**Atmospheric
Chemistry
and Physics**

Assessing observed and modelled spatial distributions of ice water path using satellite data

S. Eliasson[1], S. A. Buehler[1], M. Milz[1], P. Eriksson[2], and V. O. John[3]

[1]Department of Space Science, Luleå Univ. of Technology, Kiruna, Sweden
[2]Department of Earth and Space Sciences, Chalmers Univ. of Technology, Göteborg, Sweden
[3]Met Office Hadley Centre, Exeter, UK

Abstract. The climate models used in the IPCC AR4 show large differences in monthly mean ice water path (IWP). The most valuable source of information that can be used to potentially constrain the models is global satellite data. The satellite datasets also have large differences. The retrieved IWP depends on the technique used, as retrievals based on different techniques are sensitive to different parts of the cloud column. Building on the foundation of Waliser et al. (2009), this article provides a more comprehensive comparison between satellite datasets. IWP data from the Cloud-Sat cloud profiling radar provide the most advanced dataset on clouds. For all its unmistakable value, CloudSat data are too short and too sparse to assess climatic distributions of IWP, hence the need to also use longer datasets. We evaluate satellite datasets from CloudSat, PATMOS-x, ISCCP, MODIS and MSPPS in terms of monthly mean IWP, in order to determine the differences and relate them to the sensitivity of the instrument used in the retrievals. This information is also used to evaluate the climate models, to the extent that is possible.

ISCCP and MSPPS were shown to have comparatively low IWP values. ISCCP shows particularly low values in the tropics, while MSPPS has particularly low values outside the tropics. MODIS and PATMOS-x were in closest agreement with CloudSat in terms of magnitude and spatial distribution, with MODIS being the better of the two. Additionally PATMOS-x and ISCCP, which have a temporal range long enough to capture the inter-annual variability of IWP, are used in conjunction with CloudSat IWP (after removing profiles that contain precipitation) to assess the IWP

variability and mean of the climate models. In general there are large discrepancies between the individual climate models, and all of the models show problems in reproducing the observed spatial distribution of cloud-ice. Comparisons consistently showed that ECHAM-5 is probably the GCM from IPCC AR4 closest to satellite observations.

1 Introduction

Ice clouds are an important part of Earth's climate system. Knowledge of the distribution and properties of ice clouds is central to understanding the atmospheric water budget, as their distribution strongly affects precipitation and the water cycle. Ice clouds also have a strong effect on the radiation budget of the atmosphere. They cool the atmosphere by reflecting incoming solar radiation, but also heat the atmosphere by absorbing and re-emitting outgoing terrestrial radiation. The magnitude of both processes, hence the net radiative impact of ice clouds, depends on macro-physical properties such as cloud top temperature, and vertical and horizontal extent, and on micro-physical properties such as ice crystal shape, cloud optical thickness and effective radius (Ramanathan et al., 1989). One important ice cloud quantity is the vertical column integral of the cloud Ice Water Content (IWC), which includes all types of ice particles. This quantity is called Ice Water Path (IWP) and commonly has the units g/m^2. There are large differences in IWP between climate models (e.g., Waliser et al., 2009; John and Soden, 2006). Waliser et al. (2009) highlighted that there is a lack of adequate cloud property measurements with which to constrain the models. The differing model assumptions made about ice particle size, mass, and cross-sectional area

Correspondence to: S. Eliasson
(s.eliasson@ltu.se)

also contribute to the large biases in the models, because the assumptions directly effect particle fall velocities (Heymsfield and Iaquinta, 2000). Although global and continuous datasets of satellite retrieved cloud properties are now available, their use to validate climate models is fraught with difficulties. This is largely due to the definition of IWP itself. Models make a clear distinction between precipitating ice particles and suspended cloud ice particles, whereas retrieved IWP constitutes a mixture of both. Satellite derived IWPs, are defined in the same manner for all datasets, but are largely different from each other. This must be taken into account when comparing observational datasets to each other and to models, as done in this article.

In a qualitative IWP comparison study, Waliser et al. (2009) illustrated a large level of disagreement between the available datasets. The objective of this paper is to provide a more quantitative comparison between the observed datasets, based on the foundation provided by Waliser et al. (2009), especially in terms of climatic distributions of IWP, using monthly mean values. This information is in turn used to compare observations to models, with an intent to evaluate both the variability and the mean. Basically, the most important reason for the difference between the datasets is that retrievals based on different measurement techniques are inherently different, as they are sensitive to different parts of the cloud column (described in Sect. 2.2.6). Additionally, the retrieval accuracy is limited by the uncertainties made in the cloud microphysical assumptions, and the uncertainties in the a priori background (e.g., Wu et al., 2009; Eriksson et al., 2008). Despite these shortcomings, satellite data remain the most valuable source of information as in situ measurements of IWP are few and far apart. There are many satellite datasets available that provide an IWP product, some with a temporal coverage of up to 25 years. The range of datasets are based on techniques covering spectral bands ranging from microwave to visible using passive sensors or active instruments. Our knowledge on IWP has increased significantly through the introduction of CloudSat, and is therefore the main reference dataset in this article.

This article provides a comprehensive comparison of satellite datasets, such comparison, in the long run can potentially be used to constrain model IWP output. In particular, we anchored monthly mean IWP of long term satellite datasets to CloudSat, as CloudSat retrieval accuracies have been previously quantified using in situ data from several campaigns in Heymsfield et al. (2008). We also evaluated the performance of various climate models in terms of the distribution of monthly mean IWP on climatic time scales. The climate models used in this study are a subset of those included in the fourth assessment report (AR4) in the Intergovernmental Panel on Climate Change (2007).

Section 2 provides short descriptions of the chosen climate models and satellite datasets. Section 3 provides quantitative and qualitative results from the comparison study. Comparisons of satellite datasets, and satellite-model comparisons,

are done in parallel throughout the results section. Section 4 contains the discussion and conclusion.

2 Description of datasets

2.1 General circulation models

All models in AR4 provide monthly averages of IWP, which contains only the suspended cloud-ice portion of the column, with a temporal range of at least 100 years from 1900. For the remainder of this article, model "cloud-ice" will be simply referred to as IWP. Table 1 shows the details of the sub-set of AR4 models used in this article. The table depicts model details such as the model resolution, the full name of the model, its institute, and the short name further used in this article. They are presented in no particular order. Further descriptions of the climate models in AR4 can be found at http://www-pcmdi.llnl.gov/ipcc/model_documentation/ipcc_model_documentation.php. We have chosen a somewhat ad hoc selection of 6 models which, in themselves, roughly represent the inter-model variability of IWP of all models in AR4. ECHAM is a high resolution model with comparatively low IWP values, CCSM has the highest resolution and the lowest IWP averages of all models, CSIRO has a relatively high horizontal resolution and relatively high IWP averages, GISS has the coarsest resolution and has very high IWP averages, INM is a low resolution model with very low IWP values in general, and UKMO has a relatively high resolution and shows remarkably low IWP values in the tropics compared to outside the tropics.

2.1.1 Common model features

In general, GCMs generate ice clouds through the convergence of moist air masses leading to condensation by large-scale dynamics. This is added to the existing cloudiness in the grid box previously determined either diagnostically by relative humidity or prognostically as cloudiness from an earlier time step. Moist convection also adds cloud-ice to existing ice cloud. When comparing climate model output to satellite data, it is important to recognise that the representation of IWP is only comparable to satellite observations to a certain extent. Climate models distinguish between precipitating ice, such as snow and graupel, and cloud-ice, which remains suspended aloft. The precipitated ice is removed at each time step and it either sublimates as it falls out or reaches the ground as precipitation. Only the cloud-ice remaining from these processes is saved as a diagnostic variable, and stored in the AR4 archive. Waliser et al. (2009) suggest models may have ratios of suspended, or floating cloud ice compared to the total column of ice particles, further called cloud-ice column ratio, in the order of 0.1 to 0.3. We compared the models in terms of absolute column water mass (Precipitable Water, Liquid Water Path and IWP), and detected large differences between the models (not shown).

Table 1. IPCC Global Climate Models. This table provides an overview of the resolution and country of origin for the subset of IPCC AR4 models used in this study.

Short name	Horizontal Resolution	Atm. Layers	Model Name	Institute
ECHAM	$1.9 \times 1.9°$	19	ECHAM5/MPI-OM	Max Planck Institute for Meteorology (Germany)
CCSM	$1.4 \times 1.4°$	26	Community Climate System Model, version 3.0 (CCSM3)	National Centre for Atmospheric Research (USA)
CSIRO	$1.9 \times 1.9°$	18	CSIRO Mark 3.0, Climate System Model	Commonwealth Scientific and Industrial Research Organisation (Australia)
GISS	$5 \times 5°$	15	GISS ModelE-R atmospheric, General Circulation Mode	NASA Goddard Institute for Space Studies (GISS) (USA)
INM	$5 \times 4°$	21	INMCM3.0	Institute of Numerical Mathematics, Russian Academy of Science
UKMO	$1.25 \times 1.875°$	38	Hadley Centre Global Environmental Model, version 1 (HadGEM1)	Hadley Centre for Climate Prediction and Research and UK Met office

Therefore, statements about the absolute differences in IWP between the models should take the column cloud-ice ratio into account.

2.2 Satellite IWP datasets

IWP is retrieved over a wide radiative spectrum, ranging from microwave to visible wavelengths. We have chosen datasets with retrievals throughout this range. The satellite datasets used in this survey are presented in no particular order below. In contrast to models, satellite retrievals of ice clouds do not make a distinction between ice particles, whether precipitating or suspended, and this must also be taken into account in all model-satellite comparisons.

2.2.1 CloudSat

CloudSat data are provided by Colorado State University and NASA Jet Propulsion Laboratory. CloudSat is part of the A-train, which is a constellation of satellites flying in close formation enabling maximum collocations between these satellites (Stephens et al., 2002). To date, CloudSat provides the most advanced satellite dataset on clouds. In contrast to the other satellite datasets in the survey, which are based on passive remote sensing techniques, CloudSat has a 94GHz, 0.16° off-nadir looking Cloud Profiling Radar. It has a high vertical resolution of 500m, enabling the retrieval of cloud vertical structures, and has a small horizontal footprint of approximately 1.5km.

The IWP product used in this study is RO_ice_water_path from the 2B-CWC-RO CloudSat dataset, version 008. IWP is retrieved from the detected back-scatter from the Cloud Profiling Radar in conjunction with model temperature data from the ECMWF model (the CloudSat ECMWF AUX product). The algorithm used for this product is described in Austin et al. (2009). On-line descriptions of CloudSat products are also available at the website: http://www.cloudsat.cira.colostate.edu/dataSpecs.php.

Although precipitating ice is part of the ice column that makes up IWP, from the model standpoint, large precipitating particles in observations lead to larger IWP than intended in the definition of IWP. Therefore, for comparisons of CloudSat IWP to modelled IWP, we removed all profiles that are flagged to contain precipitation at the surface in a similar manner as done in Waliser et al. (2009). The precipitation flag used is in the 2C-PRECIP-COLUMN CloudSat dataset (version 0). This product is only valid over open ocean. (Haynes et al., 2009). The dataset created by applying the precipitation flag to CloudSat profiles is known as $IWP_{noPrecip}$.

Despite probably providing the best estimate of IWP, there are some major uncertainties that must be considered. Firstly, the retrieval uncertainty has been quantified in Heymsfield et al. (2008), where CloudSat IWC retrievals were shown to be within ±40% of in situ measurements (Austin et al., 2009; Waliser et al., 2009). However, the sample from which the 40% error estimate is determined is small, and the study used simulated satellite data rather than real data, so the true retrieval error may be larger than 40%. Secondly, another important, but unquantified source of uncertainty is the cloud phase classification. CloudSat retrievals use a linear liquid to ice ratio as a function of modelled temperature, where all clouds warmer than 273 K are liquid clouds and all clouds colder than 253 K are ice clouds. Thirdly, although precipitation in a profile is not an error source, by attempting to remove precipitation for model comparison additional uncertainties arise. Aside from the uncertainties in determining surface precipitation, one inevitably introduces a dry bias as all cloud-ice associated with precipitation events is removed.

Despite the uncertainties in the IWP$_{noPrecip}$ dataset, even if this observed estimate of cloud-ice column ratio is off by a factor 2, the uncertainty is still small than the one or two orders of magnitude between models (Waliser et al., 2009). In the absence of other information, we assume that 40% is a realistic error estimate for CloudSat IWP, and for the model comparison we additionally use the IWP$_{noPrecip}$ dataset minus 40% for the low end of the CloudSat uncertainty. We have used CloudSat IWP data covering the temporal range of July 2006 to June 2009.

2.2.2 ISCCP

IWP data are provided by the International Satellite Cloud Climatology Project (ISCCP). The D2 dataset, which contains monthly averages, is based on one Infrared (IR) channel around 11 μm and one visible channel around 0.6 μm. This is the only dataset that uses data from geostationary satellites. Radiance data are collected from five geostationary satellites and two polar orbiting satellites, although data from polar-orbiting satellites are complementary and used mainly at high latitudes. The geostationary satellites include the satellites from the METEOSAT (EUMETSAT), GOES (USA), INSAT (India), FY-2C (China), and MTS and GMS (Japan) series. The polar orbiting satellites are from the NOAA series from USA (Rossow and Schiffer, 1991). As ISCCP is mainly based on geostationary satellites, it is best at tropical latitudes. However, until 1997 there was a systematic gap in data coverage over the Indian Ocean and this effect can be seen in the merged data prior to this date. Data from ISCCP are gridded using an equal area grid described in Rossow and Garder (1984), and have a horizontal resolution of $2.5 \times 2.5°$ at the equator. ISCCP does not provide IWP directly. Instead, Water Path monthly means are provided for 15 distinct cloud types, classified according to their cloud top temperature, cloud top pressure, and cloud optical thickness. According to these classifications, an ice cloud is a cloud that has a cloud top temperature colder than 260 K or has a cloud top pressure less than 440 hPa. In much the same approach used in Storelvmo et al. (2008), IWP is calculated from the sum of Water Path values from all cloud types classified as ice clouds and multiplied by the cloud fraction. The difference in approaches lies in that Storelvmo et al. (2008) partly reclassify ice clouds in the tropics into liquid clouds, whereas we have decided not to reclassify any clouds. Refer to Rossow and Schiffer (1991) or the web page: http://isccp.giss.nasa.gov/docs/D-toc.html for a detailed description of the algorithms used in the ISCCP retrievals. We have used data that extend from July 1983 to April 2008.

2.2.3 PATMOS-x

The Pathfinder Atmospheres- Extended (PATMOS-x) dataset is from the Cooperative Institute for Meteorological Satellite Studies, Madison USA. The IWP data product is based on visible, near infrared, and infrared radiances from the AVHRR instrument on board the NOAA Polar- orbiting Operational Environmental Satellites (POES). An algorithm based on Heymsfield et al. (2003) is used to derive IWP using solar reflection and radiance measurements. The PATMOS-x data used in this survey are monthly mean IWP, gridded on a similar grid as ISCCP, but with a resolution of 0.5° at the equator. Data from PATMOS-x span the longest time period in the survey (January 1982 to April 2008). Information about the dataset is available at the web site: http://cimss.ssec.wisc.edu/patmosx/.

2.2.4 MODIS

IWP data are also available from the NASA Moderate Resolution Imaging Spectroradiometer (MODIS). The product used here is from the MODIS Science Team. As with PATMOS-x, IWP is retrieved from visible, near infrared, and infrared channels, yet at a higher spectral resolution. The MODIS instrument is down-looking with a total of 36 spectral channels, and 13 of them are used for retrieving cloud properties. The MODIS instrument is on board the Earth Observing Satellites, Aqua and Terra, where Aqua is located in the A-train (Stephens et al., 2002). All MODIS data are screened using a cloud mask described in Ackerman et al. (1998), and details on the retrieval are described in King et al. (1997). The MODIS data used in this survey are the level 3 monthly cloud product (MYD08_03, collection version 5), column integrated IWP, with a 1° gridded spatial resolution. We have chosen to focus solely on data from Aqua as its data are collocated with CloudSat. The MODIS data used in this survey extend from June 2006 to April 2008. Information about the dataset is available at the web site: http://modis-atmos.gsfc.nasa.gov/index.html.

Table 2. Satellite datasets. The first part of the table contains the short name of the satellite dataset used in this study, the on-board instrument, and the number of channels in the radiance spectrum used for the retrieval. The second part shows the type of IWP product and the temporal range used in this study. MODIS pertains to IWP data from the Aqua satellite only and MSPPS pertains to data from NOAA18 only.

Short Name	Type	Platform(s)	Sensor	Spectrum	No. Ch.
ISCCP	VIS/IR	Geostationary: GOES, GMS, METEOSAT Polar: NOAA	Spectrometers	0.6, 11 μm	2
PATMOS-x	VIS/IR	NOAA-15,16,17,18	AVHRR	0.6–12 μm	5
MODIS	VIS/IR	Aqua	MODIS	0.6–14.2 μm	13
MSPPS	MW (passive)	NOAA-18	MHS	89–183 GHz	5
CloudSat	MW (active)	CloudSat	CPR	94 GHz	1

Short Name	IWP Data Product	temporal range
ISCCP	Level 3 mean Water Path from different cloud classes on an equal area grid (ca. 250×250 km).	198307-200804
PATMOS-x	Level 3 mean IWP product on equal area grid (ca. 50×50 km).	198201-200804
MODIS	Level 3 mean IWP product on $1°$ grid	200606-200804
MSPPS	Level 2, granule	200606-200804
CloudSat	Level 2, granule (2B-CWC-RO)	200606-200906

2.2.5 MSPPS

Satellite IWP data from passive microwave sensors are provided by the Microwave Surface and Precipitation Products System (MSPPS) from the National Environmental Satellite Data and Information Service (NESDIS). The MSPPS IWP product is retrieved using the AMSU-B sensor which has 5 channels ranging from 89 GHz to 183.3 GHz. In short, IWP is retrieved from the dampened microwave emission from the surface, where the dampening of the emitted radiation is due to scattering by large particles. AMSU is flown on the NOAA POES satellites. For further details regarding the AMSU-B instrument, refer to Atkinson (2001). For a description of the MSPPS IWP product and others derived from AMSU refer to Ferraro et al. (2005). As the satellite NOAA-18 has the largest number of collocations with CloudSat, we have chosen to focus on MSPPS data from NOAA-18 only. AMSU-B was replaced by the Microwave Humidity Sounder (MHS), which is very similar in design to its predecessor on NOAA18. MSPPS data used in this survey extend from October 2005 to December 2009.

2.2.6 Satellite dataset summary

IWP from all gridded datasets (MODIS, PATMOS-x and IS-CCP) are multiplied by the corresponding cloud fraction for comparability with granule datasets (CloudSat and MSPPS) and climate models. Refer to Table 2 for an overview of the datasets used in this survey. The table provides the short name, the instruments used in the retrievals, the satellite platforms and the institutions associated with these products. Comparing satellite datasets of retrieved IWP is not straightforward. One major retrieval uncertainty comes from the uncertainty in determining the cloud top and cloud base from retrievals (Wu et al., 2009). This uncertainty is inevitable as cloud-ice signals are generally stronger for high frequencies (VIS, IR) than for low frequencies (passive microwave, radar). This is due to the relation of wavelength and particle size, influencing the particle's scattering properties. Retrievals of IWP based on microwave frequencies such as used by MSPPS and CloudSat, have difficulties detecting thin cirrus clouds, but can retrieve cloud information from thicker ice clouds than retrievals based on IR or VIS instruments can. On the other hand, retrievals such as used in ISCCP, PATMOS-x and MODIS, detect more thin clouds. In truth, each satellite IWP retrieval portrays only part of the true IWP column (Waliser et al., 2009). Figure 1 shows a simple schematic of a cloud, that outlines approximately where cloud property information is obtained depending on the wavelength regions used for the measurement. From this, it is intuitive that retrievals based on different measurement techniques will give different solutions.

As a consequence, the distribution and, especially, the magnitude of IWP is expected to vary between datasets and this must be taken into account when comparing datasets. Although CloudSat gives the best instantaneous measurement of IWP, CloudSat's temporal range is too short to access variability in the monthly mean IWP in the models. The ENSO index had a tendency to positive values (La Nina), throughout the time period that CloudSat has been operational, leading to changes in the IWP spatial distribution. Although the models capture inter-annual variations and may predict the correct variability, the timing of these events will not necessarily coincide with the inter-annual variations of the real world. Also, for assessing the spatial variability of the models, CloudSat's sparse horizontal coverage must be

taken into account. The relative sparsity of measurements
introduces sampling effects for gridded data, that the longer
datasets, based on scanning instruments, are not suscepti-
ble to. Choosing a 5° grid for CloudSat generally mitigates
this effect. Due to this, we also use satellite datasets with
longer temporal ranges and a higher spatial coverage. As
both PATMOS-x and ISCCP have long temporal ranges, they
capture the inter-annual variations of IWP, and are there-
fore suitable for validation of models in terms of variabil-
ity. PATMOS-x, MODIS and MSPPS have high spatial cov-
erage which is useful for spatial distribution comparisons
to models.

3 Observations and results

3.1 Zonal averages

Figure 2 shows the zonal mean IWP for the satellite datasets,
compared on a common period (July 2006 to April 2008).
The uncertainty interval, indicated by the grey shaded area,
is based on the uncertainty of CloudSat measurements, esti-
mated to be 40% by Heymsfield et al. (2008). For the ob-
servational datasets in Fig. 2, the uncertainty is in the range
between CloudSat minus 40% and CloudSat plus 40%. For
example, at the Equator, the average of CloudSat IWP is
$92\,\mathrm{g/m^2}$. Multiplying by a factor 0.6 and 1.4 for the lower
and upper uncertainty bounds respectively, gives an uncer-
tainty interval of 55–$128\,\mathrm{g/m^2}$ at the Equator. For Fig. 3,
$\mathrm{IWP_{noPrecip}}$ is instead used as the reference dataset, to facil-
itate comparison to the models, which do not include pre-
cipitation in their IWP column. In the same manner as for
the uncertainty of the observations, the uncertainty is in the
range $\pm40\%$ of the zonal averages of $\mathrm{IWP_{noPrecip}}$. As this
reference dataset is only valid over ocean, IWP values over
land are masked out in both the observations in Fig. 2 and the
models in Fig. 3 for the zonal comparison.

As seen in Fig. 2, there are large differences in the IWP
magnitude between the satellite datasets. As mentioned in
Sect. 2.2.6, these differences can be attributed to instrumen-
tal sensitivities and, to some extent, the retrieval techniques
used. For the individual datasets, PATMOS-x and MODIS
(AQUA) are in best agreement with CloudSat. Within
$\pm60°$ latitude, statistical comparisons also including MODIS
data from the TERRA satellite (not shown) indicate that
PATMOS-x lies between the two MODIS datasets in terms
of correlation, and root mean square difference compared to
CloudSat. MODIS (TERRA) is always un-collocated with
CloudSat, and as mentioned earlier, MODIS (AQUA) is col-
located with CloudSat and PATMOS-x contains both collo-
cated and un-collocated data. PATMOS-x has a very large
positive bias outside $\pm60°$ latitude, and as MODIS's zonal
average generally lies inside the uncertainty interval at all
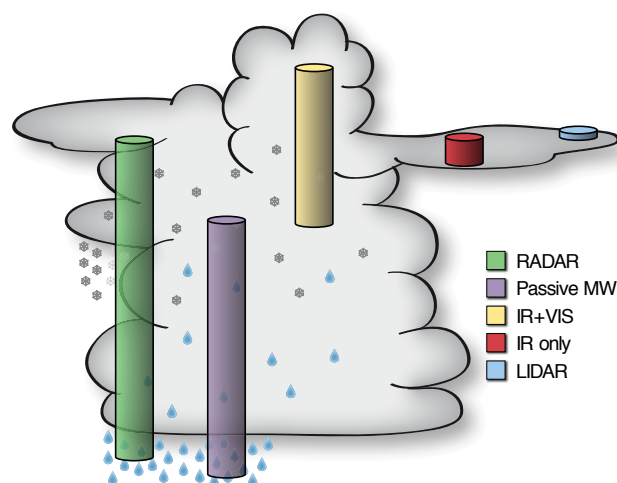latitudes, it is in better agreement with CloudSat overall.

Fig. 1. A schematic figure of a thick cloud. The columns indicate
approximately where in the vertical cloud the different measure-
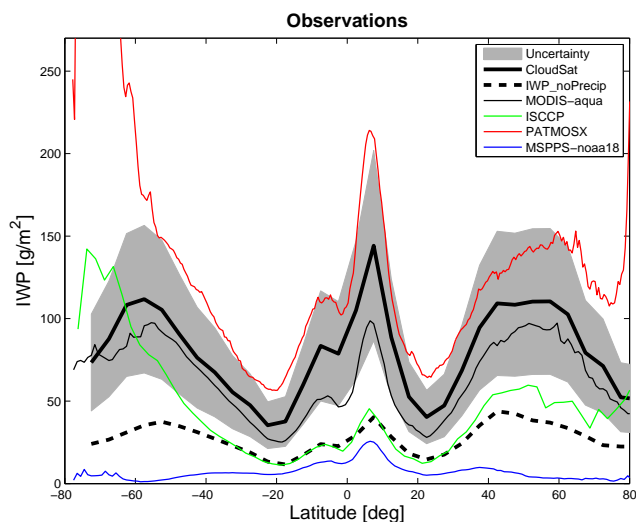ment techniques are sensitive (figure by Oliver Lemke).

Fig. 2. Zonal averages of IWP of satellite datasets over ocean. The
grey shaded area is the uncertainty interval, as described in the text.
The averages are for the period July 2006 to April 2008.

ISCCP appears to have low IWP values in general com-
pared to the other datasets, with particularly low values in
the tropics. Importantly, ISCCP appears to be in good agree-
ment with the $\mathrm{IWP_{noPrecip}}$ dataset in the tropics, but has larger
values outside the tropics. MSPPS has the lowest zonal aver-
ages of IWP, probably because many clouds are completely
undetected. MSPPS is far below both the lower bounds of the
uncertainty interval outside the tropics. In mid-latitudes re-
gions, which are dominated by less thick ice clouds, MSPPS
has very low values. In contrast to ISCCP, MSPPS has its
largest IWP averages in tropical areas. It is close to the
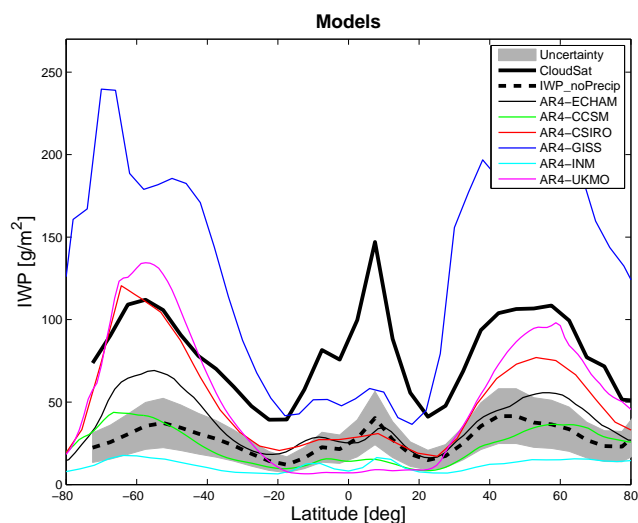$\mathrm{IWP_{noPrecip}}$ dataset in the tropics in terms of absolute IWP,

Fig. 3. Zonal averages of IWP for climate models from 100 years of monthly mean data. The uncertainty level is based on CloudSat data and applied to the $IWP_{noPrecip}$ dataset, as described in the text. A factor 0.5 is applied here to AR4-GISS in order to visualise it in the domain of this figure.

although for different reasons than ISCCP. This can be explained insofar as MSPPS retrievals mainly have sensitivity to large ice particles associated with convective clouds. All passive satellite datasets perform very poorly over snow covered regions, including ice shelves (not shown), as a signal from the cold surface is often indecipherable from a signal from thick ice clouds, and therefore erroneously detected as clouds with high IWP values. For this reason most comparisons are for data inside the latitude range 60°S–60°N, a region henceforth known as the "total"-region,

Figure 3 shows the zonal mean IWP for the models using all available data, albeit for ocean grid boxes only. To reiterate from Sect. 2.2.1, the $IWP_{noPrecip}$ dataset should be a low estimate of IWP as it contains an additional dry bias introduced by removing all cloud-ice where there was surface precipitation. It is clear that there are large differences between observed and modelled IWP averages (Figs. 2 and 3), but also between the models themselves. Within tropical latitudes, about half of the models are below or far below the uncertainty range defined using $IWP_{noPrecip}$, indicating an underestimation of IWP. Outside the tropics, the models are widely spread and, as they do not contain precipitation, probably overestimate IWP. GISS is the exception, as it clearly overestimates IWP at all latitudes (a factor 0.5 is applied to GISS averages in Fig. 3). However, statements on overestimation or underestimation of IWP in the models is somewhat uncertain, as the actual cloud-ice column ratio (suspended amount divided by the total) of the models is not provided.

Additional uncertainties in comparing satellite to model data were highlighted in Waliser et al. (2009). One must account for satellite overpass times when pairing datasets and

models. If the sampled data have different overpass times from the model run time, they are essentially sampling different parts of the diurnal cycle. IWP may fluctuate on the order of ±50% from the mean in the tropics. They also mentioned the need to take sensor and algorithm sensitivities into consideration when comparing satellite to model measurements. There are maximum and minimal retrieval values, which depend on the sensitivities of the retrieval and instruments. The models don't suffer from this, and therefore one should remove modelled cloud properties outside the detectable range of the dataset before comparison. However, we have decided to neglect these effects, as this paper concerns the comparison of modelled monthly mean IWP, to several satellite data on a climatic time scale. Waliser et al. (2009) reported that even amongst some more sophisticated models, their zonal averages of IWC were in disagreement with CloudSat data. Figures 2 and 3 verify this, but information on the longitudinal inhomogeneity of IWP is lost. Therefore, in order to learn more about the differences in IWP between datasets, we also assess the spatial distribution of IWP. In this way we may understand if the differences are larger in some regions than others.

3.2 Spatial distribution of IWP

Figure 4 shows the spatial distribution of CloudSat IWP (left) and the spatial distribution of the ratio of CloudSat $IWP_{noPrecip}$ to IWP (right) between 60° N and 60° S on a 5° grid. The ratio between these two datasets provides a rough estimate of the cloud-ice column ratio, in order to relate the IWP from CloudSat to the model version of IWP (which has no precipitation). According to CloudSat data, the area around the Indonesian archipelago has the highest IWP globally. This region is often referred to as the Tropical Warm Pool (TWP). High values of IWP are also observed at the Inter-Tropical Convergence Zone (ITCZ) and mid-latitude storm tracks, and the smallest IWP averages are found over the subsidence regions in the South Eastern Pacific and Atlantic oceans. Figure 4 (right) shows that for regions with the strongest convection, e.g. associated with the ITCZ and the TWP, the mean of $IWP_{noPrecip}$ is around 20% of the mean IWP. The fraction is around 40% for regions of westerlies in the northern and southern hemispheres, and in the subsidence regions the fraction is closer to 90%. These cloud-ice column ratios are roughly consistent with cloud-ice column ratios of the RAVE-GCM as described in Waliser et al. (2009), and especially its latitudinal dependence (see Fig. 10e therein).

To enable a better qualitative inter-comparison of the spatial distributions of IWP, the data shown in Figs. 5 and 6, are in percentiles. If the data were presented as absolute IWP, the large differences in IWP magnitudes, and the outlier values for each dataset, impact in such a way that many spatial features in the distribution of IWP are harder to discern from the figure. The distribution of gridded IWP values is not

Table 3. Quartile values of the cumulative distribution of gridded IWP data, and the total mean as shown in Figs. 5 and 6. All datasets and statistics are based on a 5° grid, and the unit for IWP is g/m^2.

| | Observations | | | | | |
	CloudSat	IWP_noPrecip	MODIS	ISCCP	PATMOS-x	MSPPS
1st quartile	41	15	35	17	75	3
Median	72	29	63	34	115	7
Mean	78	27	60	35	117	11
3rd quartile	108	38	85	50	152	16

| | Models | | | | | |
	ECHAM	CCSM	CSIRO	GISS	INM	UKMO
1st quartile	21	11	20	81	7	8
Median	35	19	38	180	12	25
Mean	36	20	47	232	11	46
3rd quartile	49	28	70	331	15	78

Fig. 4. CloudSat IWP (left), and the ratio (%) of CloudSat IWP$_{noPrecip}$ and IWP (right) from the period July 2006–June 2009.

Gaussian and have quite different shapes from each other. Therefore they are not readily normalised using mean and standard deviations. The quartiles and the mean are given in Table 3.

Figure 5 shows the spatial distribution of IWP for the satellite datasets for the period July 2006 to April 2008. From left to right, the top panels are IWP and IWP$_{noPrecip}$ data (both from CloudSat), the middle panels are MSPPS and ISCCP, and the bottom panels are MODIS and PATMOS-x. In general, the satellite datasets agree on the spatial distribution of key dynamical features such as the extent of the TWP and ITCZ, but do not show the same level of agreement at higher latitudes. As mentioned in Sect. 2.2, differences in the absolute IWP are expected due to instrument sensitivities. Consequently, judging from the spatial distribution of MSPPS, especially the relatively high IWP values of the tropical continental regions, indicate that its strength lies in detecting IWP of thick ice clouds, e.g. associated with deep convection. This is due to the passive microwave radiation's sensitivity to large ice particles. As indirectly indicated from Fig. 4, deep convective clouds probably contain the highest fraction of ice particles which are large enough to induce detectable dampening of the microwave emission through scattering. Con-

versely, outside the tropics and in the oceanic subsidence regions, clouds are largely undetected, resulting in very low absolute values compared to the other observations and in the median IWP being well below the mean IWP. The spatial distribution of datasets using IR/VIS techniques (PATMOS-x, ISCCP and MODIS) have more likeness to CloudSat. In terms of magnitude, ISCCP has low IWP averages in general compared to CloudSat, PATMOS-x and MODIS, and has particularly low IWP in the convection dominated regions of the tropics. Waliser et al. (2009) argue that ISCCP IWP is expected to roughly coincide with the suspended cloud-ice portion of the atmospheric column, as it is insensitive to precipitation. ISCCP's close agreement with IWP$_{noPrecip}$ shown in Fig 2 support this notion. Although PATMOS-x exhibits larger values of IWP while MODIS somewhat lower values than CloudSat, both of their relative spatial distributions appear to be in general agreement with CloudSat. As can be seen in Fig. 5, the IWP$_{noPrecip}$ dataset has a mean at about 1/3 of the total IWP from CloudSat over the latitude interval, but has relatively less IWP in the convective regions in the tropics.
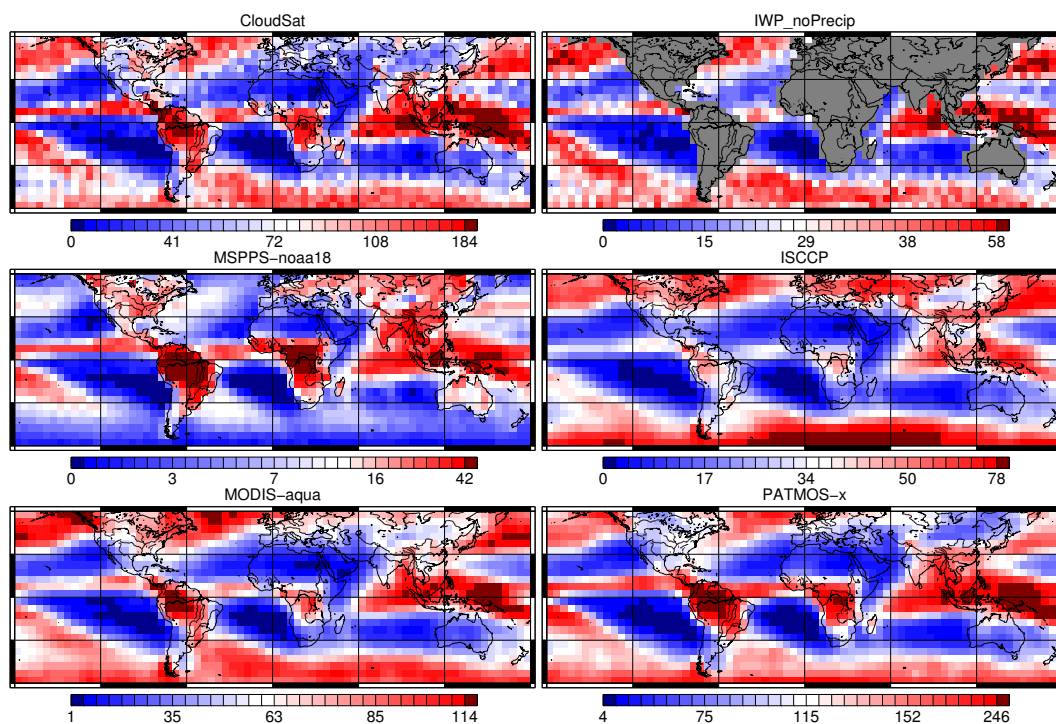
Fig. 5. IWP spatial distribution of satellite data for their common temporal range (July 2006 to April 2008). Based on a polar colour table, the colour "steps" represent $100 \times \frac{1}{32}$-percentile intervals, the "colours" are in reference to the percentile box that contains the mean (which is coloured white), and the values in the legend are the absolute IWP [g/m^2], corresponding to the percentiles at 0%, 1st quartile, median, 3rd quartile and at the $100 \times \frac{31}{32}$ percentile (96%). See also Table 3 for the gridded statistics referred here.
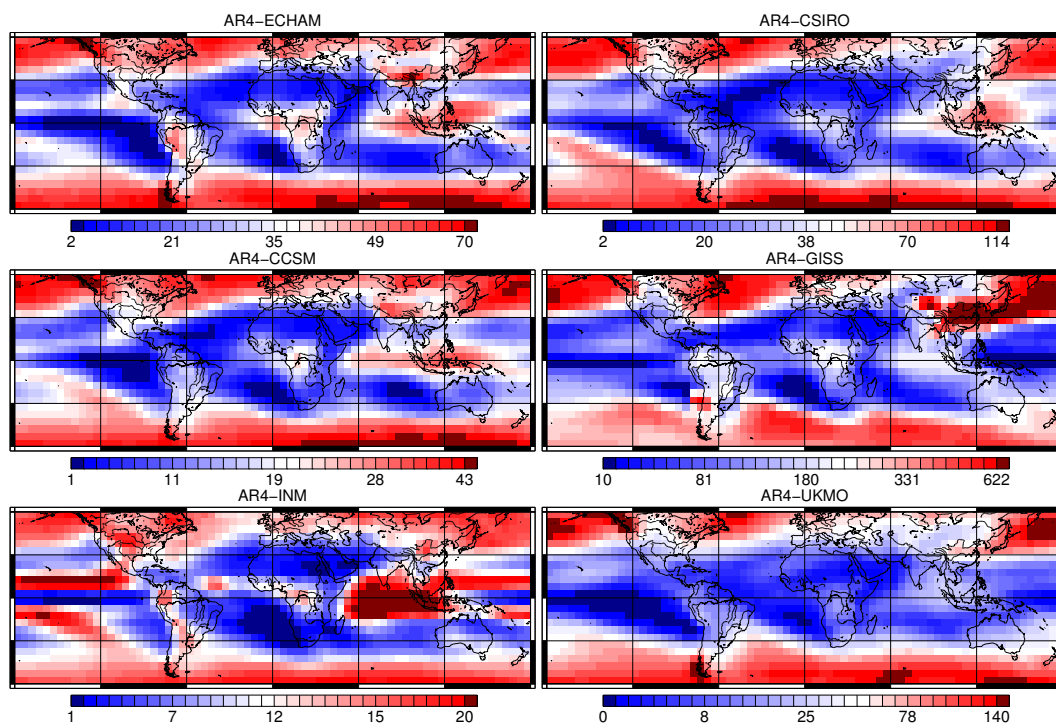

Fig. 6. Same as Fig. 5, except for model data, and all available monthly IWP data are used.

Table 4. IWP Statistics using monthly means [g/m^2] for selected satellite datasets and climate models. Statistics are provided for the total region (60° S–60° N) and the tropical region (30° S–30° N). The Satellite data comparison is from 2007, using a common spatial resolution of 5°, and with CloudSat as the reference dataset. Model statistics are based on the month averages derived from 100 years of data, which has been re-gridded to a 5° grid for comparison. As model IWP can not be readily compared to satellite data, and the IWP$_{noPrecip}$ dataset is only valid over ocean, ECHAM has been arbitrarily chosen as the reference model for the subset of models. RMSD is the root mean square difference and R stands for the spatial and temporal correlation of a dataset compared to the reference dataset.

			Total					Tropics		
Datasets	Mean	Std	Bias	RMSD	R	Mean	Std	Bias	RMSD	R
CloudSat	77	90	0	0	1.00	75	104	0	0	1.00
ISCCP	33	28	−44	88	0.61	23	23	−52	101	0.79
PATMOS-x	109	83	32	71	0.73	105	97	31	72	0.80
MODIS-aqua	58	44	−19	67	0.75	48	47	−27	76	0.82
MSPPS-noaa18	12	18	−65	103	0.64	15	22	−60	108	0.74
AR4-ECHAM	34	18	0	0	1.00	24	13	0	0	1.00
AR4-CCSM	18	11	−15	17	0.93	12	7	−12	14	0.82
AR4-CSIRO	42	31	8	19	0.90	25	17	1	10	0.80
AR4-GISS	211	197	178	257	0.62	115	127	91	154	0.29
AR4-INM	11	5	−23	27	0.67	10	6	−15	18	0.65
AR4-UKMO	39	43	6	30	0.82	11	10	−13	19	0.35

Figure 6 shows the spatial distribution of IWP percentiles for the selected climate models, averaged over all available years (~100 years). The size and location of key dynamical features differs between the models. All models show elevated IWP averages, in relation to their total mean, around the oceanic westerlies of both hemispheres, although the magnitude of absolute IWP varies considerably. The models also depict, to a varying extent, large regions of low IWP, in relation to their total mean, in the tropical and subtropical Eastern Pacific and Indian Oceans. Notably, the ratio of IWP magnitudes between tropical regions and temperate regions varies greatly from model to model. Figure 6 (and Table 4, introduced later) indicate that it is at tropical latitudes that the models are in greatest disagreement. Compared to observed dynamical features seen in Fig. 5 (e.g. TWP, ITCZ), some models, such as GISS and INM, show large areas of relatively elevated IWP, misplaced well west of the TWP. Some models, such as GISS, present a region of relatively low IWP along the equator deep into the TWP, also diminishing the ITCZ. In an absolute IWP sense, the GISS model uniquely has higher IWP values than the observations. Some models have strong local maximum over Southern China, whilst other models completely lack this feature. The GISS model has extremely elevated values in this area (several orders of magnitude higher than its total absolute mean IWP). Such features are not seen in any of the satellite datasets. It is also clear that some models, such as UKMO, have very low IWP values in the Tropics, both in a relative and absolute sense. This indicates that the models are probably incorrect or the model data are incomplete here. Compared to observations, large-scale dynamical features of modelled IWP are seen to largely deviate from observations, even after attempting to remove the difference caused by precipitation, as done in the IWP$_{noPrecip}$ dataset.

3.3 Monthly mean IWP

Modelled monthly mean IWP only very roughly represents the amount of ice in the atmosphere, where small scale processes, such as convection, and processes dependent on the diurnal cycle are parametrised. Figure 7 shows the histogram distributions of CloudSat IWP (level 2, granule data) for March 2007. The figure on the left shows data from combined regions with equatorial continental convection in South America and Africa, called trop_cont (see Fig. 9). The figure on the right shows the global IWP distribution. The total mean (black), the cloudy mean (red) and the cloudy median (blue) IWP value for the month is shown as vertical lines in the figure.

For the given month, frequent convection in the trop_cont contributes to a mean IWP of 182 g/m^2 and for cloudy footprints, defined as IWP > 0 as seen by CloudSat, the average IWP is 420 g/m^2. Despite the highest IWP values in this region, the ratio of cloudy pixels is only ~44%. The median IWP value of the cloudy pixels is only 39 g/m^2. By comparison, the global mean of all CloudSat IWP measurements for the same month is 73 g/m^2, and 190 g/m^2 for the mean cloudy IWP. The median cloudy IWP value is 33 g/m^2, and the ratio of cloudy pixels is 38%.

From Fig. 7, it is easy to understand that the monthly mean IWP is generally far from the IWP expected for the most common atmospheric states. For the studied region, the common state is cloud free (CloudSat IWP = 0), or cloudy,
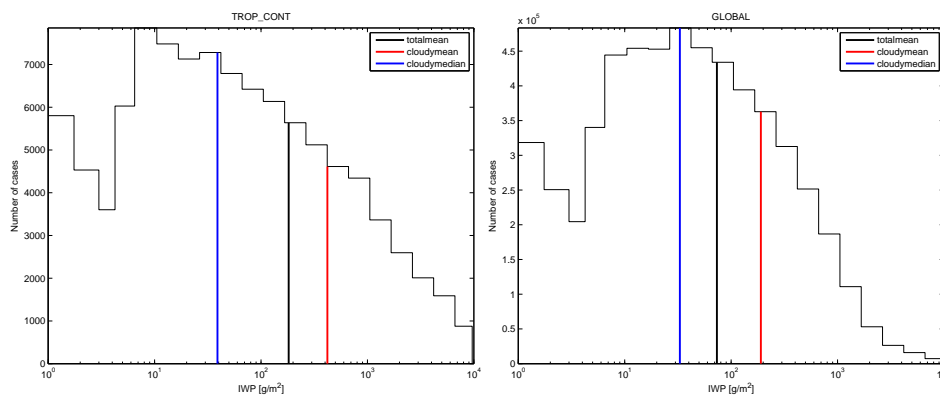
Fig. 7. Histogram of CloudSat IWP in a region of frequent strong convection (left) and the global IWP distribution (right). The region studied here is the combined region of equatorial continental convection in South America and Africa (see Fig. 9.). The IWP data (level 2, granule) are shown in intervals of $\log_{10\,(\mathrm{IWP})}$. The sample month chosen for both histograms is March 2007. The vertical lines represent: the total mean including cloud free footprints (black), the cloudy mean excluding cloud free (red), and the cloudy median (blue) for the given month.

with IWP less than $\frac{1}{4}$ of the month mean. As the distribution of IWP values is highly non-Gaussian, the use of Gaussian statistics, such as mean and standard deviation, is not entirely suitable for representing the monthly state of cloud properties, such as IWP. For example, for some physical process F (e.g. radiative forcing) that is a non-linear but monotonic function of x (e.g. IWP), the median value of x can be used to infer the median value of the F. The same cannot be said of the mean values owing to the non-linearity of F. Hence, estimating the monthly cloud radiative forcing from monthly mean IWP is inappropriate (e.g., Atlas et al., 1995). This appears to be especially serious for convective regions, where the mean and median cloud values have the largest differences. Despite this, mean IWP is the common statistic used by models and satellite datasets for representing the data on a monthly basis.

3.4 Total statistics

Table 4 contains the area-weighted mean IWP of the selected satellite and climate model datasets for the total region ($60°$ S–$60°$ N) and the tropical region ($30°$ S – $-30°$ N). These statistics are based on all monthly mean values in the given period and grid. This differs from the statistics provided in Table 3 for the spatial distribution of IWP expressed in percentiles, which are based on the average gridded IWP for the entire period. The satellite comparison is based on monthly mean values from 2007. The IWP$_{\mathrm{noPrecip}}$ dataset has not been included here as it is only valid over open ocean. As indicated, the satellite datasets are in better agreement, in terms of spatial and temporal correlation in the tropical region than in the total region. MODIS and PATMOS-x are in closest agreement with CloudSat and are within, or close to CloudSat's ±40% uncertainty interval in both regions. MODIS appears closer to CloudSat in the total region (which includes the tropics) than in the tropical region, whereas PATMOS-x

may agree slightly better with CloudSat in the tropics. However, PATMOS-x deviates strongly from the other datasets at latitudes greater than $60°$ N/S (not shown). MSPPS has a large negative IWP bias in both regions, largely due to its technical inability to detect most ice clouds. For detected clouds MSPPS generally reports a factor 5 less IWP than CloudSat (Holl et al., 2010). As previously shown in Fig. 2, ISCCP IWP is close to the averages of the IWP$_{\mathrm{noPrecip}}$ dataset, hence against CloudSat, ISCCP has a large negative bias, tending towards a larger negative bias in the tropics. A more detailed regional analysis of IWP averages is further presented in Sect. 3.6.

Monthly averages, based on model data from 1900–2000, were used for the model statistical comparison. Compared to the satellite data, the models, with the exception of GISS, clearly have much smaller IWP magnitudes, with averages ranging from 30–50% of the CloudSat average. As mentioned earlier, this is expected as model IWP pertains to the suspended cloud-ice portion of the IWP column only. The models are therefore in better agreement with the ocean only IWP$_{\mathrm{noPrecip}}$ dataset, which simulates CloudSat measurements without precipitation (but has a dry bias), in terms of magnitude. As the available datasets are not suitable for direct comparison to model data here, we have arbitrarily chosen ECHAM as a reference model for the subset of models. In terms of spatial-monthly correlation, the models appear in better agreement with each other in the total region, whereas in the tropical region, the models deviate strongly from one another. The models, except for INM, also have much higher IWP averages in the total region than in the tropical region. UKMO exhibits nearly a factor 4 higher IWP values in the total region, whereas INM has the same average for both regions. ECHAM, CCSM and CSIRO are generally well correlated in both regions, but CCSM has much lower IWP values. GISS has a very large IWP magnitude across all latitudes and clearly exhibits problems modelling this quantity.

In summary, the satellite datasets are in better agreement in the tropics than outside the tropics, especially in terms of spatial distribution of IWP. The climate models, on the other hand, appear to have the opposite relation.

3.5 IWP seasonal cycle

In this section, we provide a qualitative evaluation of the large-scale seasonal cycle of IWP in the tropics ($30°\text{S} - 30°\text{N}$), to check the model IWP variability compared to observations. As the standard deviation of IWP varies greatly in magnitude between datasets, we have normalised the seasonal cycle using the coefficient of variation (CV) (e.g., Mohapatra et al., 2007). CV is the variation of the average IWP for every month of the year (i.e. 12 values) normalised by the total mean of the period.

$$\text{CV}_{xy} = \frac{\sigma_{xy}}{\overline{\overline{I_{xy}}}}$$

where $\overline{\overline{I_{xy}}}$ is the mean IWP of all months per grid box and σ_{xy} is the standard deviation of the month averages per grid box.

Figure 8 shows the CV of observed and modelled IWP. PATMOS-x and ISCCP were chosen as observational data due to their long temporal ranges (25 and 24 years), long enough to resolve the inter-annual variability, which is present in the models. In the observed data, the spatial distribution of CV indicates broad areas around the TWP, which on a relative scale have little variability. This indicates semi-persistent high IWP conditions in this region. As expected, the yearly monsoonal activity, associated with the wandering ITCZ, gives rise to a large variability of monthly IWP near the Indian subcontinent, Northern Australia and Eastern Africa. Both PATMOS-x and ISCCP also indicate a region with a strong seasonal cycle in the Eastern Pacific. This feature is associated with an apparent "double ITCZ" that forms during March and April, which ice signature was shown in Mohr and Zipser (1996). Strong CV signals can also be seen in areas with low absolute IWP averages, which may have a short period of relatively large IWP values. This is because CV emphasises regions with low IWP.

The models diverge in regions near, or periodically near the ITCZ. The seasonal cycle of ECHAM closest resembles observations, followed by CSIRO. GISS completely lacks the area of low seasonal cycle associated with the TWP and its adjacent areas, indicating too large seasonal variability here. Most models have too small regions of low CV east and west of the TWP, where observations indicate semi-persistent high IWP. In general, the seasonal cycle of the models in the important high IWP regions are not well represented in the models. They also diverge in regions with low IWP, but these are finer details, and this test is too stringent to assess the models in these regions.

3.6 Statistics by region

Waliser et al. (2009) compared the average global, tropical ($30°\text{S}–30°\text{N}$) and extra-tropical ($> 30°\text{N,S}$) IWP for both observational datasets and GCMs. They found these averages were approximately the same for each satellite dataset, and for the models, they have a factor 2 more IWP in the extra-tropics. We have chosen to go a step further and analyse the regional averages of satellite and model data in regions that, over all seasons, contain quasi- homogeneous atmospheric conditions. We have defined 4 large regions, each with a climate dominated by a large scale vertical motion of the atmosphere, either subsiding or rising.

We have selected two large subsidence zones west of South America and west of South Africa, and these areas combined constitute the oceanic subsidence region (ocean_sub). Tropical continental South America and Africa together constitute the tropical continental convection region (trop_cont). The TWP and areas adjacent constitute the Tropical Oceanic Convection region (warm_pool). The areas of mid-latitude storms in the North Atlantic, North-West Pacific, and the mid-latitude storm tracks in the southern hemisphere, together constitute a region called westerlies. The boundaries of these regions are based on the spatial distribution of IWP of all satellite datasets, and the largest possible regions with similar IWP distributions were selected. The extent of these regions are shown in Fig. 9.

The area-weighted mean and the regional fraction of IWP was calculated for satellite datasets and models for the chosen regions and are shown in Table 5. The regional fraction in percent (IWP_{rel}) is defined:

$$\text{IWP}_{\text{rel}} = \frac{\text{IWP}_{\text{reg}}}{\text{IWP}_{\text{tot}}} \cdot 100\% \tag{1}$$

where IWP_{reg} is the mean IWP for a given region and IWP_{tot} is mean IWP from the total region ($60°\text{S}–60°\text{N}$) for each dataset. This quantity was included in Table 5 for clarity, as the difference in the mean IWP between datasets is large.

3.6.1 Regional observations

Table 5 shows the regional statistics of the observed datasets and climate models. Results show that all satellite datasets are in general agreement in the oceanic subsidence regions in terms of relative IWP, with PATMOS-x and ISCCP having a higher fraction of IWP in this region. PATMOS-x has comparatively high absolute IWP averages in this region ($3\times$ larger than CloudSat). It can detect thinner ice clouds than CloudSat, such as clouds induced from gravity waves or remnant outflow from distant convection. However, this is may not be the cause of PATMOS-x's high relative values in these regions because MODIS, based on similar spectral channels, does not show the same result. In the westerlies region, ISCCP has the highest relative fraction of the total
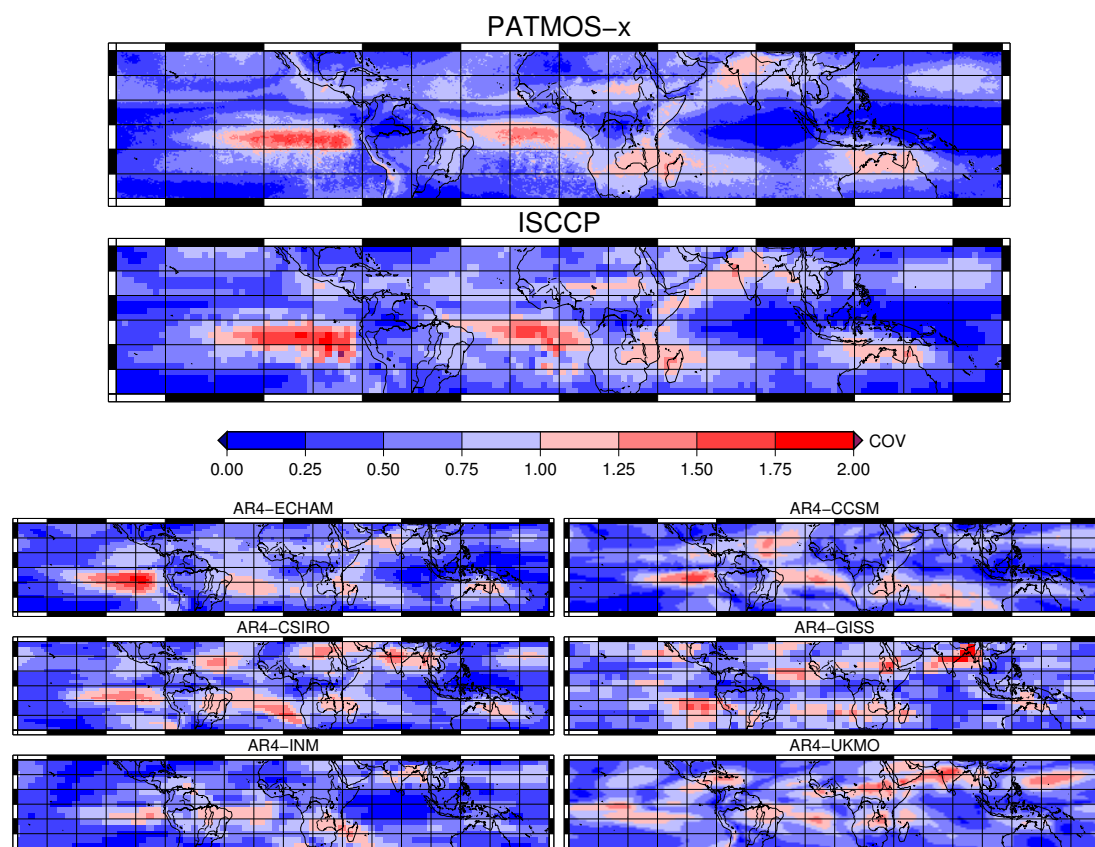
Fig. 8. Coefficient of variation, depicting the seasonal cycle of IWP in the tropics (30°S–30°N). The top two panels are the datasets, PATMOS-x and ISCCP, both of which have sufficiently long temporal ranges to capture the inter-annual variability of IWP. The bottom six panels are the subset of models used in this survey.
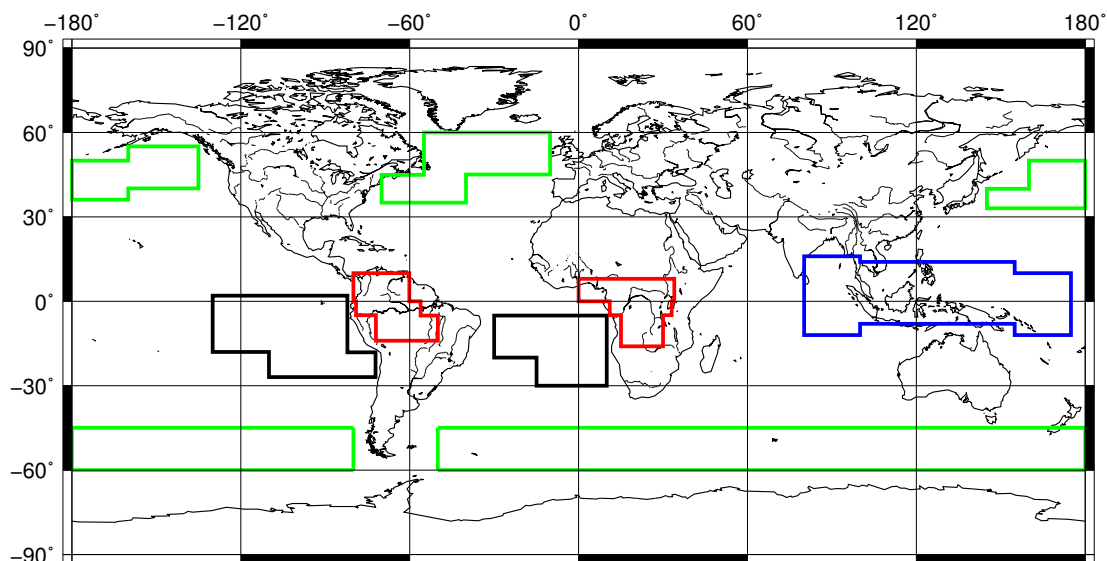

Fig. 9. Schematic of regions chosen for statistical comparison. Tropical Continental zones (red), Tropical maritime convective zone (blue), subsidence zones (black) and Westerlies (green).

average (1.9), PATMOS-x and MODIS have the same fraction (1.6), CloudSat (1.4). MSPPS has much too low IWP averages in this region both in an absolute and relative sense (0.4). As expected, the reverse is true for the warm_pool region. MSPPS and CloudSat have the highest fractions of 2.6 and 2.3 respectively, followed by PATMOS-x and MODIS with 1.9 and 1.8 respectively, and ISCCP with 1.5, has the lowest fraction. The average IWP in the trop_cont region is approximately the same fraction as for the warm_pool region, albeit a little less for all datasets, with the strong exception of MSPPS. MSPPS average IWP for the tropical continental region is nearly a factor five higher than its total average. This is likely an indication of the deep convection which is generally more intense and more frequent over land than over ocean (Hong et al., 2005). The ocean-only IWP$_{noPrecip}$ dataset has in general a smaller fraction of IWP in the tropics and a larger fraction in the westerlies, than the other datasets. As mentioned earlier, ISCCP and IWP$_{noPrecip}$ mainly represent suspended cloud-ice and CloudSat and MSPPS mainly represent larger particles. In light of this, it appears that clouds may have a higher proportion of precipitating cloud particles to small cloud-ice particles in the tropical convection regions compared to that of the westerlies regions.

Problems associated with comparing datasets in regions where there are large diurnal variations of IWP, such as tropical convective regions, are assumed to be small for CloudSat, MODIS, and MSPPS. The CloudSat and Aqua (MODIS) platforms are in the A-train and the NOAA18 (MSPPS) platform is close to the A-train, and therefore fairly collocated. The local equatorial passing times for the A-train are around 13:45 and 01:45. ISCCP, based on geostationary satellites and PATMOS-x, based on composite data from several sun-synchronous satellites, may deviate from the other datasets due to a different diurnal sampling. The regional comparison of the observations further confirm the general statements made on the differences between satellite datasets presented in Waliser et al. (2009).

The models have the same IWP values as the observations in the oceanic subsidence regions. As seen in Fig 4 (right) the cloud-ice column ratio is likely to be high in the subsidence regions indicating that the models are may be close to agreement here. However in terms of IWP$_{rel}$, the models have too much IWP here in comparison to their mean global IWP. All models except UKMO, have a relative IWP of between 27%–42%, whereas UKMO has a relative IWP value of 7%. The models have similar to observed absolute IWP values in the Westerlies regions, indicating that, in relation to the total mean IWP, all models tend to overestimate IWP in these regions. The exceptions are GISS and UKMO. GISS has very large absolute IWP values and UKMO has a relative IWP a factor three larger than its total average in the westerlies region. In the convective regions all of the models have particularly low absolute, but also relative IWP compared to the satellite datasets. This is especially pronounced in the tropical continental regions. UKMO stands out with very low

IWP in these regions, merely around 30% of its total mean IWP. The low modelled IWP averages in this region are at least in part due to the low cloud-fraction, but in addition it may also indicate insufficient moist convection.

It has previously been shown that climate models have problems modelling the diurnal cycle, of which indicates problems with convection (e.g., Dai and Trenberth, 2004; Eriksson et al., 2010). The study in Eriksson et al. (2010) showed that the models they studied captured the diurnal cycle to a varying degree. Their results in conjunction with the results presented in this study indicate that a model which is good at one aspect of the cloud ice (such as IWP) appears to also be good at other aspects (such as diurnal phase). Additionally, the model regional results may reflect the better availability of surface and atmospheric data at mid-latitudes, which enables better parametrisations and tuning for the mid-latitudes. The tropics are largely made up of oceans and monitored by a very sparse network of measurement sites, leading to more uncertain results.

As mentioned earlier, Waliser et al. (2009) showed tropical, extra-tropical and global averages for all models and datasets. For example, by attaining a tropical average (30° S–30° N) by latitude only, large areas with high IWP, such as trop_cont and warm_pool are averaged with large dry regions, such as ocean_sub along with large deserts (e.g Sahara). Whereas, the average IWP of the extra-tropics is attained from regions with, in general, less extreme IWP differences. This regional comparison indicates that additional important information can be attained by comparing datasets and models on a regional basis, such as done here. Our results show for CloudSat, the IWP magnitude of the tropical "wet" regions is approximately a factor 2 higher than the extra-tropical "wet" regions. The GCMs show IWP associated with the westerlies which are generally twice as high as the total average, in accordance with Waliser et al. (2009). This survey showed additionally that modelled IWP in tropical continental regions is markedly lower than over tropical maritime convective regions, in accordance with ice clouds associated with convection over land probably having smaller cloud-ice column ratios by comparison. Despite the large uncertainties for determining which models are closest to observations, judging in terms of absolute and relative IWP values and spatial distribution of large dynamical features, explained in Sect. 3.2, we believe ECHAM followed by CSIRO are closest to observations.

In this sense UKMO and GISS are in largest disagreement with observations, in that it UKMO has much too low tropical IWP values and has quite high values, and in relation to its total mean, very high values in the westerlies. GISS has a very high absolute IWP bias compared to all datasets at all latitudes, and in the placement of key dynamical features of IWP. Most likely the IWP data provided to the AR4 archive from these two models have serious problems, the causes of which are outside the scope of this article.

Table 5. Area-weighted mean IWP [g/m^2] for selected regions. The regional fraction IWP$_{rel}$ (in %), follows the regional mean in brackets. The mean IWP for the total region has been included in the table as a reference (far right). All available data have been used to retrieve the mean values.

Datasets	ocean_sub	trop_cont	warm_pool	westerlies	total
CloudSat	5.8 (7)	156.1 (202)	176.9 (229)	106.8 (138)	77.2
IWP$_{noPrecip}$	2.4 (9)		50.7 (195)*	38.3 (148)	25.9
ISCCP	3.9 (12)	42.6 (135)	46.5 (147)	59.3 (187)	31.7
PATMOS-x	15.0 (14)	182.8 (176)	197.3 (190)	166.3 (160)	104.1
MODIS-aqua	4.3 (8)	95.6 (165)	104.7 (181)	91.9 (159)	57.9
MSPPS-noaa18	0.6 (5)	58.9 (481)	30.6 (249)	4.7 (39)	12.2
AR4-ECHAM	9.1 (27)	39.3 (117)	43.5 (130)	61.9 (184)	33.6
AR4-CCSM	4.8 (26)	14.9 (81)	22.9 (125)	36.5 (198)	18.4
AR4-CSIRO	7.2 (17)	26.0 (62)	52.1 (124)	98.1 (234)	41.9
AR4-GISS	57.0 (27)	183.7 (88)	93.1 (44)	433.4 (207)	209.3
AR4-INM	4.6 (43)	9.8 (91)	16.0 (148)	15.0 (138)	10.9
AR4-UKMO	2.7 (7)	10.2 (26)	12.8 (32)	118.2 (300)	39.4

* IWP$_{noPrecip}$ is an ocean only dataset which effects its regional averages in the trop_cont and warm_pool regions.

4 Discussion and summary

From a subset of 6 models out of 20 in the AR4 archive, large discrepancies between modelled IWP distributions are apparent. The models, which were selected in a way that their the inter-model variability is close to that of all models in the archive, reflect the difficulties faced in modelling this quantity. A major contributing factor to this is the lack of detailed cloud ice measurements to constrain the models with. Although satellite datasets of IWP now span 25 years, no satellite datasets provide consistent and accurate IWP retrievals for the entire cloud ice column. Rather, depending on the remote sensing technique used, different portions of the column are retrieved. The introduction of the CloudSat Profiling Radar is a leap forward for the modelling community in addressing the problems in cloud ice simulation. In contrast to the IWP data from passive sensors, CloudSat provides information on the vertical structure of clouds, which is essential information in modelling fields such as IWC.

Waliser et al. (2009) used CloudSat as the main reference of IWP, and compared the distribution of modelled IWC in two prognostic GCMs. They showed large differences in vertical distribution of IWC between the models and Cloud-Sat. CloudSat IWP represents the total mass of ice, including graupel and snow, which are not included in model IWP. As discussed in detail in Waliser et al. (2009), the distinction between different ice categories is natural and simple from a model point of view, but hard or even impossible from a measurement point of view. The only reasonably well defined quantity to measure is total IWP. Therefore, to aid in the model comparison to observations we added a dataset, called IWP$_{noPrecip}$, based on CloudSat measurements where profiles containing surface precipitation were removed. This

dataset represents the lower end of the uncertainty as it naturally has a dry bias because all clouds strong enough to support precipitation are removed from the total.

The satellite datasets have very different magnitudes and distributions of IWP, due largely to the different retrieval techniques used. MODIS is in best agreement with CloudSat over all latitudes, with a negative bias. PATMOS-x is also relatively close to CloudSat within 60° S–60° N, but with a positive bias compared to CloudSat. ISCCP has a considerable low bias compared to CloudSat but is in good agreement after removing the precipitation, i.e. with IWP$_{noPrecip}$.

Also in Sect. 3.1 and 3.2, it is shown that the climate models disagree on the magnitude and spatial distribution of IWP, which confirms the results of Waliser et al. (2009). As the modelled IWP and observed IWP are not defined equally, not much can be said about the absolute IWP of the models. However, it is shown that the models are in larger disagreement to observations in regions with persistent convection.

The regional comparison study in Sect. 3.6 showed that in relation to the observed mean IWP in 60° S–60° N, many models have particularly low IWP ratios in the tropical regions compared to the westerly storm-track regions. The GISS model is the exception, it has much higher absolute IWP values than seen in satellite observations across all latitudes, depicting difficulties modelling this quantity. In this comparison ECHAM and CSIRO were the two models which were closest to the "observed" data in terms of absolute magnitude and spatial correlation. There may be some merit for the models having higher relative values outside the tropics than inside as model IWP is suspended cloud ice only, and the simulated suspended cloud ice to IWP fraction relationship shown in Fig. 4 suggest that cloud ice fractions

are indeed higher outside the tropics. This relationship was first indicated by zonal averages using two models in Waliser et al. (2009).

As this study concerns the atmospheric state monthly mean, we assessed the distribution of CloudSat level 2, granule IWP data for one month. Section 3.3, showed that mean for March 2007, is the mean of very different atmospheric states, and its distribution is highly non-Gaussian. This is also most apparent in convective regions.

For the climatology of IWP to be modelled correctly, they must model the large scale processes that effect the seasonal variability of IWP. It is expected that there is a high seasonal variability of IWP in certain regions of the tropics such as the Indian sub-continent and Northern Australia, and a low variability in regions such as near the equator over the Indonesian archipelago. Using long term datasets of IWP from PATMOS-x and ISCCP, we assessed the coefficient of variation, a measure of the seasonal variability, of the gridded monthly averages of IWP in the models. By using datasets with long time series, the inter-annual variability is averaged in the model and observations statistics. Figure 8) in Sect. 3.5 shows that most models differ from the observed data in terms of where and how strong their seasonal variability of IWP is, especially in and around the ITCZ region. This may indicate that important seasonal features such as the Indian monsoon may not be simulated satisfactorily.

5   Conclusions

Although several IWP observational datasets exist, they are more or less inherently different from one another, as they retrieve information from different parts of the cloud column. This is not purely a limitation. With the knowledge on the observational datasets' limitations and sensitivities, they may be used in combination to enhance our knowledge on the total IWP column. We have evaluated five IWP datasets by comparing them to CloudSat in their period of overlap, in terms of distribution and magnitude of monthly mean IWP. ISCCP, one of the longest datasets, has low IWP values by comparison and has particularly low values in the tropics. MSPPS has the lowest IWP values of the datasets. This dataset, based on passive microwave measurements has very low values in regions outside the tropics, but tends to detect IWP in deep convective clouds. MODIS and PATMOS-x are closest to CloudSat, with MODIS being in better agreement. PATMOS-x has a temporal range of 25 years compared to MODIS approximately 10 years. Especially for the model to observations comparisons, we attempted to remove the precipitation bias in CloudSat. Using a ground precipitation flag, we created the IWP$_{noPrecip}$ dataset. ISCCP IWP is notably in good agreement with this dataset. It is well documented that it is difficult to compare retrieved IWP to modelled IWP. By utilising long term datasets such as PATMOS-x and ISCCP, the climate model community has 25 years of

data with which to constrain their IWP distributions. We compared the magnitude and distribution of modelled IWP of a subset of models from the IPCC AR4 to observed IWP distributions. Our results are consistent with John and Soden (2006) and Waliser et al. (2009) that there are large discrepancies between the climate models used in AR4. We showed that all models appear to have problems modelling the spatial distribution of IWP. This is especially true in tropical convective regions.

In future studies, we would like to enhance our understanding of the differences between the observational datasets. This should be done on a pixel level on coincidental data, rather than using monthly means, or preferably using satellite simulators to better understand the observations. We believe that utilising the long satellite datasets of IWP and understanding how the datasets complement each other is the key to constraining model IWP distributions. Future satellite sensors that measure IWP more directly than current sensors, such as the one proposed in Buehler et al. (2007), would also be desirable.

References

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear-sky from clouds with modis, J. Geophys. Res., 103(D24), 32,141–32,157, 1998.

Atkinson, N. C. (2001), Calibration, monitoring and validation of AMSU-B, Adv. Space. Res., 28(1), 117–126.

Atlas, D., Matrosov, S. Y., Heymsfield, A. J., Chou, M.-D., and Wolff, D. B.: Radar and radiation properties of ice clouds, J. Appl. Meteorol., 34, 2329–2345, 1995.

Austin, R. T., Heymsfield, A. J., and Stephens, G. L.: Retrievals of ice cloud microphysical parameters using the Cloud-Sat millimeter-wave radar and temperature, J. Geophys. Res., 114, D00A23, doi:10.1029/2008JD010049, 2009.

Buehler, S. A., Jiménez, C., Evans, K. F., Eriksson, P., Rydberg, B., Heymsfield, A. J., Stubenrauch, C., Lohmann, U., Emde, C., John, V. O., Sreerekha, T. R., and Davis, C. P.: 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, 2007.

Dai, A. and Trenberth, K. E.: The diurnal cycle and its depiction in the community climate system model, J. Climate, 17, 930–951, 2004.

Eriksson, P., Ekström, M., Rydberg, B., Wu, D. L., Austin, R. T., and Murtagh, D. P.: Comparison between early Odin-SMR, Aura MLS and CloudSat retrievals of cloud ice mass in the upper tropical troposhere, Atmos. Chem. Phys., 8(7), 1937–1948, doi: 10.5194/acp-8-1937-2008, 2008.

Eriksson, P., Rydberg, B., Johnston, M., Murtagh, D. P., Struthers, H., Ferrachat, S., and Lohmann, U.: Diurnal variations of humidity and ice water content in the tropical upper troposphere, Atmos. Chem. Phys., 23, 11,519–11,533, doi:10.5194/acp-10-11519-2010, 2010.

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

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

Heymsfield, A. J. and Iaquinta, J.: Cirrus crystal terminal velocities, J. Atmos. Sci., 57, 916–938, 2000.

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

Heymsfield, A. J., Protat, A., Austin, R., Bouniol, D., Hogan, R., Delanoë, J., Okamoto, H., Sato, K., van Zadelhoff, G.-J., Donovan, D., and Wang, Z.: 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, 2008.

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

Hong, G., Heygster, G., and Kunzi, K.: Intercomparison of deep convective cloud fractions from passive infrared and microwave radiance measurements, IEEE Geosci. R. S. Le., 2, 18–24, doi: 10.1109/LGRS.2004.838405, 2005.

Intergovernmental Panel on Climate Change: Fourth Assessment Report: Climate Change 2007: The AR4 Synthesis Report, Geneva: IPCC, 2007.

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

King, M. D., Tsay, S.-C., Platnick, S. E., Wang, M., and Liou, K.-N.: Cloud retrieval algorithms for modis: Optical thickness, effective particle radius, and thermodynamic phase, Tech. rep., MODIS Science Team, mODIS Algorithm Theoretical Basis Document No. ATBD-MOD-05, 1997.

Mohapatra, G. N., Panda, U. S., and Mohanty, P. K.: Annual cycle of surface meteorological and solar energy parameters over orissa, Indian J. Radio Space, 36, 128–144, 2007.

Mohr, K. I. and Zipser, E. J.: Defining mesoscale convective systems by their 85-GHz ice-scattering signatures, Bull. Amer. Met. Soc., 77(6), 1179–1189, 1996.

Ramanathan, V., Cess, R. D., Harrison, E. F., Minnis, P., Barkstrom, B. R., Ahmad, E., and Hartmann, D.: Cloud-radiative forcing and climate: Results from the earth radiation budget experiment, Science, 243, 57–63, 1989.

Rossow, W. B. and Garder, L.: Selection of map grid for data analysis and archival, J. Clim. Appl. Mete., 23, 1253–1257, 1984.

Rossow, W. B. and Schiffer, R. A.: ISCCP cloud data products, Bull. Amer. Met. Soc., 72, 2–20, 1991.

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'Connor, E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., Mitrescu, C., and the CloudSat Science Team: The cloudsat mission and the A-train, Bull. Amer. Met. Soc., 83, 1771–1790, 2002.

Storelvmo, T., Kristjánsson, J. E., and Lohmann, U.: Aerosol influence on mixed-phase clouds in cam-oslo, J. Atmos. Sci., 65, 3214–3230, doi:10.1175/2008JAS2430.1, 2008.

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

Wu, D. L., Austin, R. T., Deng, M., Durden, S. L., Heymsfield, A. J., Jiang, J. H., Lambert, A., Li, J.-L., Livesey, N. J., McFarquhar, G. M., Pittman, J. V., Stephens, G. L., Tanelli, S., Vane, D. G., and Waliser, D. E.: Comparisons of global cloud ice from MLS, CloudSat, and correlative data sets, J. Geophys. Res., 114, D00A24, doi:10.1029/2008JD009946, 2009.