranges where it may be valid to make comparisons with DARDAR, the SRBS data sets are determined to cover a large range of IWP values. In general, the retrievals at low IWP values (approximately $\leq$30 g m^{-2}) are not correlated at all between the compared data sets and DARDAR, meaning that either neither is sensitive to such thin clouds or only one of them is (probably DARDAR). There appears to be no clear upper limit to the SRBS data sets, although the data sets are slightly less correlated above approximately 2000 g m^{-2}. At greater than these values, the systematic errors between the SRBS data sets and DARDAR rapidly increase. Because all three data sets are based on similar instruments, the retrievals also have approximately the same IWP ranges. However, there are still some notable differences between these data sets, which are best visualized in Figure 7.

[60] Figure 7 shows the median and percentiles of the ratio between SRBS data sets and DARDAR IWP as a function of DARDAR IWP. Curiously, within the valid ranges, MODIS reports more IWP than DARDAR from approximately 20 to 100 g m^{-2}, whereas CMSAF and PATMOS-x do not. In the vicinity of 20 g m^{-2}, MODIS has a factor of 2 higher IWP values, and at the end of the range around 5500 g m^{-2},

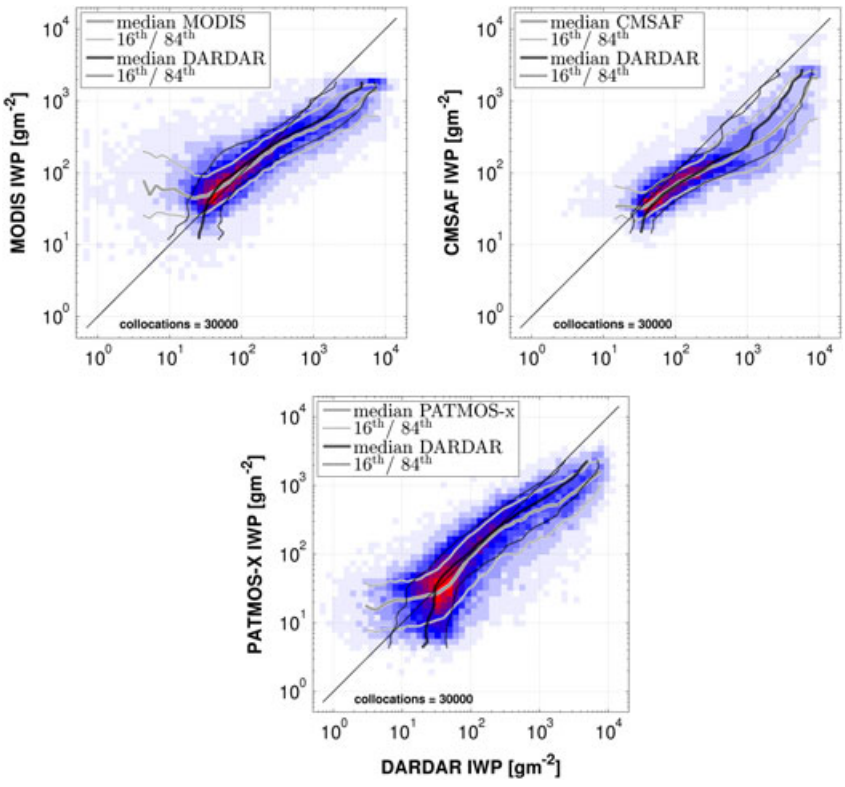

Figure 6. Tropical IWP comparison 2007 for SRBS techniques: two-dimensional histograms as in Figure 4. MODIS IWP (top left), CMSAF IWP (top right), and PATMOS-X IWP (bottom).

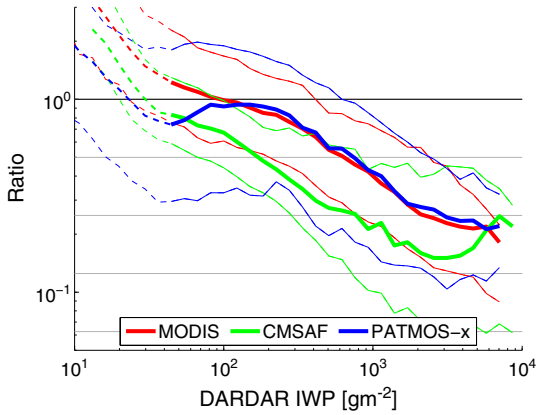

Figure 7. IWP from passive SRBS data sets. The median and "pseudo 1-sigma" of the ratios of MODIS versus DARDAR, PATMOS-x versus DARDAR, and CMSAF versus DARDAR are shown together. Thick lines indicate medians; think lines indicate "pseudo 1-sigma"; dashed lines indicate IWP ranges where the data sets are uncorrelated.

MODIS reports IWP that is lower by around a factor of 4. PATMOS-x and CMSAF have a systematic negative bias for all DARDAR values, but from approximately 100 g m^{-2} the bias of PATMOS-x is nearly the same as the bias of MODIS, whereas CMSAF has nearly a factor of 2 more systematic bias than MODIS and PATMOS-x for the whole IWP range.

[61] The combined random uncertainty (see section 3.1) of MODIS-DARDAR, depicted as thin red lines in Figure 7, is very constant and nearly parallel with the median throughout the valid range and ranges from a factor of 1.5 to 2. PATMOS-x-DARDAR has a somewhat larger but also quite constant random uncertainty (approximately a factor of 2–3) for the whole IWP range. CMSAF has a low random uncertainty below approximately 200 g m^{-2} (approximately a factor of 1.3), but then steadily increases to around a factor of 3 to 4 for the highest IWP values.

[62] Taking MODIS-DARDAR as an example, *Cooper et al*. [2003, 2007] showed that the uncertainties in MODIS retrievals of $\bar{r}_e$, used to derive IWP, are of the order of 30 to 40%, and DARDAR's random uncertainty was shown to have an overall random uncertainty of around 20 to 50% (see section 4.1). However, such general uncertainty estimates per data set are inadequate for IWP retrievals because they are clearly not valid for all IWP ranges. Using the approach described in this article (see section 3.1), the combined random error for the entire valid IWP range can be estimated as a function of IWP.

[63] Assuming that the random errors between these two data sets are uncorrelated, the maximum random uncertainty of either data set is concluded to be smaller than a factor of 2 (100%). Because the combined error of MODIS-DARDAR is not too large, covering a wide range of IWP values, this demonstrates the likely strength of both data sets. This also helps to justify our choice of DARDAR as the reference data set, because if it was a bad reference, the combined uncertainties between DARDAR and any other data set would always be large. If both were assumed to have the same uncertainty, using equation 2, their uncertainty would be smaller than 1.63. An uncertainty of 1.63 on the ratio corresponds to a relative error of 63%, which is not too far from the earlier

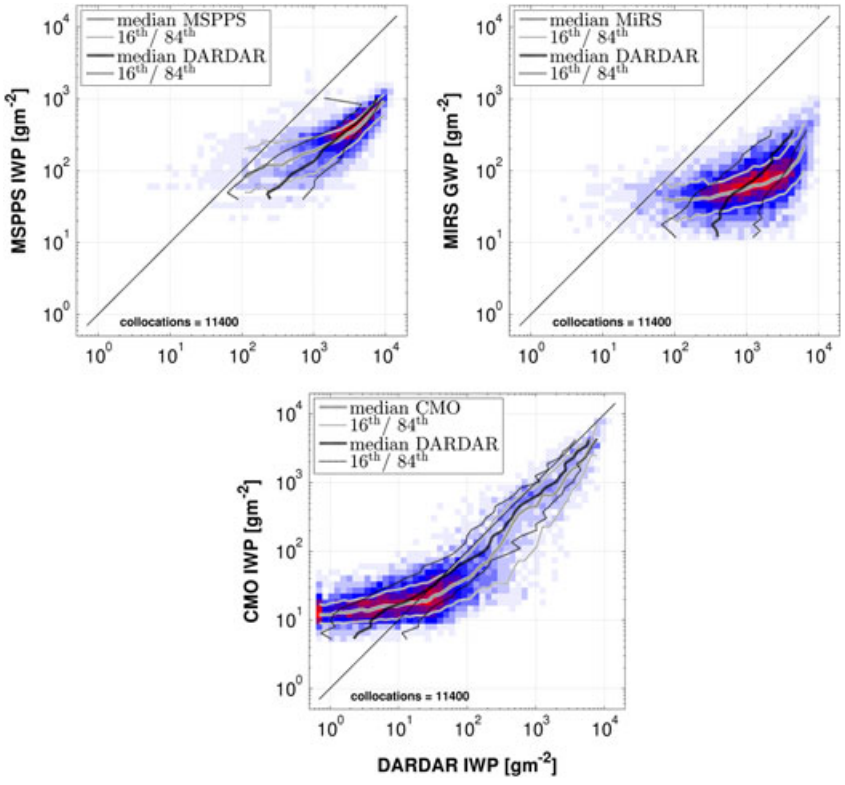

Figure 8. Tropical IWP comparison 2007 for passive microwave techniques: two-dimensional histograms as in Figure 4. MSPPS IWP (top left), MiRS graupel water path (top right), and CMO IWP (bottom).

uncertainties from *Cooper et al.* [2003] and *Austin et al.* [2009] cited earlier for MODIS and DARDAR, respectively. Errors caused by collocations are neglected in this argument.

[64] The differences between the SRBS data sets can be explained by mainly three factors. First, although PATMOS-x and CMSAF are based on the same measurements, the difference between them is systematic and large. This is due to the choice of ice particle model (see section 2.2.3). The largest uncertainties in these retrievals are a direct result of the assumed microphysics [*Zhang et al.*, 2009]. Second, because MODIS IWP retrieval algorithm can use five channels compared with two for CMSAF and PATMOS-x, it is conceivable that it has better constrained retrievals. Third, the different sizes of the pseudo footprints of roughly 10×10 km, 4×1 km, and 1×1 km for PATMOS-x, CMSAF, and MODIS, respectively, may also play a role, and because larger footprints should have more collocation errors (see section 3), this may be the reason PATMOS-x has a larger random uncertainty than MODIS, although the systematic errors are the same.

4.5. Passive Microwave Measurements

[65] Three data sets that retrieve IWP using passive microwave data are shown in Figure 8. Using the medians as before, comparisons of IWP can be made in the ranges from approximately 900 g m^{-2} for MSPPS (left), from approximately 700 g m^{-2} for MiRS (right), and from approximately 10 g m^{-2} for CMO. There is no clear upper bound to valid IWP ranges for these data sets. It is well established that IWP retrievals using passive microwave are only sensitive to clouds that have precipitation-sized ice particles, and this is reflected in this comparison. As shown in the histogram of CMO versus DARDAR in Figure 8, CMO reports IWP even where DARDAR reports very low IWP values. This is purely an artifact of this current version of the data set because the retrieval will always retrieve IWP greater than 0.

[66] As shown in both Figures 8 and 9, the systematic error between MSPPS versus DARDAR and the systematic error between MiRS versus DARDAR is very large even for clouds where both the DARDAR and these microwave data sets have sensitivity to clouds. The MSPPS and MiRS data sets are probably not suited for such a comparison, but it is worth noting that when clouds have an IWP larger than approximately 2000 g m^{-2}, they do appear to be fairly correlated with DARDAR, albeit offset to a factor of 5 to 6 lower IWP values. The correlation improves with increasing IWP because of the steady increase of precipitation. Once the IWP is larger than approximately 2000 g m^{-2}, the random error of the collocated microwave data sets is not particularly large. MSPPS has a random error of mostly less than a factor of 1.5, and MiRS has a larger random error. The increased random error of MiRS is likely due to the much larger footprint sizes of MiRS compared with MSPPS (see section 3.1).

[67] The third microwave data set, CMO stands out because of its different retrieval approach and higher frequency microwave channels, and the valid IWP range defined here is very large (see Figures 8 and 9). Note that CMO is developed based on collocations with CPR; therefore, the random errors are not assumed to be uncorrelated. CMO reports higher values than DARDAR in the low IWP range. CMO is known not to be retrieving clouds here because it is insensitive to clouds in this range. The version compared in this study does not include any cloud detection and retrieves IWP in logarithmic space. Therefore, it retrieves a low value of IWP even when there is no ice at all. A good way of cutting off those values is still under development. Although the data set is still in a fairly early development stage, it already looks promising.

Table 3. Approximate Valid IWP Ranges: For Each Data Set Based on the IWP Bins or DARDAR, and the Range Selection Criteria Presented in Section 4.2.1[a]

Data Set	Minimum IWP [g m^{-2}]	Maximum IWP [g m^{-2}]
RO	10	9000
IORO	10	9000
MODIS	40	—
CMSAF	40	—
PATMOS-x	40	—
MSPPS	900	8000
MiRS	700	—
CMO	10	—

[a]The missing upper bounds for CMO and MiRS indicate that a clear upper limit to their valid ranges was not found.

5. Conclusions

[68] This article is a necessary continuation of the *Eliasson et al.* [2011] study where the satellite observations were compared in terms of monthly mean IWP. The results presented in this article help explain the differences between the L3 observational data sets found in *Eliasson et al.* [2011], which could only be inferred using the monthly mean data. However, comparing monthly mean products is problematic because different approaches are used to go from an L2 ("footprint" resolution) to L3 (gridded monthly mean) prod-

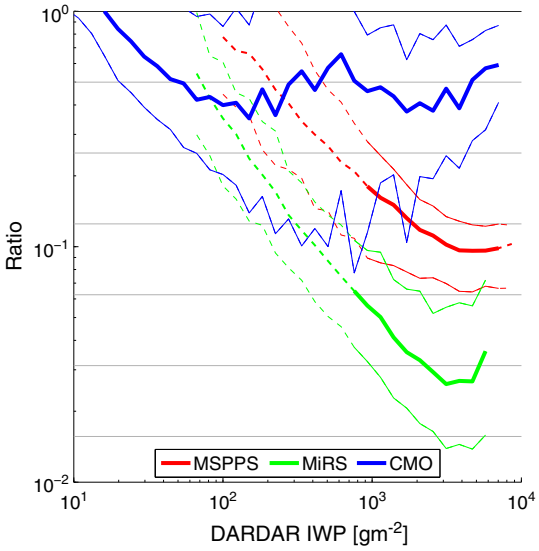

Figure 9. IWP from passive microwave data sets. The median and "pseudo 1-sigma" of the ratios of MSPPS versus DARDAR, MiRS versus DARDAR, and CMO versus DARDAR are shown together. Thick lines indicate medians; thin lines indicate "pseudo 1-sigma"; dashed lines indicate IWP ranges where the data sets are uncorrelated. Note that the *y* axis differs from Figures 5 and 7.

uct. IWP derived from different techniques were compared with DARDAR at a "footprint" resolution, using data from 2007 between ±30° latitude. In this study, DARDAR is assumed to provide the most accurate retrieval of IWP, because of using the combination of CALIPSO and CloudSat measurements. DARDAR IWP is further assumed to be valid over a wide range of IWP, thus being ideal for the investigations presented. Using the fractional errors of DARDAR IWC, we found that the 1-sigma fractional errors in column integrated IWP range between 20 and 50% and are generally decreasing with increasing IWP (see section 4.1).

[69] The range of IWP where both DARDAR and the collocated data set are sensitive to clouds are first identified for each collocated data set and summarized in Table 3. The focus of this study is on the assessment of various available IWP data sets, i.e., investigating their valid IWP range using DARDAR as reference. In these IWP ranges, we have looked at the systematic bias of IWP and the combined random errors. These are then used to assess the quality of the data sets in Figures 5, 7, and 9.

[70] The combined uncertainties in log space of the collocated data sets are reported in section 4. It is shown that data sets based on solar reflectances can be compared with DARDAR over a large IWP range starting from approximately 30 g m^{-2}. The combined uncertainty of DARDAR and MODIS is less than a factor of 2, and this sets the upper limit for the random error of either data set. CMSAF and PATMOS-x have larger combined uncertainties (see Figure 7).

[71] All SRBS data sets' systematic errors increase to a factor of 4 negative bias as their signals start to attenuate at high IWP values. CMSAF is biased low compared with MODIS and PATMOS-x because of different assumptions on the ice particle models in the retrievals (see section 2.2). Data sets based on passive microwave measurements are only comparable with DARDAR in ranges starting from approximately 700 g m^{-2}. In this range, their combined uncertainty is not larger than the SRBS data sets despite their larger footprints, but the systematic errors are very large. The bias of MSPPS/MiRS compared with DARDAR ranges from a factor of 5 to 7 and 8 to 10, respectively.

[72] We have concluded that all assessed data sets provide valuable information on IWP. The results presented in Table 3 suggest that a synergistic passive retrieval using microwave and shortwave may be able to retrieve IWP for a larger range of values than retrievals based on either alone. Such a retrieval could be developed in a similar way to the CMO data set with a DARDAR reference, therefore truly exploiting the benefits of active and passive shortwave and microwave, and thus constrain IWP as much as possible from a spaceborne platform.

[73] It is encouraging that MODIS, CMSAF, and PATMOS-x show reasonably good agreement over a large IWP range despite the remaining large systematic biases. This is especially encouraging because CMSAF and PATMOS-x are based on AVHRR measurements that span over 30 years. These data sets, are suitable for long-term climate applications. Future work may also include comparing the data sets separated into, e.g., cloud types or specific atmospheric scenarios.

[74] **Acknowledgments.** We thank the Swedish Science Council for cofunding the earlier phase and the Swedish National Space Board for cofunding the later phase of this study. We thank the ICARE Data and Services Center and LAADS web for providing access to the data used in this study. We thank Patrick Eriksson from Chalmers University of Technology, Sweden, and Andy Heidinger from the Center for Satellite Applications and Research for numerous e-mail discussions related to this study. We thank the Swedish National Research School in Space Technology for offering courses that have helped S.E. and G.H. in their research.

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.

Baum, B. A., A. J. Heymsfield, P. Yang, and S. T. Bedka (2005), Bulk scattering properties for the remote sensing of ice clouds. Part I: Microphysical data and models, *J. Appl. Meteorol.*, *44*, 1885–1895.

Boukabara, S.-A., K. Garrett, W. Chen, F. Iturbide-Sanchez, C. Grassotti, C. Kongoli, R. Chen, Q. Liu, B. Yan, F. Weng, R. Ferraro, T. J. Kleespies, and H. Meng (2011), MiRS: An all-weather 1DVAR satellite data assimilation and retrieval system, *IEEE Trans. Geosci. Remote*, *49*(9), 3249–3272, doi:10.1109/TGRS.2011.2158438.

Buehler, S. A., E. Defer, F. Evans, S. Eliasson, J. Mendrok, P. Eriksson, C. Lee, C. Jimenéz, C. Prigent, S. Crewell, Y. Kasai, R. Bennartz, and A. J. Gasiewski (2012), Observing ice clouds in the submillimeter spectral range: The CloudIce mission proposal for ESA's Earth Explorer 8, *Atmos. Meas. Tech.*, *5*, 1529–1549, doi:10.5194/amt-5-1529-2012.

Buehler, S. A., C. Jiménez, K. F. Evans, P. Eriksson, B. Rydberg, A. J. Heymsfield, C. Stubenrauch, U. Lohmann, C. Emde, V. O. John, T. R. Sreerekha, and C. P. Davis (2007), A concept for a satellite mission to measure cloud ice water path and ice particle size, *Q. J. R. Meteorol. Soc.*, *133*(S2), 109–128, doi:10.1002/qj.143.

Comstock, J. M., R. D. Entremont, D. DeSlover, G. G. Mace, S. Y. Matrosov, S. A. McFarlane, P. Minnins, D. Mitchell, K. Sassen, M. D. Shupe, D. D. Turner, and Z. Wang (2007), An intercomparison of microphysical retrieval algorithms for upper-tropospheric ice clouds, *Bull. Am. Meteorol. Soc.*, pp. 191–204, doi:10.1175/BAMS-88-2-191.

Cooper, S. J., T. S. L'Ecuyer, P. Gabriel, A. J. Baran, and G. L. Stephens (2007), Performance assessment of a five-channel estimation-based ice cloud retrieval scheme for use over the global oceans, *J. Geophys. Res.*, *112*, D04207, doi:10.1029/2006JD007122.

Cooper, S. J., T. S. L'Ecuyer, and G. L. Stephens (2003), The impact of explicit cloud boundary information on ice cloud microphysical property retrievals from infrared radiances, *J. Geophys. Res.*, *108*, doi:10.1029/2002JD002611.

Delanoë, J., and R. J. Hogan (2008), A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer, *J. Geophys. Res.*, *113*, D07204, doi:10.1029/2007JD009000.

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.

Devasthale, A., and M. A. Thomas (2012), Sensitivity of cloud liquid water content estimates to the temperature dependent thermodynamic phase: A global study using CloudSAT data, *J. Climate*, doi:10.1175/JCLI-D-11-00521.1.

Dybbroe, A., K.-G. Karlsson, and A. Thoss (2005a), NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative modelling - part I: Algoritm description, *J. Appl. Meteorol.*, *44*, 39–54.

Dybbroe, A., K.-G. Karlsson, and A. Thoss (2005b), NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative modelling - part II: Tuning and validation, *J. Appl. Meteorol.*, *44*, 55–71.

Eliasson, S., S. A. Buehler, M. Milz, P. Eriksson, and V. O. John (2011), Assessing observed and modelled spatial distributions of ice water path using satellite data, *Atmos. Chem. Phys.*, *11*, 375–391, doi:10.5194/acp-11-375-2011.

Ferraro, R. R., F. Weng, N. C. Grody, L. Zhao, H. Meng, C. Kongoli, P. Pellegrino, S. Qiu, and C. Dean (2005), NOAA operational hydrological products derived from the advanced microwave sounding unit, *IEEE Trans. Geosci. Remote*, *43*(5), 1036–1049.

Hartmann, D. L., M. E. Ockert-Bell, and M. L. Michelsen (1992), The effect of cloud type on earth's energy balance: Global analysis, *J. Climate*, pp. 1281–1304.

Heidinger, A. K., W. C. Straka, C. C. Molling, J. T. Sullivan, and X. Q. Wu (2010), Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record, *Int. J. Remote Sens.*, *31*, 6493–6517, doi:10.1080/01431161.2010.496472.

Hess, H., R. B. A. Koelemeijer, and P. Stammes (1998), Scattering matrices of imperfect hexagonal crystals, *J. Quant. Spectrosc. Radiat. Transfer*, *60*, 301–308.

Heymsfield, A. J., S. Matrosov, and B. Baum (2003), Ice water path - optical depth relationships for cirrus and deep stratiform ice cloud layers, *J. Appl. Meteorol.*, *42*(20), 1369–1390.

Heymsfield, A. J., and G. M. McFarquahar (2002), Mid-latitude and tropical cirrus, in Cirrus, edited by D. K. Lynch, K. Sassen, D. Starr, and G. Stephens, pp. 78–101, Oxford University Press, New York.

Heymsfield, A. J., A. Protat, R. Austin, D. Bouniol, R. Hogan, J. Delanoë, H. Okamoto, K. Sato, G.-J. van Zadelhoff, D. Donovan, and Z. Wang (2008), Testing IWC retrieval methods using radar and ancillary measurements with in situ data, *J. Appl. Meteorol. Clim.*, *47*(1), 135–163, doi:10.1175/2007JAMC1606.1.

Holl, G., S. A. Buehler, B. Rydberg, and C. Jiménez (2010), Collocating satellite-based radar and radiometer measurements – methodology and usage examples, *Atmos. Meas. Tech.*, *3*, 693–708, doi:10.5194/amt-3-693-2010.

Iturbide-Sanchez, F., S.-A. Boukabara, R. Chen, K. Garrett, C. Grassotti, W. Chen, and F. Weng (2011), Assessment of a variational inversion system for rainfall rate over land and water surfaces, *IEEE Trans. Geosci. Remote*, doi:10.1109/TGRS.2011.2119375.

Jiang, J. H., et al. (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA a-train satellite observations, *J. Geophys. Res.*, *117*, D14105, doi:10.1029/2011JD017237.

John, V. O., G. Holl, S. A. Buehler, B. Candy, R. W. Saunders, and D. E. Parker (2012), Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses, *J. Geophys. Res.*, *117*(D2), D02305, doi:10.1029/2011JD016349.

John, V. O., and B. J. Soden (2006), Does convectively-detrained cloud ice enhance water vapor feedback?, *Geophys. Res. Lett.*, *33*, L20701, doi:10.1029/2006GL027260, see corrections in John and Soden (2006), *Geophys. Res. Lett.*, *33*, L23701, doi:10.1029/2006GL028663.

Khvorostyanov, V. I., and K. Sassen (2002), Microphysical processes in cirrus and their impact on radiation, in Cirrus, edited by D. K. Lynch, K. Sassen, D. Starr, and G. Stephens, pp. 397–432, Oxford University Press, New York.

King, M. D., S.-C. Tsay, S. E. Platnick, M. Wang, and K.-N. Liou (1997), Cloud retrieval algorithms for modis: Optical thickness, effective particle radius, and thermodynamic phase, *Tech. Rep. ATBD-MOD-05*, NASA Goddard Space Flight Center MODIS Algorithm Theoretical Basis Document.

King, M. D., W. P. Menzel, Y. J. Kaufman, D. Tanre, B.-C. Gao, S. Platnick, S. S. Ackerman, L. A. Remer, R. Pincus, and P. A. Hubanks (2003), Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS, *IEEE Trans. Geosci. Remote*, *41*(2), 442–458.

King, M. D., S. Platnick, P. A. Hubanks, G. T. Arnold, E. G. Moody, G. Wind, and B. Wind (2006), Collection 005 change summary for the MODIS cloud optical property (06_OD) algorithm, *Tech. Rep.*, 23 pp, NASA Collection 005 Change Summary Document.

Koh, T.-Y., Y. S. Djamil, and C.-K. Teo (2011), Statistical dynamics of tropical wind in radiosonde data, *Atmos. Chem. Phys.*, *11*, 4177–4189, doi:10.5194/acp-11-4177-2011.

Li, J.-L. F., 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.

L'Ecuyer, T. S., N. B. Wood, T. Haladay, G. L. Stephens, and P. W. S. Jr. (2008), Impact of clouds on atmospheric heating based on the r04 cloudsat fluxes and heating rates data set, *J. Geophys. Res.*, *113*, D00A15, doi:10.1029/2008JD009951.

Mace, G. G., and S. Benson (2008), The vertical structure of cloud occurrence and radiative forcing at the SGP ARM site as revealed by 8 years of continuous data, *J. Climate*, *21*(11), 2591–2610, doi:10.1175/2007JCLI1987.1.

McFarquhar, G. M., and A. J. Heymsfield (1998), The definition and significance of an effective radius for ice clouds, *J. Atmos. Sci.*, *55*, 2039–2052.

Meyer, K., P. Yang, and B.-C. Gao (2006), Tropical ice cloud optical depth, ice water path, and frequency fields inferred from the MODIS level-3 data, *Atmos. Res.*, *85*, 171–182, doi:10.1016/j.atmosres.2006.09.009.

Nakajima, T., and M. D. King (1990), Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: Theory, *J. Atmos. Sci.*, *47*(15), 1878–1893.

Protat, A., J. Delanoe, E. J. O'Connor, and T. S. L'Ecuyer (2010), The evaluation of cloudsat and calipso ice microphysical products using ground-based cloud radar and lidar observations, *J. Atmos. Oceanic Technol.*, *27*, 793–810, doi:10.1175/2009JTECHA1397.1.

Roebeling, R. A., A. J. Feijt, and P. Stammes (2006), Cloud property retrievals for climate monitoring: Implications of differences between SEVIRI on METEOSAT-8 and AVHRR on NOAA-17, *J. Geophys. Res.*, *111*, D20210, doi:10.1029/2005JD006990.

Smith, G. L., B. R. Barkstrom, E. F. Harrison, R. B. Lee III, and B. A. Wielicki (1994), Radiation budget measurements for the eighties and nineties, *Adv. Space Res.*, *14*(1), 81–84.

Stammes, P. (2001), IRS 2000: Current Problems in Atmospheric Radiation, pp. 385–388, A. Deepak Publ., Hampton, VA.

Stein, T. H. M., J. Delanoë, and R. J. Hogan (2011), A comparison among four different retrieval methods for ice-cloud properties using data from CloudSat, CALIPSO, and MODIS, *J. Appl. Meteorol. Clim.*, *50*, 1952–1969, doi:10.1175/2011JAMC2646.1.

Stephens, G. L., S. Tsay, P. W. Stackhouse Jr., and P. J. Flatau (1990), The relevance of the microphysical and radiative properties of cirrus clouds to climate and climatic feedback, *J. Atmos. Sci.*, *47*(14), 1742–1753.

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, *Bull. Am. Meteorol. Soc.*, *83*, 1771–1790.

Vivekanandan, J., J. Turk, and V. N. Bringi (1991), Ice water path estimation and characterization using passive microwave radiometry, *J. Appl. Meteorol.*, *30*, 1407–1421.

Waliser, D. E., J.-L. F. Li, C. P. Woods, R. T. Austin, J. Bacmeister, J. Chern, A. D. 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.

Walther, A., and A. K. Heidinger (2012), Implementation of the daytime cloud optical and microphysical properties algorithm (DCOMP) in PATMOS-x, *J. Appl. Meteorol. Clim.*, doi:10.1175/JAMC-D-11-0108.1.

Wu, D. L., et al. (2009), Comparisons of global cloud ice from MLS, CloudSat, and correlative data sets, *J. Geophys. Res.*, *114*, D00A24, doi:10.1029/2008JD009946.

Wyant, M. C., C. S. Bretherton, J. T. Bacmeister, J. T. Kiehl, I. M. Held, M. Zhao, S. A. Klein, and B. J. Soden (2006), A comparison of low-latitude cloud properties and their response to climate change in three agcms sorted into regimes using mid-tropospheric vertical velocity, *Clim. Dyn.*, *27*, 261–279, doi:10.1007/s00382-006-0138-4.

Zhang, Z., P. Yang, G. Kattawar, J. Riedi, L. C. Labonnote, B. A. Baum, S. Platnick, and H. L. Huang (2009), Influence of ice particle model on satellite ice cloud retrieval: Lessons learned from MODIS and POLDER cloud product comparison, *Atmos. Chem. Phys.*, *9*, 7115–7129, doi:10.5194/acp-9-7115-2009.

Zhang, Z., S. Platnick, P. Yang, A. K. Heidinger, and J. M. Comstock (2010), Effects of ice particle size vertical inhomogeneity on the passive remote sensing of ice clouds, *J. Geophys. Res.*, *115*, D17203, doi:10.1029/2010JD013835.

Zhao, L., and F. Weng (2002), Retrieval of ice cloud parameters using the advanced microwave sounding unit, *J. Appl. Meteorol.*, *41*, 384–395.