Surveys in Geophysics (2024) 45:1999–2048
https://doi.org/10.1007/s10712-024-09824-0

# Lessons Learned from the Updated GEWEX Cloud Assessment Database

**Claudia J. Stubenrauch**[1] · **Stefan Kinne**[2] · **Giulio Mandorli**[1] · **William B. Rossow**[3] · **David M. Winker**[4] · **Steven A. Ackerman**[5] · **Helene Chepfer**[1] · **Larry Di Girolamo**[6] · **Anne Garnier**[4,7] · **Andrew Heidinger**[8] · **Karl-Göran Karlsson**[9] · **Kerry Meyer**[10] · **Patrick Minnis**[4,7] · **Steven Platnick**[10] · **Martin Stengel**[11] · **Szedung Sun-Mack**[4,7] · **Paolo Veglio**[5] · **Andi Walther**[5] · **Xia Cai**[7] · **Alisa H. Young**[12] · **Guangyu Zhao**[6]

Received: 25 September 2023 / Accepted: 8 January 2024 / Published online: 29 February 2024

## Abstract

Since the first Global Energy and Water Exchanges cloud assessment a decade ago, existing cloud property retrievals have been revised and new retrievals have been developed. The new global long-term cloud datasets show, in general, similar results to those of the previous assessment. A notable exception is the reduced cloud amount provided by the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) Science Team, resulting from an improved aerosol–cloud distinction. Height, opacity and thermodynamic phase determine the radiative effect of clouds. Their distributions as well as relative occurrences of cloud types distinguished by height and optical depth are discussed. The similar results of the two assessments indicate that further improvement, in particular on vertical cloud layering, can only be achieved by combining complementary information. We suggest such combination methods to estimate the amount of all clouds within the atmospheric column, including those hidden by clouds aloft. The results compare well with those from CloudSat-CALIPSO radar–lidar geometrical profiles as well as with results from the International Satellite Cloud Climatology Project (ISCCP) corrected by the cloud vertical layer model, which is used for the computation of the ISCCP-derived radiative fluxes. Furthermore, we highlight studies on cloud monitoring using the information from the histograms of the database and give guidelines for: (1) the use of satellite-retrieved cloud properties in climate studies and climate model evaluation and (2) improved retrieval strategies.

**Keywords** Satellite remote sensing · Atmosphere · Cloud cover · Cloud height · Cloud radiative properties · Microphysical properties · Cloud types

---

Extended author information available on the last page of the article

## Article Highlights

- Since the first Global Energy and Water Exchanges (GEWEX) cloud assessment a decade ago, existing cloud property retrievals have been revised and new retrievals have been developed. Eleven climate data records of cloud properties are investigated, and a common database is made available
- While the spread in the absolute values between the datasets is still large, their latitudinal and seasonal variations agree well, except at the polar latitudes
- Though it is very challenging to detect small inter-annual global and zonal changes, similar geographical change patterns between the datasets demonstrate that the data can be used to study climatological change patterns, in particular those for specific cloud types
- The similar results of the two assessments indicate that further improvement, in particular on vertical cloud layering, needs combining complementary information

## 1 Introduction

Clouds are important regulators of the energy and water cycle. While surface observations provide the morphological structure of the cloud types, satellite observations make it possible to get an overview of the synoptic cloud fields (Fig. 1). Furthermore, the satellite observation era, beginning in the 1970s and spanning more than 40 years, allows us to monitor physical cloud properties over the entire globe. However, it is challenging to evaluate cloud properties on a global scale, as no unique dataset offers the ultimate truth, due to associated uncertainties.

The Global Energy and Water Exchanges (GEWEX) Cloud Assessment (Stubenrauch et al. 2013) provided the first coordinated inter-comparison of publicly available, global cloud products (gridded, monthly statistics). In addition to self-assessments (Annex I of Stubenrauch et al. 2012), indicating the maturity of the datasets, this inter-comparison has

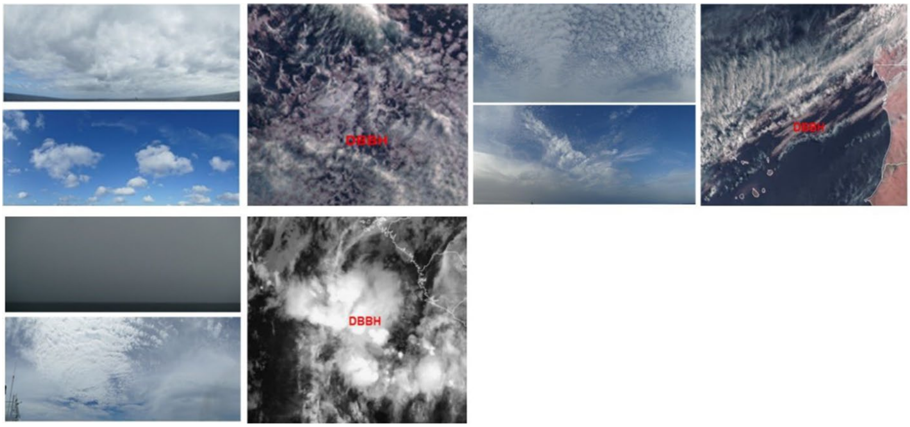


**Fig. 1** Cloud types from surface observation (morning and afternoon) and synoptic cloud fields from Meteosat satellite images, for specific days in May 2023 during the science cruise M189/2 EARS. Examples show: stratocumulus fields off the coast of Namibia (top left), Cirrus and low-level clouds off the Sahara (top right) and convective cells and anvils in the ITCZ over the Atlantic (bottom). Position of the German research vessel Meteor, at 7:15 UTC:DBBH

shown how cloud property averages and distributions are affected by instrument characteristics and retrieval methods. The database associated with the GEWEX Cloud Assessment has revealed its usefulness for the assessment of new datasets. Since then, several teams have revised their retrievals, and new global long-term datasets have emerged. The updated cloud database contains data until 2020, in the same format as the original GEWEX Cloud Assessment database, with monthly statistics (averages, variability and histograms) per cloud property, at a spatial resolution of 1° latitude × 1° longitude. Since cloud property statistics are stored in files per year, we foresee a yearly extension of this database, if the cloud teams are able to provide the files.

These satellite data are very valuable for climate studies and for the evaluation of climate models, *if one is aware of the limits specific to the applied dataset*. In the following, we provide insight on how and why the retrieved cloud properties of these datasets differ. In this overview, we concentrate on cloud amount, cloud height as well as cloud radiative and microphysical properties. For the inter-comparison of averages and spatial distributions, we used the common 8-year period of 2008–2015.

After a short description of the different sensors, retrieval methods and the database in Sect. 2, we compare cloud amounts and cloud properties of the individual datasets in Sect. 3. This section also presents statistics of specific cloud types from two-dimensional histograms in the space of near-cloud-top pressure and visible cloud optical depth. As passive remote sensing only provides information on the uppermost cloud layers, Sect. 4 illustrates how complementary information can be used to estimate the full amount of all cloud layers, including the amount hidden under clouds aloft. Section 5 discusses the monitoring of cloud type occurrences with respect to time and to global surface temperature. A summary and guidelines are given in Sect. 6.

## 2 Sensors, Retrieval Methods and Database

The updated GEWEX cloud assessment database includes monthly cloud property statistics of eleven cloud climatologies, based on measurements from five different sensor types:

- *multi-spectral imagers*, with a maximum of 5 channels between visible (VIS) and infrared (IR), covering 0.5 µm to 12 µm, and a footprint size of about 4 km (Sect. 2.1),
- *advanced multi-spectral imagers*, like the Moderate Resolution Imaging Spectroradiometer (MODIS) with 36 channels between VIS (0.4 µm) and IR (14.5 µm) and a nadir footprint size from 250 m to 1 km (Sect. 2.2),
- *IR sounders*, with channels along the 15 µm $CO_2$ absorption band and a footprint size of about 13 km (Sect. 2.3)
- *multi-angle solar-spectral imager*, with 4 solar-spectral channels and a footprint size of about 1 km (Sect. 2.4)
- *active lidar*, along nadir tracks with a footprint size of about 90 m (Sect. 2.5)

The longest time series are available since the 1980s, from instruments aboard meteorological polar orbiting and geostationary satellites. Instruments operating since the 2000s have increased retrieval capabilities, in particular the spaceborne lidar and radar missions CALIPSO and CloudSat launched in 2006 (Stephens et al. 2018a, b).

All datasets of this assessment database, except ISCCP, are built from measurements aboard polar orbiting satellites. Table 1 recapitulates the datasets of this assessment

**Table 1** Participating datasets, type of sensors, nominal local observation times and time period in the updated GEWEX Cloud Assessment Database

| | | | | |
|---|---|---|---|---|
| ISCCP | Version H | Multi-spectral imagers | 1:30 AM/PM | 1984–2015 |
| Young et al. (2018) and Rossow et al. (2022) | | | 7:30 AM/PM | 1984–2015 |
| AVHRR PATMOS-x | Version 5 | Multi-spectral imagers | 1:30AM/PM | 1983–2020 |
| Heidinger et al. (2012), Walther and Heidinger (2012) | | | 7:30 AM/PM | 1992–2020 |
| AVHRR-CLARA | Version 2 | Multi-spectral imagers | 1:30 AM/PM | 1982–2015 |
| Karlsson et al. (2017) | | | 7:30 AM/PM | 1991–2015 |
| AVHRR-Cloud_cci | Version 3 | Multi-spectral imagers | 1:30 AM/PM | 1982–2016 |
| Stengel et al. 2020 | | | 7:30 AM/PM | 1992–2016 |
| MODIS CERES Science Team | Edition 4 | Advanced multi-spectral imagers | 1:30 AM/PM | 2000–2020 |
| Trepte et al. (2019), Minnis et al. (2021) | | | 10:30 AM | 2002–2020 |
| MODIS CLDPROP | Continuity version | Advanced multi-spectral imagers | 1:30 AM/PM | 2002–2020 |
| Frey et al. (2020); Platnick et al. (2021) | | | | |
| AIRS-CIRS, IASI-CIRS | Version 2 | Advanced IR sounders | 1:30 AM/PM | 2003–2018 |
| Stubenrauch et al. (2017) and Guignard et al. 2012 | | | 9:30 AM/PM | 2008–2018 |
| MISR | Version 2 | Multi-angle SW imager | 10:30 AM/PM | 2000–2017 |
| Di Girolamo et al. (2010) | | | | |
| CALIPSO-GOCCP | Version 2.9 | Lidar | 1:30 AM/PM | 2007–2017 |
| Chepfer et al. (2010) | | | | |
| CALIPSO Science Team | Version 4.2 | Lidar | 1:30 AM/PM | 2007–2016 |
| Liu et al. (2019) | | | | |
| IIR CALIPSO Science Team | Version 4.2 | Lidar/IR imager | 1:30 AM/PM | 2007–2016 |
| Ganier et al. (2021) | | | | |

database, with their nominal local observation times and temporal periods. The original ISCCP data, which provide cloud properties at eight specific universal times, have been processed for this database in such a way that the observation times are close to 1:30 and 7:30 local time (LT), AM and PM. Thus, with maximal four measurements per day, the analyses in this article are only able to partially resolve the diurnal cycle of the cloud properties. When presenting the optical and microphysical cloud properties as well as specific cloud types and their monitoring in Sects. 3.4 to 5, we only concentrate on analyses at one specific time of the day: 1:30 PM LT, because VIS optical depth is not available at night. This means that our analysis in the latter sections provides a snapshot and does not consider systematic diurnal variations of the specific cloud types (cf. Rossow and Schiffer 1991, 1999; Eastman et al. 2011; Eastman and Warren 2014; Rossow et al. 2022).

Furthermore, the National Oceanic and Atmospheric Administration (NOAA) satellites drifted by several hours during their life time (e.g. Foster et al. 2023) and the morning orbit time changed with advent of NOAA-17 in 2002 from 7:30 LT to 10:00 LT, which perturbs the values of the cloud properties retrieved from Advanced Very High Resolution Radiometer (AVHRR) measurements according to their diurnal variation and therefore may introduce biases in the time series (see Sect. 5).

Sections 2.1–2.7 shortly describe the different datasets. Table 2 summarizes their retrieval characteristics. Section 2.8 describes the database and gives some insight on the interpretation of the satellite-retrieved cloud properties. Not all datasets provide the complete set of variables shown in Table 3. An appendix of all acronyms is given in the supplement.

## 2.1 Long-Term Cloud Data from Multi-spectral Imagers

The *International Satellite Cloud Climatology Project* (***ISCCP***) uses a combination of polar orbiting and geostationary satellites in order to get a three-hourly diurnal sampling. ISCCP has recently released a new dataset (*Version H*, Rossow 2017, Young et al. 2018, Rossow et al. 2022), with a better spatial sampling (10 km instead of 30 km) and more up-to-date ancillary products (topography, land, water, snow and ice mask, atmospheric temperature-humidity profiles, ozone and aerosols). Night-time cloud properties are adjusted, based on the daytime differences between VIS/IR and IR-only results and the time interpolation of the VIS retrievals over the night-time. Changes in the polar regions are due to the removal of near-IR radiances in the retrieval over ice.

The version of ISCCP-H provided for this assessment is not the standard version as the monthly averages at four specific local times have been built from the ISCCP-HGG data, at 1° spatial resolution, given at eight specific universal times per day in such a way that the results of the two closest times were averaged. In the polar regions during periods without sunlight, only the IR-based results are available in the full ISCCP products, but in the GEWEX database, the height-stratified cloud amounts from VIS/IR analysis are not reported. Furthermore, only the 2-dimensional COD-CP and CEM-CP histograms are constructed from the initial 10 km pixels, while the 1-dimensional histograms are constructed from the 1° averages.

Measurements by the Advanced Very High Resolution Radiometers (AVHRR) aboard the NOAA and EUMETSAT polar orbiting satellites, integrated by ISCCP, have also been used on their own to produce cloud products by three different retrieval methods. The latter, listed below, are taking full advantage of all five channels and are based on very similar calibrated radiances of all available AVHRR instruments, on both

**Table 2** Summary of Cloud Property Retrieval Characteristics of the Datasets Participating in the updated GEWEX Cloud Assessment Database. In bold are the cloud height parameters directly retrieved. The 3.7 micron channel has been replaced by another channel during the daylight morning period (see Sect. 2.1)

| Dataset | Spatial resolution | Cloud detection | Variables | Retrieval method | Ancillary input |
|---|---|---|---|---|---|
| ISCCP-H | 5 km,10 km (sampled)<br>50–75 km clear sky estimation | 1 VIS 1 IR window time–space variances | COD, **CT** → CP, CZ | TB(11 µm) → CT, VIS → COD, CT correction for COD < 15 | NNHIRS retrieved *T/RH* profiles, ozone, aerosols,<br>rad transfer + cloud layer/particle model, surface properties<br>(including snow/ice, topography) |
| | | | CEM | (1-exp(-COD/const(W/I))) | |
| | | | Phase (W/I) | ice: CT < 253K | |
| | | | CWP | fct(COD,phase,fixed CRE) | |
| PATMOS-x | 1 km × 4 km Sampling distance: 3 km along, 5 km across track | 6 Bayesian classifiers derived from CALIPSO | CEM, **CT** → CP | Optimal estimation (11,12 µm) | NCEP reanalysis profiles (V1) *MODIS snow mask, rad. transfer + particle model (mixed habits tor ice)* |
| | | | Phase (W/I) | Spectral differences | |
| | | | COD, CRE | LUT approach (0.6, 3.7 µm) | |
| | | | CWP | CLWP = fct(CODW, CREW);<br>CWP = fct(CODI) | |
| AVHRR-CLARA | 1 km × 4 km Sampling same as PATMOS-x | Multi-spectral thresholding VIS/NIR/IR<br>Threshold tuning against CALIPSO | **CT** → CZ,CP<br>COD, CRE Phase (W/I) | IR matching of RTM-simulations + IR split-window histogram with ERA-interim T/WV-profiles as reference LUT approach (0.6, 3.7, 11 µm) spectral differences | ERA-interim *reanalysis profiles, rad. transfer + particle model* |
| | | | CWP | fct(COD, CRE, phase) | |
| AVHRR-Cloud_cci | 1 km × 4 km, sampling same as PATMOS-x | VIS/NIR/IR neural network | COD, **CP**, CRE, Phase → CT | Optimal estimation on VIS/NIR/IR (0.7, 0.9, 3.7, 11, 12 µm) | ERA-interim reanalysis *profiles, rad. transfer + particle model* |
| | | | CWP | fct(COD), CRE, phase) | |

**Table 2** (continued)

| Dataset | Spatial resolution | Cloud detection | Variables | Retrieval method | Ancillary input |
|---|---|---|---|---|---|
| MODIS-CE | 1 km, 4 km (sampled) 32 km clear sky estimation | multi-spectral IR/NIR/VIS (5 channels similar to VIRS) | CEM, **CT** → CZ,CP | IR split-window; lapse rate (7.1 K/km) + T profile | GMAO GEOS G5.4.1-CERES reanalysis profiles, *rad. transfer + particle model (mixed habits for ice)* |
| | | | Phase, COD, CRE | CT + LUT approach (0.6, 2.1 µm), (3.8 µm) | |
| | | | CWP | CWP = fct(COD, CRE, phase) | |
| MODIS-CLDPROP | 1 km | multi-spectral IR/NIR/VIS (16 channels)<br>+ time–space variances | **CP**,CEM → CT | $CO_2$ slicing for CP < 650 hPa, TB(11 µm) | NCEP GDAS reanalysis *profiles, 16 day spectral* surf. albedo climatology, *rad. transfer + particle model (mixed particle habits for ice)* |
| | | | Phase (W/I) | VIS/NIR/IR spectral differences | |
| | | | COD, CRE | LUT approach (0.7, 0.9, 1.2, 1 6 µm), (21 µm) | |
| | | | CWP | fct(COD, phsse, CRE) | |
| AIRS-CIRSv2<br>IASI-CIRSv2 | 13.5 km 12 km | a posteriori:<br>coherence of 6 spectral cloud emissivities (9–12 µm) &CEM > 0.05/0.1 (H / ML) | CEM & **CP**, CP → CT,CZ, CEM → COD | weighted $\chi^2$ method on 8 $CO_2$; absorption channels, T/virt. T profile for CT/CZ; COD = —c × In (1-CEM) | ERA-interim reanalysis *profiles, spectral surf. emissivities: IASI climatology;. 4A/OP* + DISORT rad. transfer + ice crystal SSPs of hex. columns and aggregates |
| | | | Phase (W/I) | ice: CT < 250 K liquid: CT > 260 K | |
| | | | CREIH, CIWPH | LUT approach on 6 spectral emissivities (9–12 µm) | |
| MISR | 1 km | multi-specral + angle VIS/NIR | **CZ** | stereoscopic cloud top height | |
| CALIPSO-GOCCP | 0.09 km, sampling: 0.34 km | Lidar VIS backscatter vertical averaging | **CZ** → CT | cloud mean altitude,<br>uppermost cloud layer (for GEWEX) | GMAO MERRA2 reanalysis profiles |

**Table 2** (continued)

| Dataset | Spatial resolution | Cloud detection | Variables | Retrieval method | Ancillary input |
|---|---|---|---|---|---|
| CALIPSO-ST | 0.09 km, sampling: 0.34 km | Lidar VIS backscatter horizontal averaging | **CZ** → CT | cloud top, uppermost cloud layer (for GEWEX) | GMAO MERRA-2 reanalysis profiles |
| | | | Phase (W/I) | ice: 532 nm depolarization | |

**Table 3** Variable names of the cloud properties with statistics also distinguished by altitude (H: CP < 440 hPa, M: 440 hPa < CP < 680 hPa, L: CP > 680 hPa) and by thermodynamic phase (W: water clouds, I: ice clouds, IH: ice clouds with CP < 440 hPa) available in the updated GEWEX Cloud Assessment database

| Variable | Total | *H* | *M* | *L* | *W* | *I* | IH |
|---|---|---|---|---|---|---|---|
| Cloud amount | CA | CAH | CAM | CAL | CAW | CAI | CAIH |
| Effective cloud amount | CAE | CAEH | CAEM | CAEL | CAEW | CAEI | CAEIH |
| Relative to total CA | | CAHR | CAMR | CALR | CAWR | CAIR | CAIHR |
| Cloud temperature | CT | CTH | CTM | CTL | CTW | CTI | CTIH |
| Cloud pressure | CP | | | | | | |
| Cloud height | CZ | | | | | | |
| Cloud emissivity | CEM | CEMH | CEMM | CEML | CEMW | CEMI | CEMIH |
| Cloud optical depth | COD | CODH | CODM | CODL | CODW | CODI | CODIH |
| Cloud water path | | | | | CLWP | CIWP | CIWPH |
| Cloud effective radius | | | | | CREW | CREI | |

morning (nominal equator crossing at 7h30 AM and PM LT) and afternoon (nominal equator crossing 1h30 AM and PM LT) satellites. The observation times slowly drift in time, in particular the ones of the afternoon satellite orbits (cf. Foster et al. 2023). Furthermore, the orbit time of the morning satellites changed from about 7:30 LT to about 10:00 LT with the advent of NOAA-17 in 2002, and during daytime the 3.7-µm channel was replaced by a 1.6-µm channel.

The *Pathfinder Atmospheres Extended* (***PATMOS-x***, *Version 5*) participated with the same retrieval version in the original GEWEX cloud assessment. Cloud detection is based on Bayesian classifiers derived from CALIPSO (Heidinger et al. 2012), and the retrieval is based on an optimal estimation approach (Heidinger and Pavolonis 2009). First cloud pressure (CP) and cloud emissivity (CEM) are retrieved using two IR channels at all times of day. Then, cloud optical depth (COD) and cloud particle effective radius (CRE) are obtained from solar channels during daytime so that finally cloud water path (CWP) can be derived from COD and CRE (Walther and Heidinger 2012).

The ***Climate Monitoring Satellite Application Facility*** (CM SAF) ***Cloud, Albedo and surface Radiation dataset*** from AVHRR data (***AVHRR-CLARA***, *Version 2*, Karlsson et al. 2017) detects clouds based on multi-spectral thresholding, which has been tuned through comparisons with cloud observations from the CALIPSO lidar. Cloud top heights are retrieved using radiance matching with radiances simulated by a radiative transfer model for opaque clouds and a split-window histogram approach for semi-transparent clouds. Cloud phase is interpreted from further spectral analysis, and a look-up table (LUT)-based retrieval of COD and CRE, following the classical Nakajima–King approach, is used. Finally, the cloud water path (CWP) is derived as a function of COD, CRE and cloud phase.

The ***ESA Cloud Climate Change Initiative*** has generated two cloud data records from AVHRR (AVHRR-AM and AVHRR-PM, *Version 3*, Stengel et al. 2020) which are combined in this study (***AVHRR-Cloud_cci*** hereafter). The cloud detection and cloud thermodynamic phase determination are based on artificial neural networks (ANNs), which were trained using data from CALIPSO. For the cloud property retrieval, the optimal-estimation-based Community Cloud retrieval for CLimate (CC4CL) was employed (Sus et al. 2018; McGarragh et al. 2018). AVHRR-Cloud_cci covers the time period 1982–2016. An extension until 2020 exists, which is of lower quality and not provided for this database.

## 2.2 Cloud Properties from Advanced Multi-spectral Imagers

The ***MODIS CERES Science Team*** (***MODIS-CE***) retrieval was recently upgraded (*Edition 4*, Trepte et al. 2019; Minnis et al. 2021), using new calibrations for solar channels, including the 1.24-µm channel for COD retrieval over snow and a new ice crystal reflectance model. The use of $CO_2$ absorbing channels and a regional lapse rate technique are applied to improve the height retrieval of high- and low-level clouds, respectively. CA, CEM, CT, COD and CRE are directly retrieved, and all other variables are deduced from these using ancillary data and parameterizations. MODIS data have been processed from instruments aboard Terra (equator crossing 10h30 AM and PM LT) and Aqua (equator crossing 1h30 AM and PM LT).

The new cloud property continuity product from the National Aeronautics and Space Administration (NASA) ***MODIS Science Team*** (***MODIS-CLDPROP***) is designed to provide continuity between MODIS and the Visible Infrared Imaging Radiometer Suite (VIIRS) with an algorithm that uses only a subset of similar spectral channels available on both imagers. Note that VIIRS is missing key atmospheric absorbing spectral channels used in the MODIS Standard Product (MOD/MYD06) cloud retrievals, in particular the 13–14 µm $CO_2$ slicing channels. CLDPROP (Meyer et al. 2020; Platnick et al. 2021), and the companion common algorithm continuity CLDMSK cloud mask product (Frey et al. 2020), are based on heritage algorithms used in the MODIS Collection 6.1 cloud product suite (Platnick et al. 2017) with the exception of the cloud top property datasets that are based on an optimal estimation algorithm (Heidinger and Pavolonis 2009; Heidinger et al. 2019). CLDPROP is in production for both MODIS Aqua and, to date, VIIRS on SNPP and NOAA-20. For convenience, CLDPROP in this article will also refer to the CLDMSK product and only the MODIS CLDPROP production stream is discussed. COD, CRE and CWP are retrieved using a LUT approach based on solar reflectance and midwave IR channels (Platnick et al. 2021). Only microphysical retrievals that use the MODIS 2.1-µm channel are currently reported as part of the GEWEX dataset (the CLDPROP production version includes retrievals based on other SWIR and MWIR channels). Note that the gridded sampling of COD, CRE and CWP may be a subsample of the CP, CEM, CT pixel population because retrievals of optical and microphysical properties are only attempted for masked pixels that meet stricter requirements than that of the cloud top properties algorithm. Further the optical/microphysical pixels population is partitioned into two separate datasets, a dataset for pixels that appear to be overcast at the spatial resolution of MODIS (1 km at nadir along with 250 m observations over the ocean) and a dataset for pixels that are likely to be partly cloudy (Platnick et al. 2017). As such, files named ***MODIS-CLDPROP*** provide optical/microphysical cloud properties only for the so-called overcast pixels having a successful retrieval (considered highest quality), while ***MODIS-CLDPROP_ALL*** files, also part of the updated cloud assessment database, *provide statistics that include the successful optical/microphysical partly cloudy pixel retrieval population as well.*

### 2.3 Cloud Properties from IR Sounders

The original GEWEX cloud assessment database includes three different IR sounder cloud climatologies: HIRS-NOAA (Wylie et al. 2005), TOVS Path-B (Stubenrauch et al. 2006) and AIRS-LMD (Stubenrauch et al. 2010).

The measurements of the High Resolution Infrared Radiation Sounder (HIRS) instruments, flown on, aboard 16 satellites from 1980 through 2015, have been recently reprocessed by Menzel et al. (2016), after having been carefully intercalibrated. The cloud retrieval method is similar to the one of HIRS-NOAA, but is now applied to HIRS measurements that have been identified as cloudy by the PATMOS-x cloud mask (Heidinger et al. 2012). This dataset is not part of the updated GEWEX cloud assessment database.

The ***CIRS (Clouds from IR Sounders)*** retrieval is an updated version of the AIRS-LMD retrieval. It has been developed for the application on any IR sounder data and has been applied so far to observations from the Atmospheric Infrared Sounder (AIRS), aboard the NASA platform Aqua, and the Infrared Atmospheric Sounding Interferometers (IASI) aboard the EUMETSAT platforms Metop-A and Metop-B (equator crossing at 9h30 AM and PM). A weighted $\chi^2$ method (Stubenrauch et al. 1999) using eight channels along the $CO_2$ absorption band simultaneously provides CP and CEM. CT and CZ are derived from CP, using ancillary atmospheric T and water vapour profiles. T inversions stronger than 2 K are accounted for, setting the near cloud-top to the level of the inversion and scaling the cloud emissivity. An ‘a posteriori’ cloud detection relies on the spectral coherence of the retrieved cloud emissivity. The original CIRS version only kept clouds with CEM > 0.1 (Stubenrauch et al. 2017). Here we present results of CIRS *version 2*, with CEM > 0.05 for high-level clouds (as in AIRS-LMD), and CEM > 0.1 for all lower clouds (CP > 440 hPa). The synergy of AIRS-CIRS and IASI-CIRS (Feofilov et al. 2019) was achieved by using the same ancillary data, in particular atmospheric profiles of the meteorological reanalyses ERA-Interim (Dee et al. 2011). Since the production of ERA-Interim ceased in August 2019, the CIRS retrieval has been recently adapted to ancillary data of ERA5 (Hersbach et al. 2020), and the whole data record until present is being produced for AIRS and for IASI at the French data centre AERIS. This dataset will be part of this updated GEWEX cloud assessment database in the near future.

### 2.4 Cloud Properties from a Multi-angle Solar-Spectral Imager

The Multi-angle Imaging SpectroRadiometer (***MISR***, *version 2*) provides a Cloud Fraction by Altitude product derived from four measurements in the spectral domain between VIS and shortwave-infrared (SWIR) that are collected by nine separate cameras ranging in viewing zenith angles from 0° to 70.5° in the along-track direction (Di Girolamo et al. 2010). Cloud detection is achieved using spectral, spatial, angular signature, and stereoscopic tests, while cloud top height (CZ) is achieved using stereoscopy. This cloud height retrieval does not rely on ancillary products and is independent of radiometric calibration. The solar wavelengths are especially useful for retrieving CZ of low-level clouds, also in the presence of thin higher clouds above. A comparison with Cloud-Aerosol Transport System (CATS) lidar data (Mitra et al. 2021) has shown that MISR detects the lower cloud in a two-layered system, provided top-layer optical depth $< \sim 0.3$. MISR participated in the original GEWEX cloud assessment, and since then the stereoscopic technique has been improved (Mueller et al. 2013). Furthermore, the additional dataset ***MISR_RC*** introduces a *resolution-corrected cloud amount*. MISR_RC uses a machine learning technique that has been trained using collocated MISR and ASTER (15 m resolution) data to correct for the resolution bias in cloud amount caused by coarse (~ 1 km) resolution measurements. This effectively retrieves cloud amount as would be measured at 15 m resolution. Cloud amount is estimated by only counting fully cloudy footprints of a size of 15 m (Jones et al. 2012; Dutta et al. 2020) and therefore provides a lower limit on cloud amount.

### 2.5 Cloud Data from Active Lidar

Two products, using different approaches, are based on lidar data of the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) mission, and both approaches participated in the original GEWEX cloud assessment. Differences between these products in the retrieved cloud properties are related to differences in the detection and cloud–aerosol classification algorithms used. In order to be more compatible with the results from passive remote sensing, both datasets provide a version which only keeps the information of the uppermost cloud layer in the case of multiple cloud layers.

The ***GCM-Oriented CALIPSO Cloud Products*** (***CALIPSO-GOCCP***, *Version 2.9*) are derived from the CALIPSO L1 Version 4 data, after vertically averaging single-shot profile data (0.33 km along track × 0.090 km cross-track) to 0.48 km. The method used is the same as described in (Chepfer et al. 2010), using a solar noise correction for daytime data described in the first GEWEX Cloud Assessment report (Stubenrauch et al. 2012). Cloud amounts are computed as the number of cloudy profiles within a grid cell divided by the total number of profiles.

The ***CALIPSO Science Team*** (***CALIPSO-ST***, *Version 4.2*) uses horizontal averaging (5, 20 or 80 km) to detect weakly scattering, thin cirrus, while only single-shot data are used for water clouds and dense ice clouds. The 90 m receiver footprints of the CALIPSO lidar are separated by 333 m along-track (Winker et al. 2010). Cloud amounts are computed as the number of cloudy profiles within a grid cell divided by the total number of profiles. The revised dataset uses CALIPSO L2 Version 4.2 data products, which feature an improved aerosol-cloud discrimination algorithm (Liu et al. 2019), an improved ice-water classification (Avery et al. 2020) relative to the original version of CALIPSO-ST, and an improved selection of cirrus lidar ratio constrained by infrared observations (Garnier et al. 2015).

For the updated cloud assessment database, in addition to the cloud amounts of only the uppermost cloud layer (here reported as CALIPSO-ST_top), CALIPSO-ST reports the cloud amounts of all detected cloud layers within the atmospheric column (CALIPSO-ST_column), one based on only the highest cloud layer in each column having COD > 0.3 (CALIPSO-ST_passive) and one based only on cloud layers which attenuate the lidar surface return signal to the point where it can no longer be detected (CALIPSO-ST_opaque), approximately at COD > 3 (Garnier et al. 2021a). *CALIPSO-ST_passive is meant to give an indication of the effect on cloud amount due to sensitivity limits of a typical passive imager, recognizing that this COD threshold detection limit depends on passive instrument and technique, underlying surface characteristics, cloud type, aerosol loading, sun-view geometry, and thermal structure of atmosphere and surface. However, CALIPSO-ST_passive reports the actual lidar-detected cloud top altitudes (as do all CALIPSO-ST products) and therefore does not mimic the cloud heights retrieved by passive sensors* (Sect. 2.8).

### 2.6 Cloud Data from Combined Active Lidar and IR Imager

Since the IR spectrum is sensitive to the size of ice crystals (up to a radius of about 50 μm), an infrared imager radiometer (IIR) was developed to accompany the lidar of the CALIPSO mission. It has three channels between 8 and 12 μm and a footprint size of 1 km. The IIR_CALIPSO-ST products, new to this assessment, include IIR ice and water microphysical retrievals for single-layer scenes identified by the CALIPSO lidar. Cloud radiative temperatures are estimated from the CALIPSO lidar profiles (Garnier et al. 2021b). IIR_CALIPSO-ST provides CT and CEM, further converted to COD, but only where microphysical properties are successfully retrieved in single-layer ice or water clouds.

### 2.7 Cloud Data from Combined Active Radar and Lidar

While the lidar cannot probe the atmosphere below optically dense clouds, the radar cannot detect thinner layers at any level and has a reduced sensitivity near the ground (1 km), because of surface clutter. Therefore, the most complete picture of the cloud vertical structure is obtained by combining the lidar with radar measurements, both part of the A-Train mission since 2006 (Stephens et al. 2018a). Compared to passive remote sensing, which does not provide such a complete picture of vertical structure, lidar and radar sampling provides a more sophisticated vertical view but only along narrow nadir tracks. These measurements are very sparse in comparison with observations of cloud properties retrieved from passive remote sensing. This very sparse sampling leads to uncertainties in regional mean cloud properties (e.g. Astin et al. 2001; Kotarba 2022). In particular, individual radar–lidar transects need to be averaged at least over an area 10° × 10° in longitude and latitude in order to achieve an accuracy of 1% (for CA) or 150 m (for CTH). The annual mean CA estimate is very sensitive to infrequent sampling, resulting in 14% or 7% average uncertainty over 1° or 2.5°, respectively.

For the comparisons in Sect. 4, we analysed CloudSat-CALIPSO GEOPROF data (*version R04*, Mace et al. 2009) of a dataset provided by J. Mace for this data assessment. It is available at https://macegroup.chpc.utah.edu/qzhang/occ/occ/occ.jsp and includes monthly statistics on maximal two vertical cloud layers, at 1° latitude × 1° longitude. Cloud top statistics has been stratified into eight categories with values of 17, 12, 8, 4.5, 2.5, 1.5, 0.75 and 0.25 km. Seven different cloud layer thicknesses are assigned with each cloud top, given by values of 0.125, 0.375, 1.00, 2.25, 4.50, 8 or 12 km. Since only maximal two cloud layers are given in this dataset, we extended the cloud layer thickness towards the surface for those cloud layers with a vertical extent of at least 8 km. For a comparison of cloud amounts stratified into high-, mid- and low- level clouds according to pressure (with boundaries at 440 and 680 hPa), polar boundaries at 3 and 6 km were increased towards 3.6 and 7.2 km at lower latitudes (35 S–35 N). Compared to version R04, the new version R05 (Mace and Zhang 2014) shows reduced false detections of the CloudSat cloud mask but at a cost of a significant loss in the true weak signals. By comparing with the CALIPSO lidar vertical feature mask, Hu et al. (2020) have improved this scheme for a future CloudSat data processing which will lead to a better detection of thin clouds.

### 2.8 Database

The database covers the same variables as the original GEWEX cloud assessment database (Stubenrauch et al. 2012). These variables are listed in Table 3.

In addition to averages over grid boxes of 1° latitude × 1° longitude, the database also contains monthly histograms of cloud properties given at their retrieval spatial resolution. Exceptions are the MODIS-CE and ISCCP histograms, containing values which have been first averaged over 0.2° (CERES sensor footprint) and 1° (ISCCP map grid resolution), respectively. Only the 2-dimensional COD-CP and CEM-CP histograms of ISCCP were constructed from the initial 10 km pixels.

*Cloud amount* (CA) is difficult to quantify at high accuracy, as it *depends on satellite sensor ability to detect optically thin clouds and partial cloud cover over the instantaneous field of view (IFOV)*. In general, for the retrieval of cloud properties it is assumed that the IFOV is completely cloud-covered. Overall, the detection thresholds should be chosen to balance between detecting and overestimating thin cirrus or broken low-level clouds (e.g. Wielicki and Parker 1992). However, there is no guarantee that this balance can be achieved to reach unbiased CA at regional scales (e.g. Zhao and Di Girolamo 2006).

Fortunately, those clouds which are the most difficult to detect are also those with the smallest radiative effect (e. g. Zhang et al. 2004), but they may still play a role in other applications (e.g. Koren et al. 2008; Pincus et al. 2012).

*Height-stratified cloud amounts* (CAH, CAM, CAL) give separate amounts of high-, mid- and low-level clouds. Satellite sensors observe the clouds from above, and hence, in the case of multiple layers, passive remote sensing only yields the properties of the uppermost cloud layer or of the whole cloud column. This means that the amounts of cloud top layers sum up to total cloud amount: $\mathrm{CAH}+\mathrm{CAM}+\mathrm{CAL}=\mathrm{CA}$, while the real amount of the underlying cloud layers may be larger due to parts hidden by the layers above. The thresholds of 440 hPa ($>6$ km) and 680 hPa ($<3$ km), initially defined for ISCCP (Rossow and Schiffer 1991), were motivated by cloud type classification of surface observers, adjusted to approximate the difference between cloud base and cloud top locations. These designations have been consistently followed for both versions of the GEWEX Cloud Assessment database.

The most *accurate geometrical cloud-top height* can be determined by spaceborne lidar (CALIPSO) and by a stereoscopic retrieval from multiple viewing angles (MISR). In the case of multi-layer clouds, the latter usually reports the height of the low-level clouds. Both approaches, however, come at the cost of low spatiotemporal sampling, particularly CALIPSO. Passive IR sensors are usually better in this respect, though the remote sensing techniques applied to their measurements provide a '*radiative cloud height'*. For water clouds, this is less of a problem as those clouds have in general well-defined cloud-tops. For ice clouds that have a 'fuzzy' cloud-top however, the retrieved cloud-top height may lie a few kilometres below the actual cloud-top height, depending on their extinction profile, e.g. in tropical high-level clouds (e.g. Liao et al. 1995a, b; Wylie and Wang 1997; Minnis et al. 2008; Hamann et al. 2014; Stubenrauch et al. 2010, 2017; Mitra et al. 2021). For opaque, tropical deep convective clouds, it has been demonstrated that the mean effective emission level of the MODIS 11-μm band corresponds to a level at which the cloud reaches an IR optical depth of about $0.7\pm0.3$ (Wang et al. 2014). The CIRS cloud height corresponds to a level within the cloud where the optical depth from CALIPSO reaches about 0.5 or to a level near mid-cloud for optically thin clouds (Stubenrauch et al. 2017). The difference between cloud-top height and radiative cloud height affects CAH differences between CALIPSO and the other sensors only at higher latitudes, because the fixed threshold of 440 hPa is in the tropics far below the tropopause, while at higher latitudes (in particular in winter) 440 hPa is close to the tropopause and therefore the radiative cloud height may be in the mid-level cloud category while the cloud-top is in the high-level cloud category (Fig. 6).

Other difficulties in the interpretation of retrieved cloud height arise from *differences in footprint size* (1–15 km), with larger footprints leading to a larger probability of partial overlapping multiple cloud layers and from the retrieval of different variables which represent height (cloud temperature or cloud pressure), which are then transformed to height by using *different ancillary data* (Sect. 3.2).

To diagnose cloud radiative effects, one needs in addition to cloud amount at least cloud temperature, VIS optical depth and IR emissivity. The height-stratified *effective cloud amount* (CAEH, CAEM, CAEL), defined as the cloud amount weighted by the cloud IR emissivity for the specific cloud height category, should reconcile differences in cloud amount, in particular CAH, because a smaller cloud amount due to lower sensitivity is compensated by larger cloud IR emissivity.

*Cloud water path* (CWP) and *effective particle radius* (CRE, averaged over a size distribution within the cloud) are the variables predicted or parameterized in climate models. These variables are then used in radiative transfer computations to obtain the *optical depth* COD. In general, at a constant CWP, the solar albedo increases with decreasing CRE. The latter may be retrieved from multi-spectral differences in the solar or thermal domain, using pre-computed look-up tables. The radiative transfer computations *need certain assumptions*, such as optical properties and crystal habit in the case of ice clouds and shape of the particle size distribution. These retrievals provide CRE at a certain depth within the cloud, which depends on the wavelengths used and on the opaqueness of the clouds. While for low-level liquid clouds in general the effective cloud droplet size decreases from top to base, the effective ice crystal size increases from top to base when sedimenting (e.g. Jensen et al. 2018).

# 3 Assessment of the Individual Datasets

## 3.1 Total and Height-Stratified Cloud Amounts

### 3.1.1 Averages

Figure 2 presents global averages of total and height-stratified cloud amounts of the eleven datasets participating in the updated GEWEX cloud assessment database. These averages are area-weighted, day-night averages over 1:30 AM and PM LT, except for IASI (9:30 AM and PM LT) and MISR (10:30 AM LT).

**CA:** Results are in general similar to those of the original GEWEX Cloud Assessment. The ***global cloud amount*****,** ***CA*****,** ***0.66 ± 0.04*****, is** ***slightly lower than before*** (0.68 ± 0.03), due to a much lower value from the new CALIPSO-ST version V4 (0.66 compared to 0.73 based on version V2) and due to slightly lower CA from the newly developed datasets AVHRR-CLARA and AVHRR-Cloud_cci. The decrease in CA of CALIPSO-ST is mostly due to an improved aerosol–cloud discrimination (Liu et al. 2019), which corrected for high altitude dust layers classified as cloud in earlier data versions. Keep in mind that these averages come mostly from observations at 1:30 AM and PM. Therefore, they include a

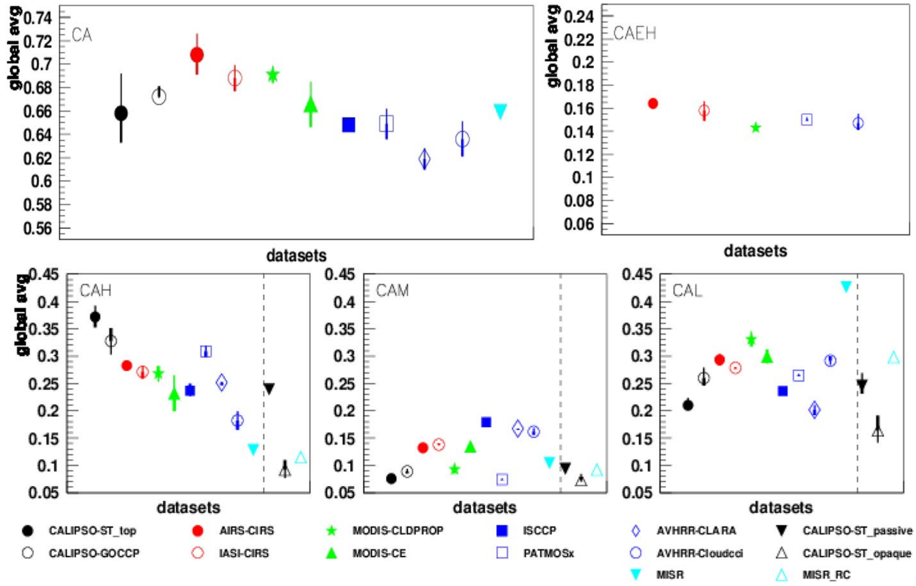


**Fig. 2** Upper left panel: Global averages of total cloud amount (CA) of the eleven datasets participating in the updated GEWEX cloud assessment database. Upper right panel: Global averages of CAEH of the five datasets providing this variable. Lower panel: Global averages of height-stratified cloud amount (CAH, CAM and CAL) of the eleven datasets. CAH and CAM of clouds with COD$>$0.3 and of clouds with COD$>$3 from CALIPSO-ST as well as results from MISR_RC are also shown for comparison. Statistics averaged over day and night (1:30AMPM LT, except IASI (9:30AMPM) and MISR (10:30 AM)) over the common period of 2008—2015. *The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers, cyan: solar-spectral imager). The vertical bars indicate day (thick) and night (thin) differences with respect to the daily mean*

small uncertainty due to the sparse time sampling. The CA averages at all available specific observation times available in this database are shown in Table S1 in the supplement.

In Fig. 2, we also show day–night differences, with the end of the thicker (thinner) bar indicating the daytime (night-time) value. A study with CATS lidar data (50 N–50 S, Noël et al. 2018) found that over ocean 1:30 PM and 1:30 AM correspond to the daily minimum and maximum cloud amount, respectively, while the diurnal cycle over land is slightly underestimated when using only these observation times. *Most datasets show indeed a slightly smaller global CA during day than during night.* The larger day-night spread for CALIPSO-ST is partly due to daytime noise from solar scattering, for which the CALIPSO-GOCCP data were corrected.

**CAH** We observe a large spread in CAH in the lower left panel of Fig. 2, where the datasets are grouped with decreasing detector sensitivity to thin high-level clouds. This spread in CAH is mainly explained by differences in instrument and retrieval method sensitivity to thin high-level clouds, including scenes with lower clouds underneath, and, in comparison with CALIPSO, by height-stratification differences according to cloud-top and radiative cloud height (see Fig. 6).

Active lidar has the highest sensitivity to thin cirrus, followed by IR sounders (with a high spectral IR resolution and no contributing VIS reflectance from underlying clouds). The IR-VIS imagers show slightly lower values of CAH (Figs. 2 and 3). While thin

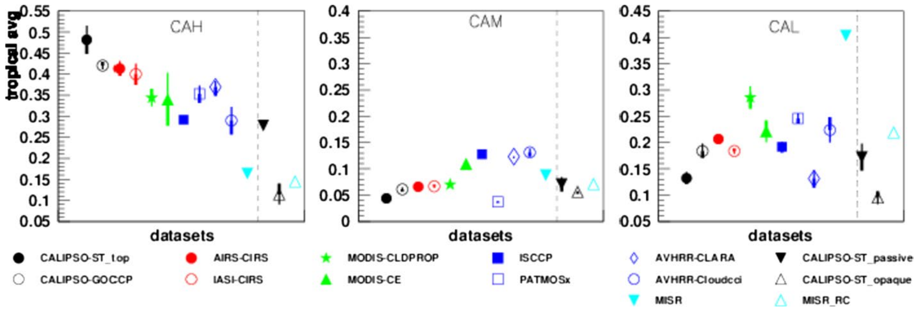


**Fig. 3** Tropical (15N–15S) averages of height-stratified cloud amount (CAH, CAM and CAL) of the eleven datasets. CAH and CAM of clouds with COD > 0.3 and of clouds with COD > 3 from CALIPSO-ST as well as results from MISR_RC are also shown for comparison. Statistics averaged over day and night (1:30AMPM LT, except IASI (9:30AMPM) and MISR (10:30 AM)) over the common period of 2008—2015. *The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers, cyan: solar-spectral imager). The vertical bars indicate day (thick) and night (thin) differences with respect to the daily mean*

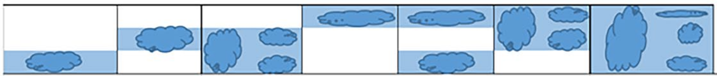
cirrus can often be identified, e. g. by spectral IR differences, their height assignment is difficult as mentioned in Sect. 2.8.

In order to quantify the two categories, 1) missed thin cirrus in a single layer and 2) mis-identified thin cirrus overlying clouds at lower levels, we use information from the different CALIPSO-ST datasets and our CloudSat-CALIPSO GEOPROF analysis, summarized in supplemental Table S2. We estimate the *amount of all thin cirrus (COD < 0.3)* as the CAH difference between CALIPSO-ST_top (0.37) and CALIPSO-ST_passive (0.24). The global average is 0.13, with largest amounts in the tropics and over Antarctica, according to Figure S1a in the supplement. Indeed, the CAH of MODIS-CE, ISCCP and AVHRR-CLARA are close to the CAH of CALIPSO-ST_passive. Yet CAH of CALIPSO-ST_passive provides only a rough estimate of what a passive imager would estimate, because the accumulative CAH changes by about 0.1 for high-level clouds with COD between 0.2 and 0.3 (Balmes et al. 2019).

We further deduce from the CloudSat-CALIPSO data that 35% of all high-level clouds are single-layer clouds (Table 4). By assuming that the overlap between cirrus and lower clouds is independent of the COD of the high-level cloud, we derive

**Table 4** Top panel: Global averages of height-stratified cloud amount, as seen from above (top-layer); bottom panel: column view with an illustration of 7 simplified cloud vertical structure classes, with columns including high-, mid- and low-level clouds (multiple cloud layers and contiguous cloud layers together). Amounts are reported as average from CloudSat-CALIPSO GEOPROF and CLDCLASS data (Sect. 2.7 and Oreopoulos et al. 2017). Statistics covers 82N-82S over 2007–2009 and 2007–2010, respectively. The ranges correspond to the difference between the average and the single analysis values and an additional 10% uncertainty due to the use of an earlier CALIPSO version (V3)

| L | M | H | top-layer |
|---|---|---|---|
| 0.24 ± 0.05 | 0.08 ± 0.02 | 0.40 ± 0.07 | CloudSat-CALIPSO |



| 1L | 1M | M L | 1H | H L | H M | H M L | column |
|---|---|---|---|---|---|---|---|
| 0.24 ± 0.05 | 0.03 ± 0.01 | 0.06 ± 0.02 | 0.14 ± 0.03 | 0.08 ± 0.02 | 0.07 ± 0.03 | 0.12 ± 0.03 | CloudSat-CALIPSO |

the *amount of single-layered thin cirrus (COD < 0.3)* as 0.35 x (CAH(CALIPSO-ST_top)—CAH(CALIPSO-ST_passive)) = 0.05. The remaining amount of 0.08 (about 60%) corresponds to *thin cirrus overlying clouds at lower levels*. The latter should then correspond to the difference between the CAM derived by imager and CAM of CALIPSO-ST_passive. This seems indeed to be the case for ISCCP, AVHRR-CLARA, AVHRR-Cloud_cci and MODIS-CE.

Furthermore, we observe *large differences due to retrieval methodology*: The CALIPSO-ST horizontal averaging scheme and detection thresholding are more sensitive to subvisible cirrus than the CALIPSO-GOCCP vertical averaging, with CAH of 0.37 and 0.33, respectively. CAH deduced from AVHRR measurements even ranges from 0.18 (AVHRR-Cloud_cci) to 0.31 (PATMOS-x). The latter value is due to an overestimation at higher latitudes, in particular in the SH (see Fig. 5).

The ISCCP results do not show any day-night difference because night-time results were adjusted by adding the daytime difference of VIS/IR and IR results interpolated over the night-time. Since the combination of VIS-only information together with stereoscopy favours the height retrieval of lower clouds, CAH from MISR (0.12) is the lowest amongst the datasets and only slightly larger than the amount of opaque high-level clouds (0.09) estimated from CALIPSO-ST_opaque.

In order to mitigate the effect of the difference between cloud-top and radiative height on CAH, we investigate the tropical averages (15N-15S) of the height-stratified cloud amounts in Fig. 3. We see indeed a smaller spread in CAH than for the global averages. The relative difference in CAH between CALIPSO-ST and CIRS decreases from 24% in global average to 15% in the tropics, and CAH of CIRS and CALIPSO-GOCCP is very close. The spread within the AVHRR datasets has also decreased. CAH of ISCCP and now also AVHRR-Cloud_cci are close to CAH of CALIPSO-ST_passive. The retrieval methods of PATMOS-x and AVHRR-CLARA seem to have a higher sensitivity towards high-level clouds, close to those from MODIS.

Similar results in CAH between CIRS and MODIS-CE during night (0.29 and 0.27 global, 0.42 and 0.41 tropical, respectively) show that the sensitivity of radiometers can be enhanced by using several IR spectral channels. However, when the retrieval method includes VIS information during the day, without considering the possibility of underlying clouds, thin cirrus may be mis-identified as a more opaque lower cloud because of a large VIS reflectance from the cloud underneath.

*Therefore, we recommend to develop retrieval methods (during daytime) which are able to distinguish single- and multi-layered cloud scenes. This may be achieved by first using the IR radiances and their spectral differences alone and then comparing the retrieved cloud properties to those obtained when including VIS reflectances.*

ISCCP is using a two-step method (first IR and then VIS), but because of building a coherent climatology based on geostationary data back to the 1980s, it is not possible to include spectral IR differences. The MODIS-CLDPROP retrieval treats IR and VIS separately, leading indeed to CEM and COD from different methods, but the sampling of cloudy scenes is different: all cloudy scenes (0.69 over 60 N–60 S) for CEM and only 'overcast' cloudy scenes (0.47 over 60 N–60 S) for COD. Therefore, their statistics are difficult to combine, but in this database products with the same IR and VIS sampling are also provided, which will be compared in Sect. 3.4.

**CAEH** The effective amount of high-level clouds, CAEH, of the datasets agrees much better (Fig. 2, upper right panel) relative to CAH, with a global average of about 0.15 ± 0.01 from the five datasets, which directly retrieve CEMH. Global averages of CAE and CEM, as well as their height-stratified averages, are shown separately in supplemental Figures S2 and S3.

**CAM** The global amount of mid-level clouds with no higher clouds above is quite small, with smallest values from CALIPSO, PATMOS-x and MODIS-CLDPROP (0.08 ± 0.01). The CAM averages of ISCCP, AVHRR-CLARA and AVHRR-Cloud_cci are largest (0.18 ± 0.01), because they likely include thin cirrus clouds for which the height assignment is biased low. Considering the tropics (Fig. 3), also CAM of CIRS is small and very close to the values of MODIS-CLDPROP and CALIPSO, while the global average is slightly larger. This may be explained by the fact that the cloud height retrieved by CIRS has a broader distribution than MODIS-CLDPROP (see Fig. 6), most probably due to the worse spatial resolution which increases the probability of partly overlapping multiple layer clouds in the field of view, and at higher latitudes the threshold of 440 hPa puts some of these clouds into a lower cloud category.

**CAL** The global amount of single-layer low-level clouds is about 0.26 ± 0.05, with CALIPSO-ST_top and AVHRR-CLARA at the lower end (0.21) and MODIS, CIRS and AVHRR-Cloud_cci at the higher end (0.30 ± 0.02). Again, the global mean of CAL provided by ISCCP is close to the one of CALIPSO-ST_passive. The low CAL of AVHRR-CLARA can be explained by the misidentification between low- and mid-level clouds, in particular over land and in the stratocumulus regions (see appendix Fig. 15). The larger values may be overestimated due to partly cloudy footprints, which have been counted as overcast. CAL of MISR (0.42) is very similar to $CAL_{col}$ (corrected for clouds underneath opaque high-level clouds, see Sect. 4). However, as stated in Sect. 2.4, MISR detects low-level clouds underneath clouds with COD < 0.3. Hence, one expects CAL of MISR to be similar to CAL of CALIPSO-ST_passive (0.25). Therefore, the combination of VIS-only reflectances together with stereoscopy seems to largely overestimate single-layer CAL. Compared to MISR, MISR_RC shows much smaller values for CAL (Figs. 2, 3). Partly cloudy fields can be identified by the ratio of CAL of MISR_RC over CAL of MISR that is smaller than 1. When considering the geographical map of this ratio in Figure S1b in the supplement, the largest changes are in the trade wind cumulus regions and over land in the tropics and subtropics.

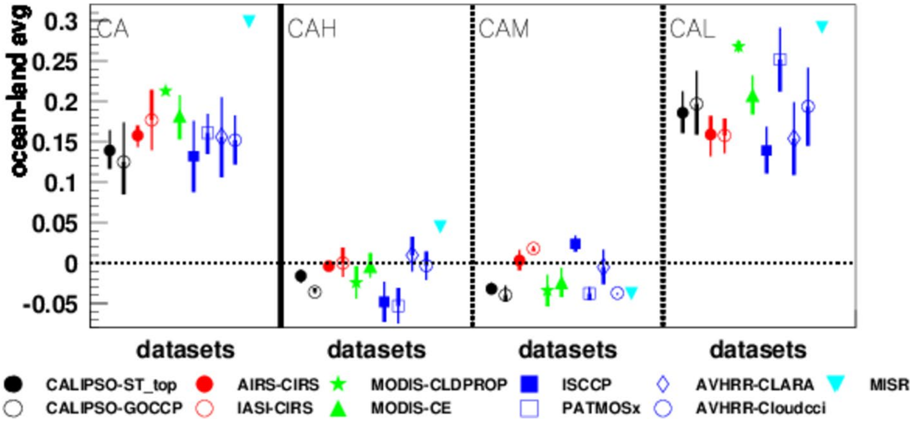


**Fig. 4** Ocean minus land differences in total cloud amount (CA) and in height-stratified cloud amount of the eleven datasets participating in the updated GEWEX cloud assessment database. Statistics averaged over day and night (1:30AMPM LT, except IASI (9:30AMPM) and MISR (10:30AM)) over the common period of 2008–2015. The same land–ocean mask was used in the analysis. *The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers, cyan: solar-spectral imager). The vertical bars indicate day (thick) and night (thin) differences with respect to the daily mean*

From Fig. 4, we deduce that the *cloud amount over ocean is larger than over land, by about 0.15 ± 0.05, which is almost entirely due to more low-level clouds over ocean.* The spread in the CAL difference is larger than 0.10. The MISR results differ from the others by a twice as large ocean-land difference in CA and a positive ocean-land difference in CAH. The geographical maps in appendix Figure 15 indicate that MISR identifies low-level clouds over ocean where other datasets identify high-level clouds, as expected since stereoscopy favours the height retrieval of lower clouds when cirrus are aloft, while PATMOS-x provides less low-level clouds over land at higher latitudes, because of identifying more high-level clouds there (due to a training with CALIPSO-ST which includes polar stratospheric clouds).

Instrument retrieval capabilities aside the observed spread in annual global ocean-land cloud amount differences do have other contributing factors tied to sampling differences between the instruments. This includes the time of day in which samples are collected (thus impacted by the diurnal cycle of clouds), the instrument swath (largely impacting the amount of coverage near the poles), and the reliance on sunlight of the retrieval algorithms (largely impacting large areas at mid to high latitudes of the winter hemisphere).

In order to estimate the uncertainty due to limited diurnal sampling, we also investigated the ocean-land differences given at 7:30 and 10:30 AM and PM in supplemental Figure S4. *The values averaged over AM and PM are very similar between 1:30 and 7:30–10:30 LT.* However, the day–night spread of the ocean-land difference in CA between 1:30 AM and 1:30 PM LT reaches 0.1 and originates mostly from low-level clouds. This is most probably due to their nearly opposite diurnal cycle over ocean and over land (cf. Eastman and Warren 2014; Rossow et al. 2022). The only dataset with a negligible day-night spread is MODIS-CLDPROP. Missing observation times, in particular early morning and late afternoon, lead to further uncertainty in these comparisons of about 0.02–0.05 (Rossow et al. 2022).

The spread of 0.06 in the ocean-land difference of CAH (excluding MISR) vanishes when considering CAEH (not shown). This indicates that the slightly larger CAH values over land are due to more semi-transparent cirrus.

### 3.1.2 Latitudinal and Seasonal Variation

Figure 5 presents the latitudinal variation of CA and CAH as well as their differences between boreal summer and boreal winter. *In general, the latitudinal variation agrees well between all datasets, except at the polar latitudes.* We concentrate our discussion in this section on CAH, because the variation of CAM and CAL is influenced by CAH and by the sensor or retrieval sensitivity of thin high-level clouds and therefore difficult to interpret. The latitudinal variation of the column amount of mid- and low-level clouds, including the amount hidden by cloud layers aloft, is discussed in Sect. 4.

The largest CA is found in the SH mid-latitudes, mostly from single-layer low-level and mid-level clouds in summer (supplement Figure S5) and from high-level (storm) clouds during winter. The seasonal oscillation of SH mid-latitude cloud tops was noted in Rossow and Schiffer (1999, Fig. 8), where it was shown that in the NH CP does not vary while CT does with air temperature but in the SH CP varies seasonally. The local peak around 5 N of the annual mean of CA is due to the high-level clouds in the Inter-Tropical Convergence Zone (ITCZ).

The peak in CAH is seen by all datasets, but with a large range from 0.35 (AVHRR-Cloud_cci) to 0.55 (CALIPSO-ST_top), and MISR with 0.2, as discussed in Sect. 3.1.1.

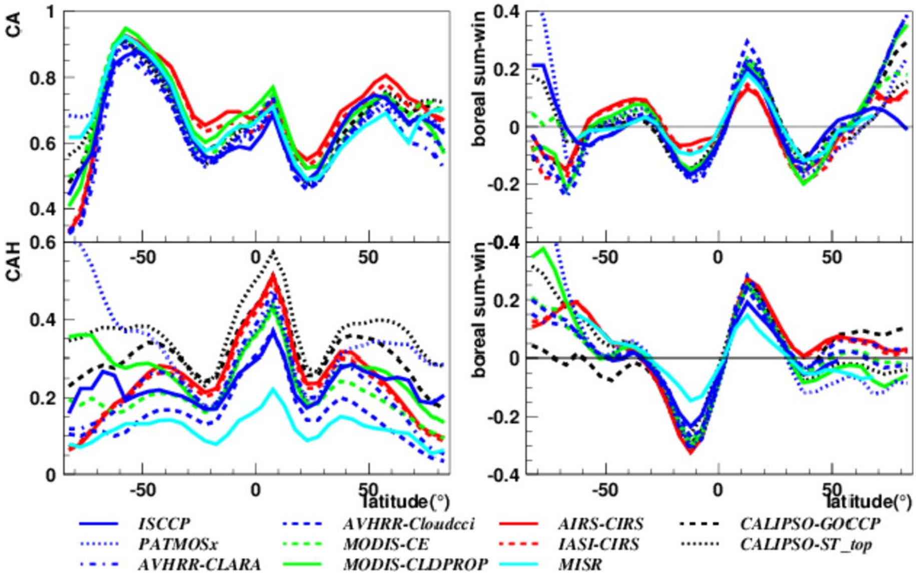


**Fig. 5** Left panel: Zonal averages of CA (top) and CAH (bottom) of the eleven datasets participating in the updated GEWEX cloud assessment database. Right panel: Differences between boreal summer (June, July, August) and boreal winter (December, January, February) of CA (top) and CAH (bottom). Statistics averaged over day and night (1:30AMPM LT, except IASI (9:30AMPM) and MISR (10:30AM)) over the common period of 2008—2015. The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers, cyan: solar-spectral imager)

The displacement of the ITCZ by about 20° between boreal summer and boreal winter is well described by all datasets. The CAH latitudinal variation of MISR is close to the one of CALIPSO-ST_opaque (not shown). Considering CAEH (supplement Figure S6), the local minima and maxima in the zonal means are at similar latitudes but they are less pronounced, and the tropical peak value lies between 0.2 and 0.35, indicating more thin cirrus in the tropics than at the other latitudes. PATMOS-x, the passive remote sensing dataset with the largest global mean CAH, agrees in the tropics with the other datasets, but has much larger values at higher latitudes. This may be related to the fact that in the training with CALIPSO polar stratospheric clouds were included. In particular, the extremely large CAH southward of 50S is identified as very thin high-level cloud (in comparison with CAEH, supplemental Figure S6).

The *spread of CAH in the polar latitudes is large*, *because* on one hand *passive remote sensing of clouds is less reliable over ice and snow,* and on the other hand *these regions are often covered by multi-layered clouds with thin cirrus (polar stratospheric) aloft low-level clouds*, deduced from cloud amounts of CALIPSO-ST_column (see Sect. 4).

The subtropics in both hemispheres have a local minimum in CA and in CAH. The NH mid-latitudes have a smaller CA in summer than in winter. At higher latitudes, we observe a large CAH difference between CALIPSO and all passive remote sensing datasets, while their CAM exceeds the one of CALIPSO (Figure S5). This is due to the retrieval of cloud-top height and radiative cloud height, respectively. At these latitudes, the latter may fall into a lower height category than the cloud-top height (see Sect. 3.2).

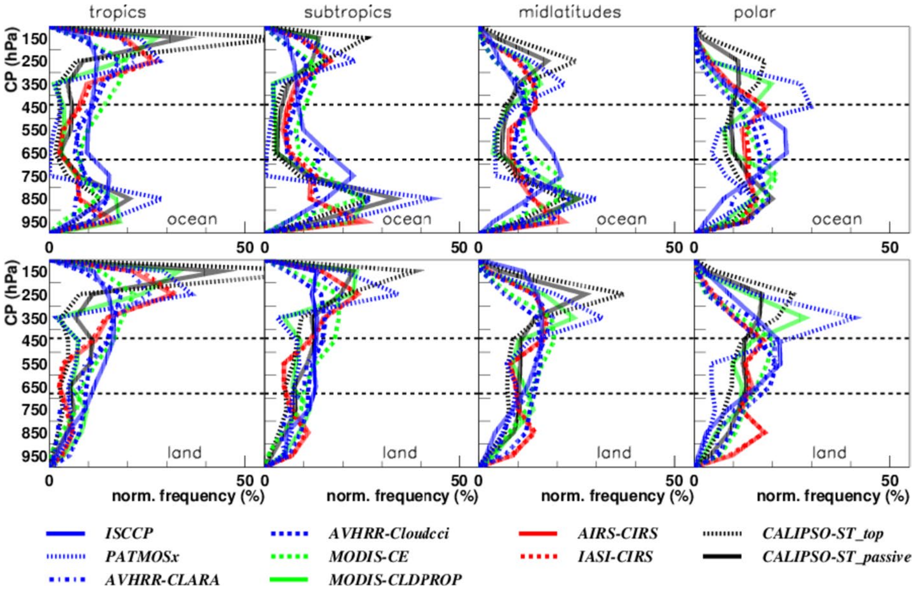


**Fig. 6** Normalized frequency distributions of near-cloud-top pressure, CP, in four latitude bands (tropics: 10° N–10° S; subtropics: 10°–30°, mid-latitudes: 30°–60°; and polar: 60°–90°), separately over ocean (top) and over land (bottom). The frequency is given per 100 hPa. Statistics averaged over 1:30AM and PM LT (9:30AM and PM IASI). These normalized distributions only reflect a view from above, not considering hidden cloud layers. The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers)

### 3.2 Height

Figure 6 presents the normalized frequency distributions of CP for four latitude bands, separately over ocean and over land. For the normalization, the histograms given in the database are divided by the total number of cloudy cases (sum over all events). Since the interval width of the histograms is 100 hPa, these frequencies are given per 100 hPa. These normalized distributions only reflect a view from above, not considering hidden cloud layers. We have also included the distributions of CALIPSO-ST_passive in order to illustrate the effect of sensitivity. Similar distributions of CZ are given in the supplement Figure S7, which includes MISR and CALIPSO-GOCCP.

*In general, the CP distributions over ocean are bimodal, with relatively large populations of cloud-tops in the upper and in the lower troposphere. Over land, the contribution of single-layer low-level clouds is smaller*, as expected from Fig. 4. *The normalized frequency of low-level clouds is highest over subtropical ocean in all datasets except for AVHRR-CLARA.*

A weaker mid-level mode appears at 5 km in the tropics in CZ for MISR and MODIS-CLDPROP and for CALIPSO-ST over land (supplemental Figure S7). A tri-modal distribution of convection with peaks at 16 km, 6 km and 2 km, corresponding to the cloud top height of Cumulonimbus, Congestus and shallow cumulus, was found by Johnson et al. (1999) with data from the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response (TOGA-COARE) Experiment. This congestus peak appears at about 7 km in MODIS-CE, while the CIRS and AVHRR datasets (CLARA and Cloud_cci) show a continuous distribution, probably linked to a coarser spatial resolution. This indicates that the tropics have a wide range of cloud heights in the convective regions, with narrow convective towers and widespread anvils and thin cirrus in the upper troposphere (UT) and congestus clouds in the middle troposphere as well as shallow cumulus and stratus in the subsidence regions (see also Sect. 4).

Compared to the original GEWEX cloud assessment (Stubenrauch et al. 2013), there are *still large differences in these profiles between the different datasets*, *but these can be understood as follows:*

1) retrieved cloud-top (CALIPSO, MISR) versus radiative height within the cloud:

In the tropical UT, the CALIPSO-ST cloud-top height peaks are narrower and at higher altitude (by about 50 hPa or 2—2.5 km) than radiative cloud height peaks obtained from passive remote sensing, which are indeed, as expected, up to a few km below the tops. An exception is MODIS-CLDPROP in the tropics, with peak values at the same height as CALIPSO-ST, but nevertheless a broader distribution. The vertical averaging of CALIPSO-GOCCP leads to a broader CZ peak than CALIPSO-ST (supplemental Figure S7).

From the tropics towards the poles, the peaks in the UT decrease in height (increase in CP) and get broader. *While the cloud-tops determined by CALIPSO stay in the UT, the radiative cloud height starts to decrease towards the middle troposphere, which partly explains an increasing height-stratified cloud amount difference between CALIPSO and passive remote sensing towards higher latitudes.*

2) different sensitivity to thin cirrus, in particular when overlying lower clouds:

From the comparison between CALIPSO-ST_top and CALIPSO-ST_passive, we deduce a large contribution of thin cirrus in the tropics close to the tropopause, but also the polar regions have a nonnegligible contribution. The polar thin cirrus clouds often appear in combination with lower clouds (estimated by the difference of CAL between corrected CALIPSO-ST_column and CALIPSO-ST_top, see Sect. 4). Part of these polar clouds of CALIPSO are stratospheric clouds, which are most likely missed by passive remote sensing. While most datasets show a bimodal distribution also over the polar latitudes, ISCCP, AVHRR-CLARA and AVHRR-Cloud_cci show profiles with one peak in the middle troposphere.

In the tropical UT, the CP peak positions and amplitudes of CIRS, PATMOS-x and AVHRR-CLARA are quite close to each other, whereas the CP peaks decrease in amplitude, by getting broader, towards MODIS-CE (in particular during daytime, not shown), AVHRR-Cloud_cci and ISCCP. Here we have to keep in mind that ISCCP in the GEWEX database provides histograms of averages over 1°, and therefore these results are worse than using the original dataset (Rossow et al. 2022).

The CZ distribution of MISR (supplemental Figure S7, zoomed version) shows a much larger low-cloud mode than the other datasets because MISR favours the accurate height retrieval of low clouds. While the peak in the low- and mid-level cloud mode between CALIPSO-ST_passive and MISR are within their reported uncertainty, the peaks in their high-level cloud mode are ~ 3 km apart in the tropics (at 15.5 km and 12.5 km, respectively). This is more than 2 km larger than the uncertainty in retrieved heights of thin cirrus for CALIPSO and MISR. This, together with the very small peak amplitude of MISR, suggests that, in the tropics, there is a much higher relative frequency of occurrence of optically thin cirrus with tops at ~ 15.5 km compared to optically thin cirrus with tops ~ 12.5 km, diurnal sampling differences notwithstanding.

3) assuming fully cloudy footprints (this effect should increase with footprint size):

*Over ocean*, the peak positions corresponding to the single-layer low-level clouds lie around 850 hPa (or 1.5 km) and agree quite well between CALIPSO, MODIS-CE, AVHRR-Cloud_cci and PATMOS-x (and MISR in CZ, supplement Figure S7). Over tropical ocean, MODIS-CLDPROP finds low-level clouds equally frequent between 950 and 850 hPa, while *CIRS peaks at 950 hPa*, most probably due to *partly covered larger footprints, which include surface emission*. This effect is added to by the common presence of a temperature inversion over low oceanic clouds that affects the ancillary temperature profiles used to convert cloud top temperature to cloud top pressure (see point 4).

4) differences in the transformation between CP, CT and CZ:

Another reason for discrepancies in CP and CZ is linked to the use of different ancillary data, in particular atmospheric profiles, which are taken from HIRS observations by ISCCP and from different meteorological reanalyses (NCEP V1, NCEP GDAS, ERA-Interim, GMAO GEOS-CERES, GMAO MERRA-2, see Table 3) by the other datasets. CALIPSO and MISR do not use ancillary products in retrieving CZ.

However, it is interesting to note that both ISCCP and AVHRR-CLARA, which first retrieve CT and then derive CP, using different atmospheric profiles (HIRS and ERA, respectively), place low-level clouds too high in height, with broader peaks. This is particularly noticeable for AVHRR-CLARA over subtropical ocean, where a major part of the low-level clouds is assigned to mid-level clouds. This bias in low-level cloud heights may be due to confusion in how to use temperature to assign height in the case of a strong inversion at the top of the marine boundary layer (e.g. Holz et al. 2008).

### 3.3 Thermodynamic Phase and Radiative Properties

While cloud droplets in liquid clouds are spherical, ice crystals exist in a diversity of shapes. The refractive index of liquid and ice particles is also different. Therefore, the thermodynamic phase is important for the determination of the radiative effect of clouds as well as for the retrieval of cloud optical depth. Among other applications, cloud phase is also important for studying cloud glaciation and precipitation.

Liquid and ice clouds can be distinguished by polarization (CALIPSO), by near cloud-top temperature (ISCCP: CTI $<$ 253 K and CTW $>$ 253 K, CIRS: CTI $<$ 250 K, CTW $>$ 260 K and mixed phase clouds assumed to exist between 250 and 260 K but not considered in the I-W separation) or by use of multi-spectral information (PATMOS-x, AVHRR-CLARA, AVHRR-Cloud_cci and MODIS-CE).

Relative water (CAWR) and ice (CAIR) cloud amounts differ between the datasets, as observed in Fig. 7, but their latitudinal behaviour is similar. Comparing CAIR and CAIHR shows that while in the tropics the vast majority of the ice clouds are high-level clouds, the relative contribution of lower level ice clouds increases towards higher latitudes (in winter, supplement Figure S8).

Normalized frequency distributions of CT are compared separately for water clouds and high-level ice clouds in Fig. 8 (left). Again, the normalization is obtained by dividing the histograms by the sum of the cloudy cases building the histograms, and since the interval widths vary, each frequency has been divided so that it corresponds to a frequency per K. This is a finer binning than in CP. The distributions of IIR_CALIPSO-ST are subsampled to single-layer clouds with a successful microphysical property retrieval. For AIRS-CIRS, we show high-level ice cloud distributions for all and only for those clouds subsampled for the retrieval of bulk microphysical properties (CTIH $<$ 235 K and 0.85 $<$ CEMIH $<$ 0.2).

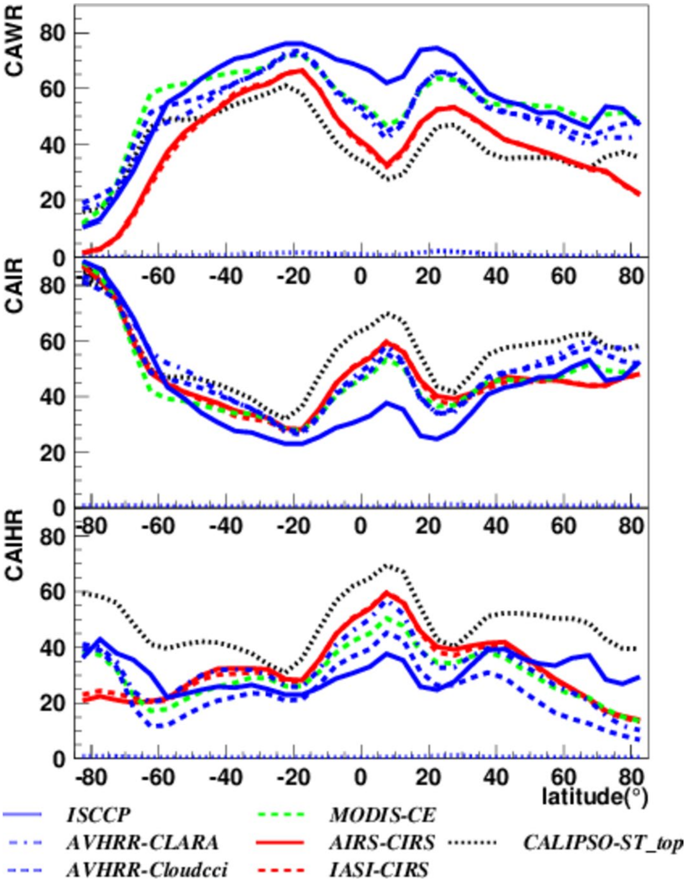


**Fig. 7** Zonal variation of CAWR, CAIR and CAIHR (in %) of seven datasets participating in the updated GEWEX cloud assessment database. Note that CAWR + CAIR ≠ 100% indicates that a fraction of the clouds have an unassigned phase (which may be interpreted as mixed phase). Statistics averaged over day and night (1:30AMPM LT, except IASI (9:30AMPM)) over the common period of 2008–2015. The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers)

CTW distributions are similar, with a peak around 285 K, slightly colder for AVHRR-CLARA and ISCCP. The CTW distribution of IIR-CALIPSO-ST shows much colder temperatures, with a peak value at 265 K, because IR microphysical retrievals tend to fail in dense and warm clouds near the surface. The CTIH distributions peak around 230 K (including IIR_CALIPSO-ST which reports radiative cloud temperatures), except for CALIPSO cloud-top temperature at 210 K and AVHRR-Cloud_cci at 240 K. Distributions of cloud-top temperatures for all high-level ice clouds (CALIPSO-ST_top) and for those with IIR_CALIPSO-ST microphysical retrievals are similar (not shown). The second peak around 195 K for ISCCP corresponds to very thin cirrus clouds, which ISCCP sets to the tropopause height. This peak is seen by CALIPSO-ST_top and not by CALIPSO-ST_passive, as expected.

CEM is retrieved at thermal wavelengths, while COD is retrieved at solar wavelengths and is therefore only available during daytime. MODIS-CLDPROP, AVHRR-Cloud_cci,

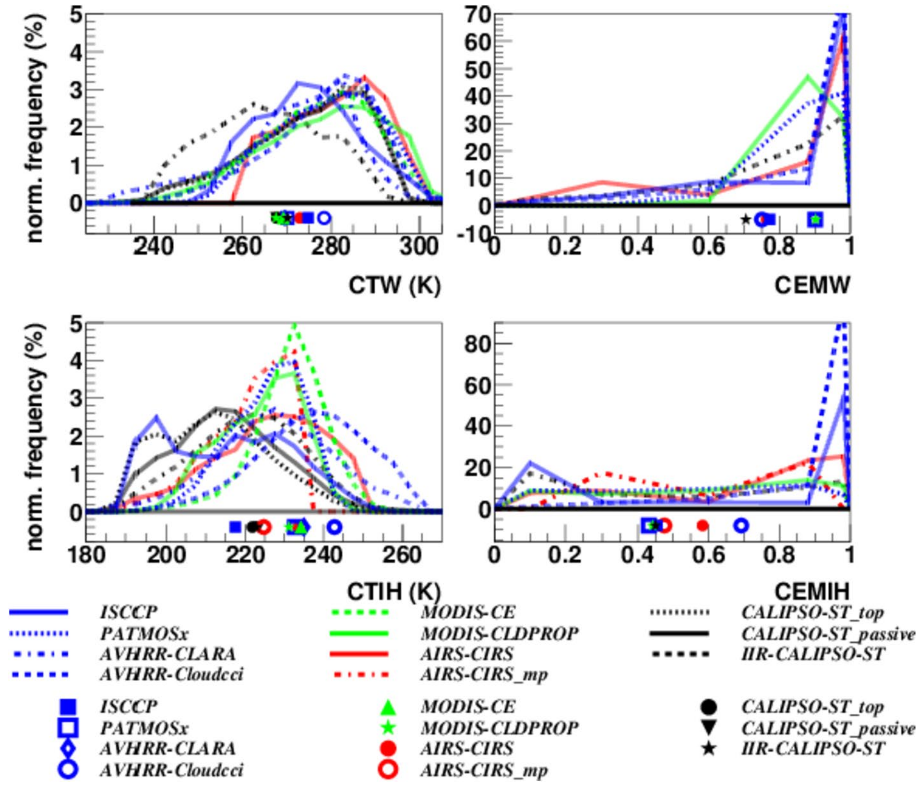


**Fig. 8** Normalized frequency distributions of radiative properties (*T* and CEM) as well as their averages (below), for liquid clouds (top) and for high-level ice clouds (bottom). The frequencies of T and CEM correspond to those per 1 K and per 0.1, respectively. Statistics is averaged over 60 N–60 S, at 1:30 PM LT. The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers)

and PATMOS-x retrieve both variables independently. ISCCP, CIRS and IIR_CALIPSO-ST only retrieve VIS COD or IR CEM, respectively, and use a conversion formula to obtain the other. AVHRR-CLARA only provides COD, and the other datasets (CALIPSO-ST, CALIPSO-GOCCP and MISR) do not provide any information on these variables.

Figure 8 (right) presents the normalized frequency distributions of CEM, separately for water clouds and high-level ice clouds. The frequencies are given per 0.1. All the CEMW distributions exhibit a peak towards large emissivity, but with a larger contribution of emissivities between 0.8 and 0.95 for MODIS-CLDPROP, PATMOS-x and IIR_CALIPSO-ST. AIRS-CIRS shows also a small contribution of emissivities smaller than 0.4. These smaller emissivities can be due to slightly overestimated cloud amount (see Sect. 3.1), and in the case of IIR_CALIPSO-ST due to the failure of the microphysical retrieval in dense clouds. The shape of the CEMIH distributions varies more. In general, there are many more semi-transparent high-level ice clouds than water clouds, as the peaks near 1 are much smaller. Only AVHRR-Cloud_cci shows a similar frequency near 1 as for CEMW and nearly no contribution of CEMIH < 0.5. Combined with the information that CTIH is also warmer than for the other datasets, the AVHRR-Cloud_cci retrieval method, based on optimal estimation with an assumption of single-layer clouds, overestimates both cloud emissivity and temperature of high-level ice clouds. The small peak of CEMIH near 0.1 of ISCCP corresponds to the very thin cirrus near the tropopause, with CTIH near 195 K, while for IIR_CALIPSO-ST these thin cirrus clouds seem to have a more variable radiative CTIH (no peak near 195 K).

Distributions of COD and bulk microphysical properties are presented and shortly discussed in the supplement (supplement Figures S9 to S12). COD distributions are difficult to present, as they are strongly skewed with some datasets showing very long tails at large COD. Though all monthly mean values are radiatively averaged, these tails yield large mean values, as for PATMOS-x, with a mean of 6.5 compared to values between 4 and 5.2 of the other datasets (supplement Figure S9). In general, the conversion of IR CEM to VIS COD underestimates the average COD, with a mean value of 2, because a CEM reaching 1 leads to a maximum COD value of about 10.

Since the distributions of all these variables are quite broad and not Gaussian, their mean values appear to be not very informative. Instead, using histograms allows one to monitor the relative changes in the distributions. This is further investigated in the next sections.

### 3.4 Relative Occurrence of Specific Cloud Types

The database also includes two-dimensional (2D) histograms in the CP-COD and CP-CEM space, respectively, which were designed for a more detailed analysis of various cloud types (i.e. Lau and Crane 1995, 1997; Rossow et al. 2005a; Tselioudis et al. 2013, 2021). Figure 9 presents the normalized relative occurrence frequency of cloud types distinguished by CP and COD in seven latitude bands, using the initial 7×7 intervals. Figures S13 and S14 present the CP-COD and CP-CEM distributions, respectively, regrouped into 3×4 intervals. Those distributions are more similar than the ones with the finer, original intervals.

As already seen in Fig. 6, the decrease of the tropopause height with latitude is visible in all datasets by a decrease in cloud height. An exception is the large proportion of very thin and high cirrus identified by ISCCP. These clouds are detected marginally in the IR and usually not in the VIS.

The ratio of optically thicker to optically thinner clouds increases towards higher latitudes, except for AVHRR-Cloud_cci, which has a large contribution in the middle COD range in the polar regions.

All datasets also show a minimum of single-layer mid-level clouds in the tropics and subtropics. Yet, by comparing datasets based on the same observations, we observe large differences in the distributions related to the retrieval method. They mostly come from multi-layered cloud fields: While the COD derived from VIS reflectances corresponds to the column optical depth, leading to correct TOA fluxes (Zhang and Rossow 2023), methods based on spectral IR alone identify the uppermost cloud and in this case the COD derived from the IR radiances corresponds to the COD of the uppermost detected cloud. This shows that these methods are complementary and their synergy should be used to describe the whole atmospheric cloud column (Sect. 4). The agreement between MODIS-CLDPROP and AIRS-CIRS is quite good, even if the footprint sizes of the two instruments are very different. *This indicates that the IR spectral information is more important than the spatial resolution to identify cirrus.*

Maps of the normalized occurrence of nine cloud types distinguished by COD and CP in Fig. 10 reveal in all datasets specific geographical patterns, such as the Stratocumulus

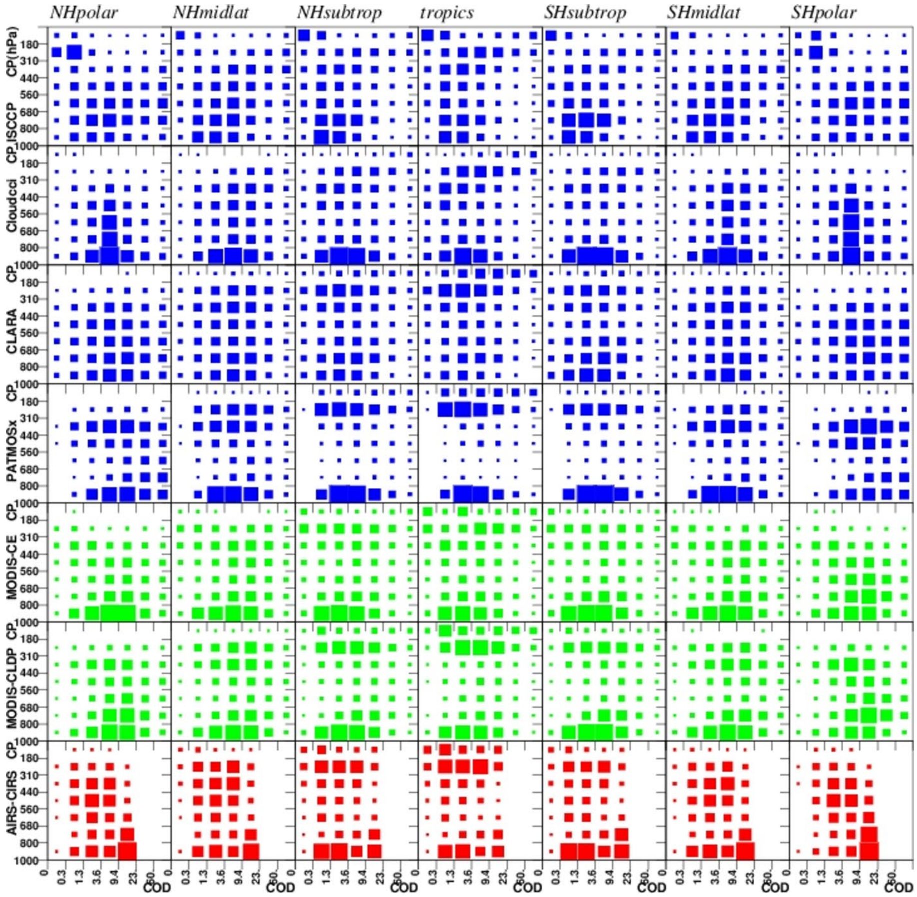


**Fig. 9** Normalized occurrence frequency of cloud types distinguished by CP and COD, for 7 latitude bands. The original binning of 7 intervals in COD and in CP is presented. Statistics averaged over 1:30PM LT, from 2008 to 2015, except for MODIS-CE (2016–2020). The size of the boxes indicates the relative frequency. Long-term datasets in blue, advanced imagers in green and IR sounder in red

regions off the Western coasts in the (COD,CP) interval (2,3), tropical cirrus anvils in the (COD,CP) interval (2,1), tropical thin cirrus in the (COD,CP) interval (1,1) or mid-latitude storm tracks in the (COD,CP) interval (2,2). Yet, some of these cloud fields appear in neighbouring COD-CP intervals as well. In particular, tropical thin cirrus and cirrus anvils also appear as mid-level clouds for ISCCP, AVHRR-Cloud_cci and AVHRR-CLARA. When considering a finer binning (not shown), the stratocumulus regions appear slightly higher and more opaque for CIRS and AVHRR-CLARA. For ISCCP they are shifted towards slightly smaller COD, while for PATMOS-x, AVHRR-Cloud_cci and MODIS-CLDPROP they are placed at larger CPs. These differences highlight the CP and COD uncertainties due to retrieval method and ancillary data. Thus, by regrouping the intervals, we were able to show a better agreement between the datasets.

The intervals with CP < 310 hPa are mainly populated by tropical and subtropical clouds. Many tropical thin cirrus are situated over the Warm Pool region. ISCCP also often

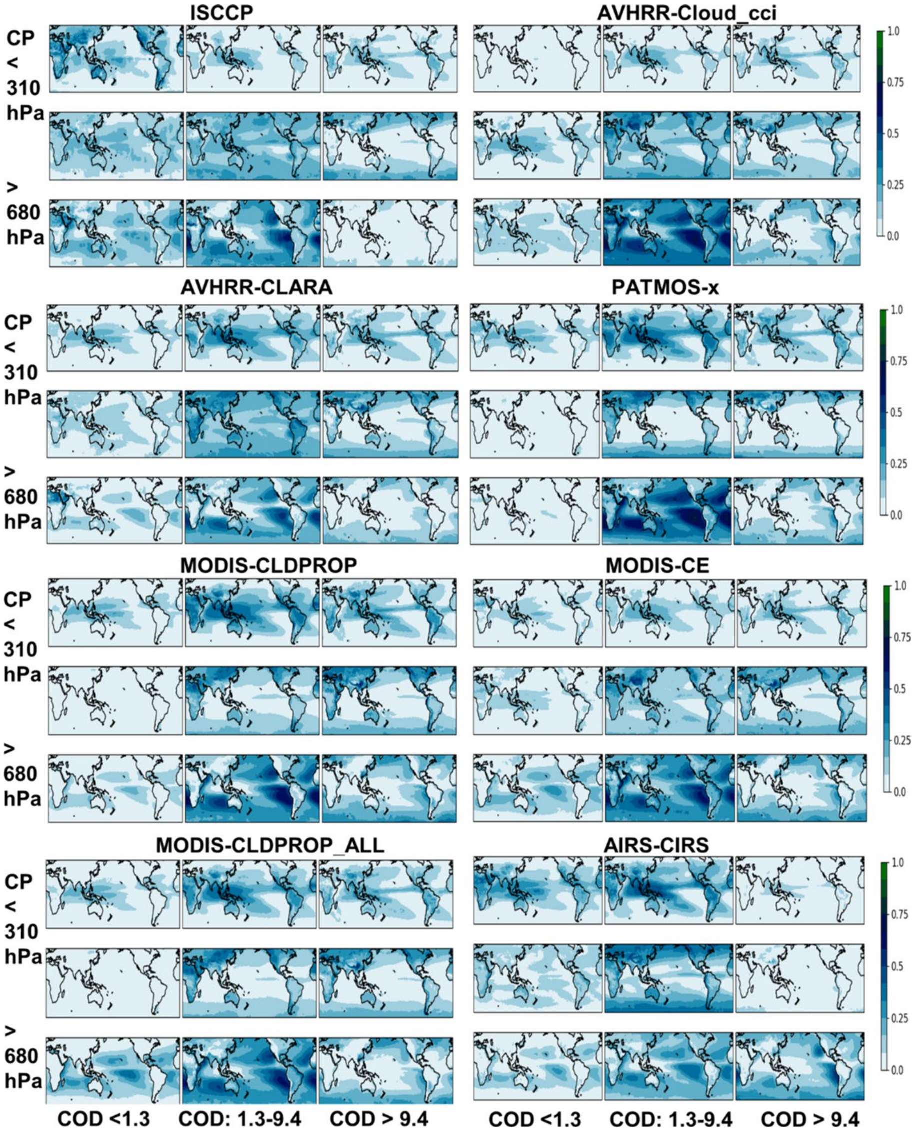


**Fig. 10** Geographical distributions (60°N-60°S) of normalized occurrence frequency of nine cloud types distinguished by COD (horizontal axis) and CP (vertical axis) from ISCCP, AVHRR-Cloud_cci, AVHRR-CLARA, PATMOS-x, MODIS-CLDPROP, MODIS-CE, MODIS-CLDPROP_ALL and AIRS-CIRS. Statistics averaged over 1:30PM LT from 2008 to 2015, except for MODIS-CE (2016–2020)

identifies this type of cloud over all continents, while the Cloud_cci retrieval method misidentifies thin cirrus as thin mid-level clouds. Both datasets seem to attribute at least a part of these tropical thin cirrus to a lower altitude ((COD,CP) interval (1,2)), whereas MODIS-CLDPROP identifies few thin clouds in the mid- to low-level range. This is explained by the fact that the MODIS-CLDPROP retrieval determines optical depth only for well-retrieved clouds, which are about 69% of all retrieved clouds. The inclusion of the other clouds (MODIS-CLDPROP_ALL) shows a relative increase in the category of thin low-level clouds, as expected. Due to the large footprint size, the CIRS cloud type occurrences are slightly shifted towards lower COD intervals, compared to MODIS-CLDPROP.

## 4 Cloud Column Structure from Complementary Information

A complete cloud column structure is desirable for evaluating climate models (e.g. Wang and Rossow 1998). First associations between the appearance of specific cloud type combinations and specific meteorological situations were identified by surface observers (e. g. Warren et al. 1985; Hahn et al. 2001). In general, random cloud overlap is a good first approximation (e. g. Chen et al. 2000). Several studies (e.g. Mace et al. 2009; Li et al. 2015; Jing et al. 2016) found slightly larger overlap between high- and mid-level clouds, especially over land, and less overlap between high- and low-level clouds, in particular over ocean at higher latitudes. Furthermore, the vertical structure depends on cloud types linked to dynamics (Naud et al. 2008) and weather states (Tselioudis et al. 2013, 2021).

Though the combination of spaceborne radar and lidar measurements provides the most complete column view, the resulting cloud statistics still depend on methods in cloud detection and cloud-type distinction. We estimate an uncertainty due to methodology by comparing results from our analyses using the CloudSat-CALIPSO GEOPROF dataset described in Sect. 2.7 and those published using CloudSat-CALIPSO CLDCLASS data (Oreopoulos et al. 2017). The latter data are based on a feature-based cloud-type classification (Sassen and Wang 2012), starting from CloudSat-CALIPSO GEOPROF data. Besides, because CloudSat and CALIPSO are flown in a single afternoon orbit with a very narrow swath width, the biases introduced by the limited diurnal sampling produce some small systematic uncertainty in the results discussed below (see also Sect. 3.1 for ocean–land contrasts).

The scheme in Table 4, introduced by Rossow et al. (2005b), illustrates cloud vertical layer structure by distinguishing seven simplified classes, based on vertical stratification in three atmospheric layers (H, M, L as described before), and associates these classes to top and column view.

The global averages summarized in Table 4 correspond to the average of two independent estimations, from our analysis using the CloudSat-CALIPSO GEOPROF data described in Sect. 2.7 and those given in Table 2 of Oreopoulos et al. 2017 using CloudSat-CALIPSO CLDCLASS data. These CloudSat-CALIPSO datasets did not use the latest version of CALIPSO data (V4.2) with improved aerosol–cloud discrimination and thus show a larger total cloud amount (0.72) than the one from CALIPSO-ST_top (0.66, Sect. 3.1). Therefore, the uncertainty coming from the retrieval method, and given in Table 4, has been estimated as the difference between the average and the single analysis values to which a 10% uncertainty has been added according to the difference between the CALIPSO versions.

The *radar–lidar data include the cloud layers hidden by those aloft*, and we can compute the amount of all associated mid-level and low-level clouds within the atmospheric column by combining the following cloud vertical structure classes to:

$$M_{\mathrm{col}} = 1M + \mathrm{ML} + \mathrm{HM} + \mathrm{HML} \;\; \text{and} \;\; L_{\mathrm{col}} = 1L + \mathrm{ML} + \mathrm{HL} + \mathrm{HML},$$

with global average $\mathrm{CAM}_{\mathrm{col}} = 0.27 \pm 0.06$ and $\mathrm{CAL}_{\mathrm{col}} = 0.50 \pm 0.10$ (CloudSat-CALIPSO).

Since the *lidar cannot probe the atmosphere below optically dense clouds*, the CALIPSO-ST_column data underestimate $\mathrm{CAM}_{\mathrm{col}}$ and $\mathrm{CAL}_{\mathrm{col}}$, with global averages of

0.18 and 0.30, respectively. We have therefore corrected these values for the amounts hidden by opaque cloud layers aloft, in guidance with the CloudSat-CALIPSO GEOPROF analysis, by using complementary information from CALIPSO-ST_opaque as well as overlap assumptions and formulas given in Appendix section "Combining CALIPSO-ST_column and CALIPSO-ST_opaque". These corrections lead to global average $\mathrm{CAM_{col}} = 0.24 \pm 0.04$ and $\mathrm{CAL_{col}} = 0.39 \pm 0.06$ (CALIPSO-ST_overlap), in closer agreement with the results from CloudSat-CALIPSO above than the initial CALIPSO-ST_column values.

As the instantaneous coverage of CALIPSO is very sparse, we also have used complementary information from passive remote sensing to estimate a complete cloud column structure. Therefore, we have combined CIRS data which identify cirrus also in the case of underlying clouds, once with data from MODIS-CE and once with data from AVHRR-Cloud_cci, both having a better spatial resolution and using additional VIS information during day which is helpful for the detection of low-level clouds.

For each 1° grid box, we choose the maximum CAH and the minimum CAL between the two considered datasets (CIRS and MODIS-CE or CIRS and AVHRR-Cloud_cci), and then, by applying cloud overlap assumptions and formulas given in Appendix section "Complementary passive remote sensing", we deduce global averages $\mathrm{CAM_{col}} = 0.25 \pm 0.03$ and $\mathrm{CAL_{col}} = 0.48 \pm 0.05$(complementary passive), in good agreement with CloudSat-CALIPSO.

A complete 3D description of clouds is also necessary in order to derive radiative heating / cooling profiles, essential for a better understanding of cloud-radiative feedbacks. While the terrestrial (LW) fluxes act primarily on the atmosphere, the solar (SW) fluxes primarily impact the surface energy budget. ISCCP provides radiative flux products (Rossow and Zhang 1995; Zhang and Rossow 2023). For their computation, a cloud vertical layer (CVL) model accounts for cloud layer overlap by assigning specific layer structures to each of the ISCCP cloud types. These types are defined by three intervals of CP and three intervals of COD (similar to the 2D histograms in Sect. 3.4), separately for each thermodynamic phase. The original model, based on the ISCCP-D version and a cloud layer climatology derived from radiosonde humidity profiles (Rossow et al. 2005b), was recently refined by combining the ISCCP-H version and CloudSat-CALIPSO GEOPROF data (Zhang and Rossow 2023). Since this model also corrects the ISCCP cloud type biases linked to multiple cloud layers, it can only be applied to ISCCP or a dataset with similar biases.

We have slightly modified the model developed for ISCCP-D and evaluated with CloudSat-CALIPSO GEOPROF data by Rossow and Zhang (2010), so that we can use the statistics within the intervals of the 2D CP-COD histograms provided in our database: We moved the lower COD threshold from 2.4 to 1.3 for ML as well as the upper COD thresholds from 2.4 to 3.6 for 1H and HL and from 9.4 to 23 for 1 M. Furthermore, we changed the thresholds to define HM from (COD 6.3–23) to (COD 3.6–9.4), and we do

**Table 5** Modified Cloud vertical layer model for ISCCP cloud types as function of CP and COD. The original model (Rossow et al. 2005b) assumes ISCCP mid-level clouds (440–680 hPa) as 1 M, HL, HL, ML, instead of HL, ML, ML, ML / 1 M in the four COD intervals, respectively

| | | | | |
|---|---|---|---|---|
| < 440 hPa | **1H** | **1H** | **HM** | **HML** |
| 440–680 hPa | **HL** | **ML** | **ML** | **ML 1 M** |
| > 680 hPa | **1L** | **1L** | **1L** | **1L** |
| CP / COD range | 0–1.3 | 1.3–3.6 | 3.6–9.4 | > 9.4 > 23 |

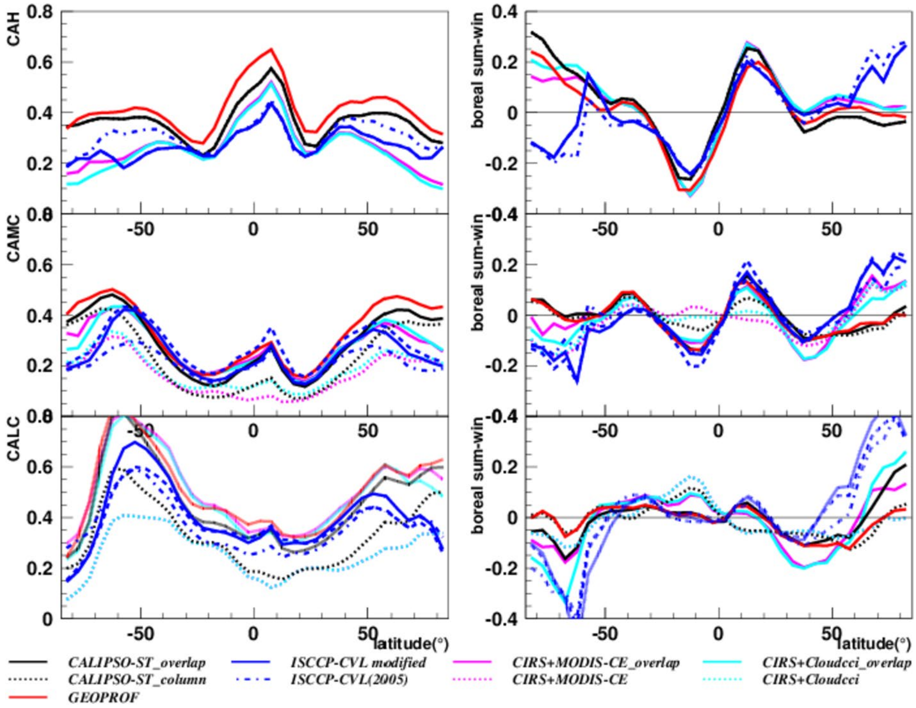


**Fig. 11** Left panel: Zonal mean values of CAH, $CAM_{col}$ and $CAL_{col}$ of GEOPROF, CALIPSO-ST_overlap, complementary passive (CIRS & MODIS-CE_overlap and CIRS & AVHRR-Cloud_cci_overlap) and of ISCCP corrected with different CVL models. CAM and CAL of CALIPSO-ST_column and complementary passive are also shown. These are uncorrected, while the overlap values include all clouds hidden by clouds aloft. Right panel: Differences of CAH, $CAM_{col}$ and $CAL_{col}$ between boreal summer (June, July, August) and boreal winter (December, January, February). Statistics averaged 1:30AMPM LT (except ISCCP, 1:30PM LT, during daytime) over the common period of 2008–2015 (except GEOPROF 2007–2009)

not distinguish between multiple layers of HML (COD 2.4–6.3) and one contiguous layer HxMxL (COD > 23), and simply define HML/HxMxL by COD > 9.4. This modified CVL model is summarized in Table 5.

We have applied this model on the statistics of the CP-COD histograms to deduce global averages CAH = 0.29, $CAM_{col} = 0.23 \pm 0.03$ and $CAL_{col} = 0.41 \pm 0.05$ (ISCCP-CVL).

The latter values agree well with the others noted above.

Figure 11 compares the latitudinal variation of CAH, $CAM_{col}$ and $CAL_{col}$ from CloudSat-CALIPSO GEOPROF, CALIPSO-ST_column, CALIPSO-ST_overlap (derived from column and opaque), complementary passive (CIRS-MODIS-CE and CIRS-AVHRR-Cloud_cci) and ISCCP-CVL. Zonal means of their differences between boreal summer and boreal winter are also reported.

Since CAH of the complementary datasets is the maximum of both, which is most often CIRS, it is relatively close to CALIPSO-ST_top, with less agreement towards the poles due to different height categories according to retrieval of radiative versus cloud-top height (Sect. 3.2). The bias-corrected CAH of ISCCP-CVL also agrees well with the other datasets, but the peak in the tropics is still slightly lower. CAH of GEOPROF is slightly larger than of CALIPSO-ST_top, because the latter has a reinforced aerosol detection.

The amounts of all mid- and low-level clouds within the atmospheric column, $CAM_{col}$ and $CAL_{col}$, agree very well among the datasets. Only ISCCP indicates a slightly smaller $CAL_{col}$ at higher latitudes, particularly between 50 and 65°S. In this region the ISCCP $CAL_{col}$ also differs more with respect to the assumptions in the two CVL models (new one and the one of Rossow and Zhang 2005b). The comparison of CALIPSO-ST_overlap (solid) with the original CAM and CAL of CALIPSO-ST_column (dotted) demonstrates that the correction of CALIPSO-ST_column due to partial obscuration by opaque clouds aloft is not negligible, particularly for low-level clouds. By comparing CAL (dotted) and $CAL_{col}$ (solid) of the complementary datasets, we deduce that *about half of the low-level clouds are obscured by clouds aloft*.

The seasonal differences are now easier to interpret: $CAM_{col}$ and even $CAL_{col}$ follow the displacement of the ITCZ, because the high-level convective clouds in the ITCZ fill the whole atmospheric column and the cloud fields around convection are complex with several layers of clouds. At the high latitudes, we observe more low-level clouds in summer than in winter. This effect is much smaller when considering only single-layer clouds (CAL, Figure S5).

## 5 Monitoring of Cloud Amount and of Relative Occurrence of Specific Cloud Types

Some of these global cloud climate data records now cover up to 40 years. Time series of the AVHRR-based datasets have already been compared with ISCCP by Karlsson et al. (2018). They found a weak negative trend in global cloud amount from all datasets over the period 1984–2009, but also noticed clear signs of various discontinuities related to satellite shifts and orbital drift effects. They state that only after the introduction of the AVHRR/3 sensors from 2001 onwards, all AVHRR Climate Data Records show more consistent results among them, for the afternoon orbits. In the supplemental Figure S15, we present the time series of the global mean CA, separately for all four observation times given in this database. It is interesting to note that the satellite drifts of the afternoon satellites are the most strongly pronounced at 1h30 PM LT, in particular for PATMOS-x.

Marvel et al. (2015) developed a multi-variate fingerprint that captures coherent, externally forced cloud changes in order to reduce the time to detect these changes. By using ISCCP and PATMOS-x data, they could still demonstrate the poleward migration in the major latitudinal features of total cloud amount. Bender et al. (2012) and Norris et al. (2016) removed artefacts of ISCCP and PATMOS-x by filtering problematic regions, spurious viewing geometry effects and specific cloud types, or by using empirical corrections, respectively. Their analyses reveal large-scale patterns of cloud change between the 1980s and the 2000s consistent with poleward retreat of mid-latitude storm tracks, expansion of subtropical dry zones, and increasing height of the highest cloud tops at all latitudes.

So far, the original ISCCP provides the best temporal sampling. Since our database does not completely resolve the diurnal cycle with four measurements per day, we concentrate in the following only on variabilities at 1h30 PM LT, as for the analysis of the specific cloud types in Sect. 3.4. Recently, Liu et al. (2023) found weak opposing trends of CA over land and over ocean under global warming, but only by using an EOF analysis which was able to separate the large signals from El-Niño-Southern-Oscillation-associated (ENSO-associated) modes in the 42 years of meteorological reanalysis data.

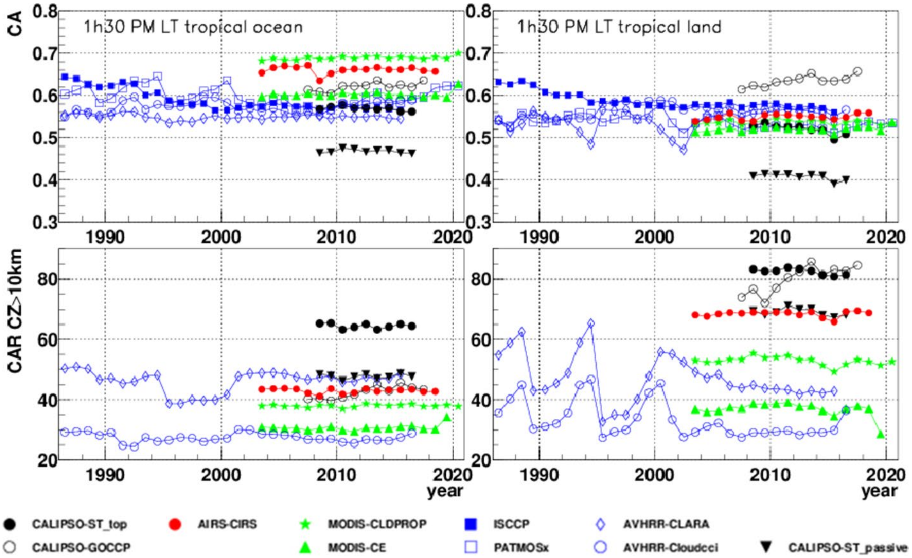


**Fig. 12** Time series of cloud amount (top) and relative proportion of clouds with CZ > 10 km (bottom), over tropical ocean (left) and tropical land (right). Statistics at 1:30PM LT over 25N-25S. The colours correspond to specific instrument types (black: active lidar, red: IR sounders, green: advanced multi-spectral imagers, blue: multi-spectral imagers)

### 5.1 Interannual Variability

The upper panel of Fig. 12 shows annual values of cloud amount for all datasets from 1986 to 2020, averaged over the latitudinal band 25 N–25 S, separately over ocean and over land. Before 2003, the AVHRR-based time series show inter-annual variations which do not coincide, without any clear trends. The three AVHRR cloud datasets use the same radiance calibration, but at this particular time of day the cloud detection of PATMOS-x (ocean) and CLARA (land) seems to be more sensitive to changes in observation time due to orbital drifts than the one by Cloud_cci. ISCCP, mostly based on geostationary imager data, shows a nearly linear decrease in cloud amount from 1986 to 2000, both over ocean and over land. The negative trend has already been thoroughly investigated for the earlier ISCCP version (Appendix 2 by W. B. Rossow in Stubenrauch et al. 2012) and also for the current ISCCP version (Rossow et al. 2022): this systematic slow decrease of CA may be slightly overestimated but it cannot be explained alone by spurious effects linked to calibration, viewing geometry changes, changes in land–ocean and day-night sampling. This conclusion is also supported by a quantitative agreement between the TOA radiative fluxes, calculated from ISCCP cloud properties and those determined from ERBS observations (Zhang et al. 2004; Norris 2005). The global cloud amount (Figure S15) has a very similar behaviour. Concerning the newer instruments, it is interesting to note that both CALIPSO datasets show a slightly different behaviour, in particular over land. The time series show in general a flat behaviour, which shows that one needs a good filtering in order to observe any changes due to specific climate modes.

In order to illustrate another application of the data, we derive the relative proportion of clouds above 10 km from the CZ histograms. The histograms allow us to define

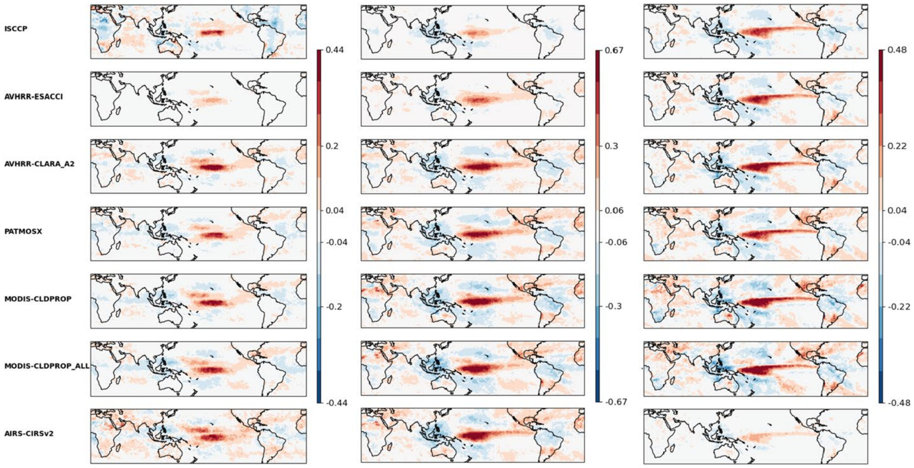


**Fig. 13** Geographical maps of linear regression slopes between monthly mean anomalies in amount of thin cirrus (left), cirrus (middle) and high opaque clouds (right), *relative to all clouds*, and global mean surface temperature anomalies. Units are in fractional change per °C warming. Cloud data are from ISCCP, AVHRR-Cloud_cci, AVHRR-CLARA, PATMOS-x, MODIS-CLDPROP_ALL, MODIS-CLDPROP and AIRS-CIRS (from top to bottom), and surface temperature from the GISS Surface Temperature Analysis, during the period 2003–2015, 40N-40S

cloud types by different thresholds, depending on the subject of study. The lower panel of Fig. 12 shows again that the data are less variable after 2003. Both CLARA and Cloud_cci have large variations over land, most probably due to the NOAA afternoon satellite drifting. In addition, CLARA also displays an abrupt decrease over ocean in 1995, due to an increase in semi-transparent clouds (see next section and supplement Figure S17). Again, the behaviour of the two CALIPSO datasets is different over land, with CALIPSO-ST exhibiting less variability.

IASI-CIRS and MISR results are presented together with the other morning measurements in Figure S16 in the supplement. Except PATMOS-x, which shows a drop in CA after 2010, the data from the morning satellites are again quite stable after 2002.

## 5.2 Anomaly Patterns with Respect to Global Surface Temperature Changes

Since the AVHRR data records are less consistent before 2000 and since the observational period of the new satellite era is still too short to directly study long-term cloud variability related to climate warming, an alternative approach is to analyse cloud type variability with respect to inter-annual global mean surface temperature anomalies. For this analysis, we use deseasonalized global monthly mean surface temperature anomalies from the GISS Surface Temperature Analysis (GISTEMP Team 2023; Lenssen et al. 2019).

We use the 2D histograms in the CP-COD space to select upper tropospheric cloud types with $\mathrm{CP} < 310$ hPa and distinguish these further into thin cirrus, cirrus and high opaque clouds, using the same three COD ranges as in Fig. 10. Then, we determine geographical change patterns in amount of these cloud types, relative to all clouds, with respect to global surface temperature: $d(\mathrm{CA}_{\mathrm{type}}/\mathrm{CA})/\mathrm{d}T_{\mathrm{s}}$. These derivatives are computed for each grid box of 1° latitude × 1° longitude, by linear regression between their anomalies

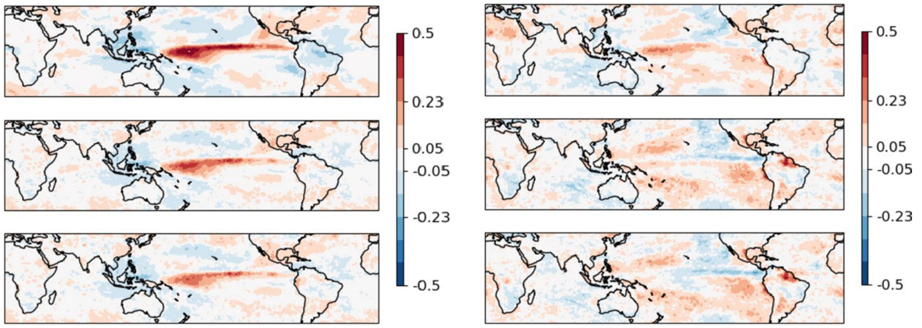


**Fig. 14** Geographical maps of linear regression slopes between monthly mean anomalies in $CAM_{col}$ (left) and $CAL_{col}$ (right) and global mean surface temperature anomalies. Units are in fractional change per °C warming. The column amounts have been obtained from complementary information, as discussed in Sect. 4, using the CVL model for ISCCP (top), and using AIRS-CIRS combined with AVHRR-Cloud_cci (middle) or MODIS-CE (bottom). Surface temperature from the GISS Surface Temperature Analysis, during the period 2003–2015

as in (Stubenrauch et al. 2017). Results are presented in Fig. 13 for the seven datasets which provide these 2D histograms over the common time period 2003–2015. All datasets show very similar geographical change patterns, with an increase per °C of warming at the equator in the tropical Pacific and a decrease over Indonesia, a pattern similar to ENSO (e.g. Stephens et al. 2018b). Because of the large footprint (15 km), the change in amount of high opaque clouds identified by AIRS-CIRS corresponds to the one of large convective areas. Their amount is much smaller than the one of convective areas identified by a smaller spatial resolution (Fig. 10). Smaller changes in cirrus by AVHRR-Cloud_cci and ISCCP can be explained by misplacing their heights. We also observe an increase in cirrus and thin cirrus in the subtropics. This may be in relation to a hypothesized widening of the Hadley cell with global warming (e.g. Staten et al. 2018). The relative amount of thin cirrus also increases around the equator, close to the increase of high opaque clouds. This can be reconciled with the findings that tropical mesoscale convective systems get higher (e. g. Stubenrauch et al. 2021) and the area of surrounding thin cirrus gets larger with their height (Protopapadaki et al. 2017). This signal decreases with decreasing sensitivity to thin cirrus, the lowest signal being given by AVHRR-Cloud_cci.

The different geographical change patterns of high opaque clouds and thin cirrus lead to variations in atmospheric heating and cooling gradients which then influence the large-scale circulation, as has already been investigated by Slingo and Slingo (1991).

It is interesting to note that the change patterns are similar when exploring a longer time period (2003–2018 for MODIS-CLDPROP and AIRS-CIRS, and 1995–2015 for the AVHRR datasets and ISCCP), as shown in supplemental Figure S17, except for AVHRR-CLARA exhibiting artificial increases of cirrus and thin cirrus from 1995 to 2001.

Changes in column cloud amount of mid- and low-level clouds with respect to global surface temperature are displayed in Fig. 14. We have used the complementary information and computations discussed in Sect. 4. The change patterns, both in $CAM_{col}$ and in $CAL_{col}$, are very similar among the different methods. The increase in column mid-level clouds in the Pacific along the equator is coherent with the change pattern of high opaque cloud amount. This indicates increasing tropical deep convective cloud amount in the tropical Pacific close to the equator, with warming. All data agree that the column low-level cloud amount decreases in the stratocumulus regions off the Californian coast and off the west coast of Australia, while it increases off the west coasts of South America and Africa as well as north and south of the increasing deep convective band in the Western Pacific.

To summarize: It is very challenging to detect small inter-annual global and zonal changes with these data, in particular with long-term datasets, which have undergone orbit drifts, instrument changes and orbit changes. However, similar geographical change patterns between the different datasets demonstrate that the data can be used to study climatological change patterns, in particular, those for specific cloud types (considering the sensitivity of a specific dataset).

## 6 Summary and Outlook

Since the original GEWEX cloud assessment a decade ago, existing cloud property retrievals have been revised and/or temporally extended, and additional retrievals from new sensor data have emerged.

In general, results have remained similar to those of the previous assessment. A notable exception is a reduced cloud amount derived by CALIPSO-ST, which is associated with an improved distinction between aerosol and clouds. Therefore, the global mean cloud amount $0.66 \pm 0.04$, estimated from the eleven datasets of the updated GEWEX cloud assessment database, is slightly lower than before ($0.68 \pm 0.03$). Both these estimations have additional uncertainties due to their limited diurnal sampling.

The absolute values of high-level cloud amount depend on instrument and retrieval sensitivity to thin cirrus, varying between 0.12 (MISR, visible information only) and 0.37 (CALIPSO-ST, active lidar). By exploiting the different CALIPSO data in combination with CloudSat-CALIPSO, we estimated the amount of thin cirrus with $\mathrm{COD} < 0.3$ roughly to 0.13, of which about 60% occur with lower clouds underneath, when assuming that multiple cloud layers are COD-independent of the higher cloud. This larger contribution compared to a much earlier result of about 30% obtained for cirrus with $\mathrm{COD} < 1$, using HIRS observations (Jin and Rossow 1997), can be partly explained by the different spatial resolutions (1° compared to 17 km) and by the assumption of COD-independent overlap made in our estimation. When the retrieval method includes IR and VIS information during day, without considering the possibility of underlying clouds, thin cirrus may be mis-identified as a more opaque lower cloud because of the large VIS reflectance from the cloud underneath.

*While the absolute values of the retrieved variables show systematic differences due to retrieval method, instrument sensitivity, spectral and ancillary data and footprint size, their geographic, synoptic and seasonal variations agree well.* Only the polar regions remain a challenge for passive remote sensing. Another issue is the difference between the radiative cloud height, retrieved by passive remote sensing (except stereoscopic height retrieval by MISR) and the cloud top height from active remote sensing. At higher latitudes, the radiative height starts to decrease towards the middle troposphere, while the cloud top height still stays in the category of the upper troposphere.

The three datasets based on AVHRR observations differ strongly. Two of these datasets have been recently reprocessed: PATMOS-x (Foster et al. 2023) and AVHRR-CLARA (Karlsson et al. 2023). Both new versions show a global CA increased by 0.03, with PATMOS-x changing from 0.65 to 0.68 and AVHRR-CLARA from 0.62 to 0.65. The AVHRR-Cloud_cci dataset is not continued.

For the determination of cloud impacts on the radiative heating profile not only altitude-stratified cloud amount but also associated values of attenuation (optical depth, emissivity) and composition (particles size, phase) matter. The presented diversity of these cloud properties is also associated with the strengths and limitations of the retrieval methods and assumptions, so that a combination of sensor capabilities is recommended.

In this article, we have highlighted how to use this updated GEWEX cloud assessment database, which includes many variables and monthly statistics in the form of averages and histograms. In particular, the two-dimensional histograms in the CP-COD or CP-CEM space allow a more detailed cloud type specification. Differences in the shape of the normalized distributions indicate the CP and COD uncertainties due to retrieval method and ancillary data. By regrouping to coarser intervals, we were able to show a better global agreement between the datasets, which indicates that the uncertainties are larger than the original intervals given in the CP-COD and CP-CEM histograms.

In spite of large uncertainties, these data are very valuable for cloud climate and process studies via analyses which cluster statistics in the CP-COD space, leading to the definition of Weather States or Cloud Regimes (e.g. Rossow et al. 2005a; Tselioudis et al. 2013; Gryspeerdt and Stier 2012; Stachnik et al. 2013; Luo et al. 2017; Tan and Oreopoulos 2019) or which construct cloud systems from adjacent measurements (e.g. Protopapadaki et al. 2017; Stubenrauch et al. 2021, 2023). Furthermore, by including radiative fluxes, as for example directly available for ISCCP (Zhang and Rossow 2023) and AVHRR-Cloud_cci (Stengel et al. 2020), cloud radiative effects may be quantified. The latter data have been used by Philipp et al. (2020) to analyse Arctic feedback mechanisms between sea ice and low-level clouds.

While passive remote sensing can address radiative (climate) impacts at the top of the atmosphere (Zhang et al. 2004), those in the atmosphere and at the surface require information on cloud-base altitude and cloud amounts hidden by higher clouds aloft, which only space-borne active radar–lidar remote sensing can offer. Still, the uncertainty in the retrieval method applied to these active data is about 10 to 20%. We have shown that insights into these properties are possible by combining complementary passive sensor data (like IR sounder and IR-VIS imager data) and an empirical overlap model or by applying a bias correction and vertical layer model specifically developed for ISCCP (Zhang and Rossow 2023) to the 2D histograms of ISCCP. Furthermore, these cloud datasets can be combined with other measurements for diagnostic purposes.

For IR-VIS radiometers, we recommend developing retrieval methods (during daytime) which are able to distinguish single- and multi-layered cloud scenes. This may be achieved by first using the IR radiances and their spectral differences alone and then comparing the retrieved cloud properties to those obtained when including VIS reflectances (e.g. Chang and Li 2005). Machine-learning methods that use VIS-IR radiances, single-layer retrievals, and ancillary data as input are another effective approach to identify overlapped clouds (e.g. Sun-Mack et al. 2023) and retrieve some of their properties (e.g. Minnis et al. 2019). It is also beneficial to combine radiances of different instruments on the same platform for a more sophisticated retrieval. Still, combining different sensors might come at a cost, e.g. in terms of spatial coverage (swath width) or temporal consistency (instrument issues). First steps in this direction have recently emerged, including the new version of the PATMOS-x cloud retrieval (Foster et al. 2023), which uses combined AVHRR and HIRS measurements, and a new cloud retrieval technique for two-layered cloud systems using the fusion of MISR VIS-stereoscopy and MODIS IR radiances (Mitra et al. 2023).

Reliable Climate Data Records require calibration between instruments and homogeneous ancillary data. EUMETSAT and NOAA/NESDIS are collaborating on generating a new GEO-RING set of data based on the recent improvements in the geostationary imager capabilities (from 2018 onwards). A next generation of the International Satellite Cloud Climatology Project (ISCCP-NG) is being developed as an application of this new GEO-RING data. The AVHRR retrieval teams in the USA and Europe are already generating prototype cloud products from ISCCP-NG. This data will have spatial resolution of 0.05° and temporal resolution of 30 min. The final resolutions are still to be determined. The goal is to have ISCCP-NG generating routine data once the Third Generation of Meteosat (MTG) has finished its first year of operations.

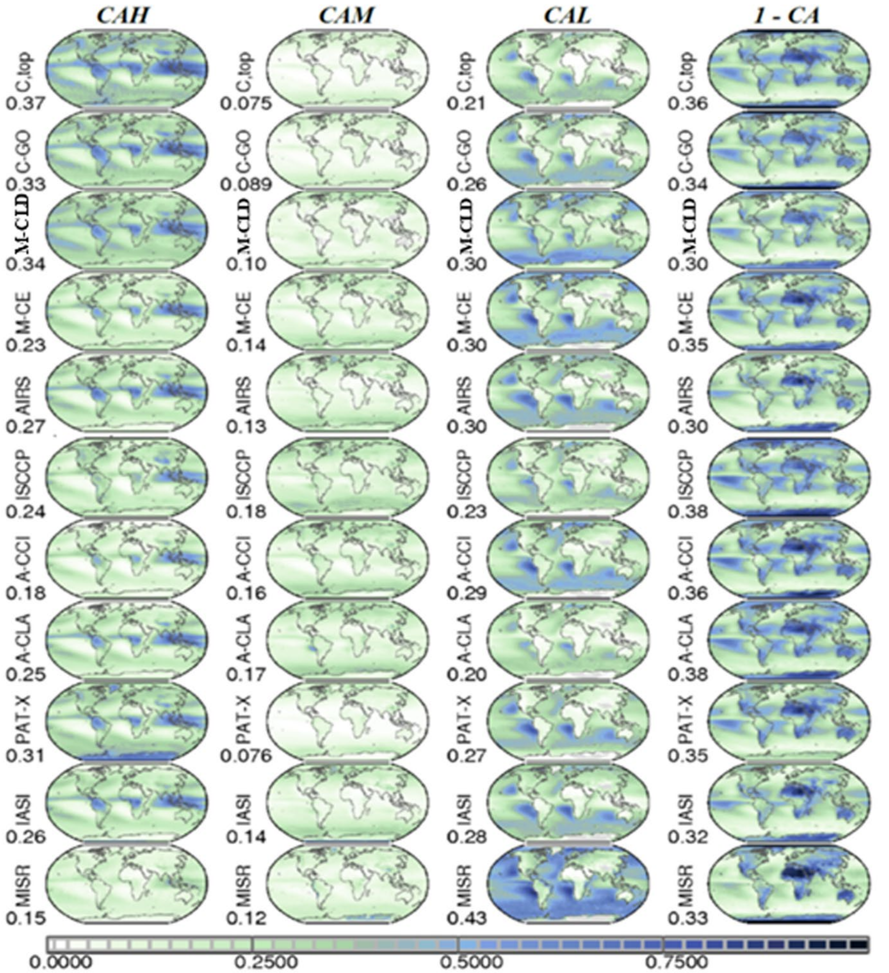
Furthermore, we strongly recommend the use of complementary information as much as possible. ISCCP employs a cloud vertical layer model based on complementary information from radar–lidar (CloudSat-CALIPSO) to compute the cloud-induced radiative fluxes, and we have proposed a scheme which combines IR sounder and VIS-IR imager cloud amounts and an empirical cloud overlap scheme to get a full cloud amount within the atmospheric column. This method assumes increasing overlap with increasing cloud emissivity of the clouds aloft and decreasing vertical distance. Both approaches should be applied on instantaneous data, but we saw also an improvement in comparisons with the CloudSat-CALIPSO data when applied to the monthly data.

As published before, global and zonal long-term trends are only coherent among the AVHRR datasets after 2002, when all cloud amounts become more stable. The decrease in total cloud amount from ISCCP from 1986 to 2018 can be explained by changes in particular cloud types (Rossow et al. 2022; Zhang and Rossow 2023). The more capable satellite sensors available since 2000 do not show any statistically significant trends in global and zonal means. Yet, regional changes in the relative occurrence of specific cloud types as well as in column low-level cloud amount can be well identified with respect to global surface temperature anomalies.

## Appendix 1: Geographical Patterns of the Individual Datasets

The geographical distributions of annual averages of CAH, CAM, CAL and clear sky amount (1-CA) are displayed in Fig. 15 for each of the twelve individual datasets.

*While the absolute values differ, the geographical patterns obtained from these data are very similar*. All data sources show the Inter-Tropical Convergence Zone (ITCZ) as the most prominent feature in CAH. Other important features are large CAL in the stratocumulus regions off the western coasts in the subtropics and over ocean in the mid-latitudes, especially in the SH. The clear sky amount (1-CA) distributions also agree well and show prominent features over the deserts and the subtropical ocean, in particular where one expects open cell cumulus clouds. Another common feature comprises the storm tracks in the zonal westerlies, which are evinced by the relatively large CAH and CAM means. In general, we observe more CAM at high than at low latitudes. This effect is stronger for the passive remote sensing, because when the tropopause is lower (down to 300 hPa) the radiative cloud height is closer to the constant threshold of 440 hPa and therefore some of the high-level clouds fall into the category of mid-level clouds (see Sect. 3.2).



| | | |
|---|---|---|
| *CALIPSO-ST top :* ***C,top*** | *MODIS longterm :* ***M-CP*** | *ISCCP :* ***ISCCP*** |
| *CALIPSO-GOCCP :* ***C-GO*** | *AIRS-CIRS :* ***AIRS*** | *AVHRR-Cloud_cci :* ***A-CCI*** |
| *MODIS-CLDPROP:* ***M-CLD*** | *IASI-CIRS :* ***IASI*** | *AVHRR-CLARA :* ***A-CLA*** |
| *MODIS-CE :****M-CE*** | *MISR :* ***MISR*** | *PATMOS-x :* ***PAT-X*** |

**Fig. 15** Geographical maps of annual averages of CAH (column 1), CAM (column 2), CAL (column 3) and (1-CA) (column 4) from eleven datasets of the updated GEWEX cloud assessment database. Statistics 2008–2015. The values to the lower left indicate global (area-weighted) averages

## Appendix 2: Estimation of Cloud Column Structure from Complementary Information

### Combining CALIPSO-ST_column and CALIPSO-ST_opaque

Since the lidar cannot probe the atmosphere below optically dense clouds, the CALIPSO-ST_column dataset underestimates $CAM_{col}$ and $CAL_{col}$ by about 25–30%, in comparison to the results from combined CloudSat–CALIPSO data in Table 6. The

**Table 6** Global averages of CAM and CAL, including the parts hidden by clouds aloft, separately over ocean and over land. Compared are results from CloudSat-CALIPSO (CC) (analysis described in Sect. 2.6 and by Oreopoulos et al. 2017), from CALIPSO-ST_column and CALIPSO-ST_overlap, and from complementary passive remote sensing (CP)

| | | ocean | | Land | | |
|---|---|---|---|---|---|---|
| | CC | C | CP | CC | C | CP |
| CAM(column) | | 0.16 | | | 0.23 | |
| $\text{CAM}_{\text{col}}$ | 0.26 ± 0.06 | 0.19/**0.23**/ 0.25 | 0.25/0.25 | 0.31 ± 0.04 | 0.26/**0.29**/0.31 | 0.26/0.26 |
| CAL(column) | | 0.37 | | | 0.15 | |
| $\text{CAL}_{\text{col}}$ | 0.58 ± 0.08 | 0.44/**0.50**/0.53 | 0.55/0.54 | 0.33 ± 0.06 | 0.19/**0.29**/0.32 | 0.32/0.30 |

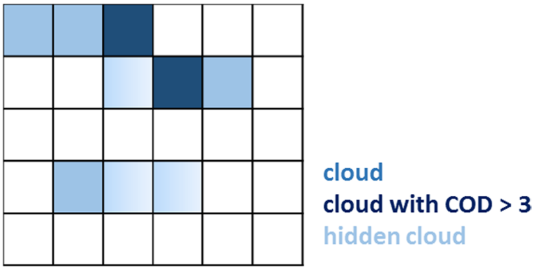
CC: averaged results from CloudSat GEOPROF-lidar and CloudSat CLDCLASS-lidar analyses

C: CALIPSO-column CAL, CAM

C: $\text{CAL}_{\text{col}}$, $\text{CAM}_{\text{col}}$ corrected for clouds hidden by opaque ones aloft under three different overlap assumptions (random, **random to maximum** and maximum overlap)

CP: complementary passive remote sensing (CIRS-MODIS-CE), under two different overlap assumptions (random and **increasing overlap with CEMH)**

**Fig. 16** Scheme of cloud overlap, with parts hidden from CALIOP by opaque cloud layers aloft



latter are given as averages from the results of 1) Oreopoulos et al. (2017), using cloud top and base heights from CloudSat-CALIPSO CLDCLASS data (Sassen and Wang 2012) and transforming these to pressure coordinates via ECMWF auxiliary data, and of 2) our analysis of CloudSat-CALIPSO GEOPROF data described in Sect. 2.6. The 10–15% differences in Table 6 may be due to the use of CP and CZ for the height-stratification and due to the different definition of cloud types in the original data.

We estimate $\text{CAM}_{\text{col}}$ and $\text{CAL}_{\text{col}}$ by correcting the CALIPSO-ST_column amounts (CAM(column) and CAL(column)) for those hidden by opaque clouds aloft (Fig. 16). CAH(opaque) (0.09) contains the vertical structures HM and HML, and CAM(opaque) (0.07) corresponds to ML (Table 4). We use formulas similar to those of Ham et al. (2021)

$$\begin{aligned}\text{CAM}_{\text{col}} &= a\text{CAH(opaque)} + \text{CAM(column)}/(1-(1-a)\,\text{CAH(opaque)}) \text{ and}\\ \text{CAL}_{\text{col}} &= a\,(\text{CAH(opaque)} + \text{CAM(opaque)})\\ &\quad + \text{CAL(column)}/(1-(1-a)\,(\text{CAH(opaque)} + \text{CAM(opaque)}))\end{aligned}$$

under three different overlap assumptions: (1) random ($a=0$), (2) *between random and maximum* ($a=0.75$), and (3) maximum ($a=1$). The computation of $\text{CAL}_{\text{col}}$ is similar, but it considers high- and mid-level opaque clouds aloft with the assumption of minimum

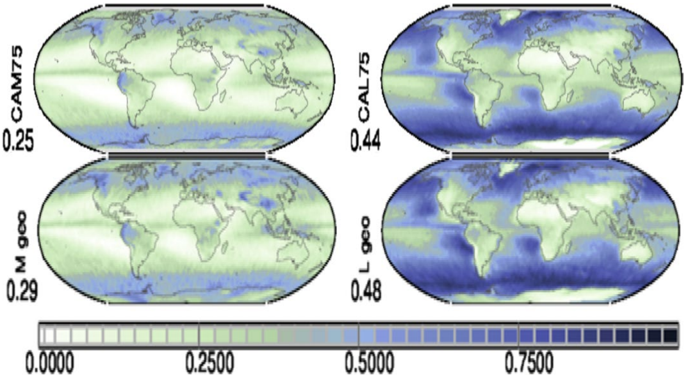


**Fig. 17** Geographical patterns of annual averages of $CAM_{col}$ (left) and $CAL_{col}$ (right), deduced from the above formulas applied on CALIPSO-ST (top) and from CloudSat-CALIPSO GEOPROF (bottom). Statistics 2007–2009. The values to the lower left indicate global (area-weighted) averages

overlap between these two. The uncertainties in $CAM_{col}$ and $CAL_{col}$ linked to overlap assumptions are about 10 to 15%, according to Table 6.

The maps in Fig. 17 reveal that the CALIPSO-ST $CAM_{col}$ and $CAL_{col}$ averages are very similar to their CloudSat-CALIPSO GEOPROF counterparts. The larger global averages can be explained by an earlier CALIPSO version used in these GEOPROF data. Strictly speaking, this method should be applied to the instantaneous gridded data. Nevertheless, it also works well on the monthly data, as demonstrated in Fig. 17.

### Complementary Passive Remote Sensing

To estimate $CAM_{col}$ and $CAL_{col}$ from passive remote sensing, we need a consistent definition for cloud layer height regimes (CAH, CAM and CAL) and assumptions about how they overlap. To minimize mis-location of thin clouds into lower height regimes, a 'best estimate' is suggested by *combining the advantages of IR sounders* (with best IR spectral resolution and hence a good sensitivity to thin cirrus: e. g. AIRS, IASI), *and the advantages of multi-spectral sensors* (with a better spatial resolution for the retrieval of smaller scale low-level cloud properties: e. g. MODIS-CE, AVHRR_Cloud_cci).

Among the two complementary datasets we choose CAH, CAM and CAL for each 1° grid cell at a specific observation time as follows:

$$CAH = \max(CAH1,\ CAH2) CAL = \min(CAL1,\ CAL2)$$

$$CAM = \max(CAH1 + CAM1,\ CAH2 + CAM2) - CAH$$

For estimates of the complete (including hidden) cloud amount at mid- ($CAM_{col}$) and low- ($CAL_{col}$) levels, we have developed an empirical formulation that considers the cloud emissivity of the high- (CEMH) and mid- (CEMM) level clouds aloft, by assuming that the probability of cloud overlap increases with increasing cloud emissivity and it decreases with increasing vertical distance:

$$
\begin{aligned}
\mathrm{CAM_{col}} &= a\mathrm{CAH} + \mathrm{CAM}/(1{-}(1-a)\,\mathrm{CAH}) \quad \text{with } a = \mathrm{CEMH} \ast\ast\ 2\\
\mathrm{CAL_{col}} &= b\mathrm{CAH} + c\mathrm{CAM} + \mathrm{CAL}/(1{-}(1-b)\,\mathrm{CAH} - (1-c)\,\mathrm{CAM}).\\
\text{with } c &= \mathrm{CEMM} \ast\ast\ 2, b = 0.75\ \mathrm{CEMH} \ast\ast\ 2
\end{aligned}
$$

These formulas allow again random ($a$, $b$, $c=0$) to maximum cloud layer overlap ($a$, $b$, $c=1$). By comparing geographical distributions with CloudSat-CALIPSO GEOPROF results we developed an empirical formula, assuming *increasing overlap with increasing emissivity of the higher clouds aloft* ($a=\mathrm{CEMH}^2$) *and decreasing overlap with increasing vertical distance between the cloud layers* (factor 0.75). These formulas lead to the global averages $CAM_{col}=0.26$ and $CAL_{col}=0.48$.

This 'best estimate' is not perfect, but by considering the complementary advantages of the two instruments, it gives an idea what to expect at the best from passive remote sensing.

In order to evaluate the stability of the cloud column structure from complementary passive remote sensing, we have also combined AIRS-CIRS with AVHRR-Cloud_cci. CAH of AVHRR-Cloud_cci is underestimated (Fig. 1), but the combination of both, built in the same manner as above, leads to results very similar to the ones from the CIRS–MODIS-CE combination, with global averages $\mathrm{CAM_{all}}=0.25$ and $\mathrm{CAL_{all}}=0.47$.

Strictly speaking, this method should be applied to the instantaneous gridded data, but again the results show that this approximation on monthly data compares well to the column cloud amounts from the active sensors (Fig. 11).

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10712-024-09824-0.

**Acknowledgements** This article was proposed to be part of the theme "Challenges in Understanding the Global Water Energy Cycle and its Changes in Response to Greenhouse Gas Emissions" within the framework of the workshop held at the International Space Science Institute (ISSI) in Bern, Switzerland (26–30 September 2022). The authors want to thank Jay Mace for the help with the specific CloudSat-CALIPSO GEOPROF dataset used for this study, as well as Genevieve Sèze for discussions on the article. Furthermore, we want to thank three anonymous reviewers for their very valuable comments which makes this manuscript easier to follow.

**Author's Contribution** All authors read and approved the final manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This study was partly supported by the Centre National de la Recherche Scientifique (CNRS), the Centre National d'Etudes Spatiales (CNES).

**Data availability** This updated GEWEX cloud assessment database is distributed by the French data centre AERIS at https://gewexca.aeris-data.fr/.

## Declarations

**Conflict of interest** The authors declare that they have no competing interests.

## References

Astin I, Di Girolamo L, Van de Poll HM (2001) Baysian confidence intervals for true fractional coverage from finite transect measurements: implications for cloud studies from space. J Geophys Res 106:17303–17310. https://doi.org/10.1029/2001JD900168

Avery MA, Ryan RA, Getzewich BJ, Vaughan MA, Winker DM, Hu Y, Garnier A, Pelon J, Verhappen CA (2020) CALIOP V4 cloud thermodynamic phase assignment and the impact of near-nadir viewing angles. Atmos Meas Tech 13:4539–4563. https://doi.org/10.5194/amt-13-4539-2020

Balmes KA, Fu Q, Thorsen TJ (2019) Differences in ice cloud opticaldepth from CALIPSO andground-based Raman lidar at the ARMSGP and TWP sites. J Geophys Res Atmos 124:1755–1778. https://doi.org/10.1029/2018JD028321

Bender FA-M, Ramanathan V, Tselioudis G (2012) Changes in extratropical storm track cloudiness 1983–2008: observational support for a poleward shift. Clim Dyn 38:2037–2053. https://doi.org/10.1007/s00382-011-1065-6

Chang F-L, Li Z (2005) A new method for detection of cirrus overlapping water clouds and determination of their optical properties. J Atmos Sci 62:3993–4009. https://doi.org/10.1175/JAS3578.1

Chen T, Zhang Y, Rossow WB (2000) Sensitivity of atmospheric radiative hetaing rate profiles to variations of cloud layer overlap. J Clim 13:2941–2959. https://doi.org/10.1175/1520-0442(2000)013%3c2941:SOARHR%3e2.0.CO;2

Chepfer H, Bony S, Winker D, Cesana G, Dufresne JL, Minnis P, Stubenrauch CJ, Zeng S (2010) The GCM oriented CALIPSO cloud product (CALIPSO-GOCCP). J Geophys Res 115:D00H16. https://doi.org/10.1029/2009JD012251

Dee DP, Uppala SM, Simmons AJ, Berrisford P, Poli P, Kobayashi S, Andrae U, Balmaseda MA, Balsamo G, Bauer P, Bechtold P, Beljaars AC, van de Bergn M, Bidlot L, Bormann J, Delsol NC, Dragani R, Fuentes M, Geer AJ, Haimberger L, Healy SB, Hersbach H, Holm EV, Isaksen L, Kallberg P, Kohler M, Matricardi M, McNally AP, Monge-Sanz BM, Morcrette J-J, Park B-K, Peubey C, de Rosnay P, Tavolato C, Thepaut J-N, Vitart F (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system. Q J R Meteorol Soc 137:553–597. https://doi.org/10.1002/qj.828

Dutta S, Di Girolamo L, Dey S, Zhan Y, Moroney CM, Zhao G (2020) The reduction in near-global cloud cover after correcting for biases caused by finite resolution measurement. Geophys Res Lett 47:e2020GL090313. https://doi.org/10.1029/2020GL090313

Eastman R, Warren SG (2014) Diurnal cycles of cumulus, cumulonimbus, stratus, stratocumulus, and fog from surface observations over land and ocean. J Clim 27:2386–2404. https://doi.org/10.1175/JCLI-D-13-00352.1

Eastman R, Warren SG, Hahn CJ (2011) Variations in cloud cover and cloud types over the ocean from surface observations, 1954–2008. J Clim 24:5914–5934. https://doi.org/10.1175/2011JCLI3972.1

Feofilov AG, Stubenrauch CJ (2019) Diurnal variation of high-level clouds from the synergy of AIRS and IASI space-borne infrared sounders. Atmos Chem Phys 19:13957–13972. https://doi.org/10.5194/acp-19-13957-2019

Foster MJ, Phillips C, Heidinger AK, Borbas EE, Li Y, Menzel WP, Walther A, Weisz E (2023) PATMOS-x version 6.0: 40 years of merged AVHRR and HIRS global cloud data. J Clim 36:1143–1160. https://doi.org/10.1175/JCLI-D-22-0147.1

Frey RA, Ackerman SA, Holz RE, Dutcher S, Griffith Z (2020) The continuity MODIS-VIIRS cloud mask. Remote Sens 12:3334

Garnier A, Pelon J, Vaughan MA, Winker DM, Trepte CR, Dubuisson P (2015) Lidar multiple scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud optical depths over oceans. Atmos Meas Tech 8:2759–2774. https://doi.org/10.5194/amt-8-2759-2015

Garnier A, Pelon J, Pascal N, Vaughan MA, Dubuisson P, Yang P, Mitchell DL (2021a) Version 4 CALIPSO imaging infrared radiometer ice and liquid water cloud microphysical properties: part II: results over oceans. Atmos Meas Tech 14:3277–3299. https://doi.org/10.5194/amt-14-3277-2021

Garnier A, Pelon J, Pascal N, Vaughan MA, Dubuisson P, Yang P, Mitchell DL (2021b) Version 4 CALIPSO imaging infrared radiometer ice and liquid water cloud microphysical properties: part I: the retrieval algorithms. Atmos Meas Tech 14:3253–3276. https://doi.org/10.5194/amt-14-3253-2021

Di Girolamo L, Menzies A, Zhao G, Mueller K, Moroney C, Diner DJ (2010) MISR Level 3 cloud fraction by altitude theoretical basis, JPL D-62358, Jet Propulsion Laboratory, Pasadena, CA, p 24. Available at https://eospso.gsfc.nasa.gov/sites/default/files/atbd/MISR_CFBA_ATBD.pdf

GISTEMP Team (2023) GISS surface temperature analysis (GISTEMP), version 4. NASA Goddard Institute for Space Studies. Dataset accessed 2023-12-21 at data.giss.nasa.gov/gistemp/

Gryspeerdt E, Stier P (2012) Regime-based analysis of aerosol-cloud interactions. Geophys Res Lett 39:21802. https://doi.org/10.1029/2012GL053221

Guignard A, Stubenrauch CJ, Baran AJ, Armante R (2012) Bulk microphysical properties of semi-transparent cirrus from AIRS: a six year global climatology and statistical analysis in synergy with geometrical profiling data from CloudSat-CALIPSO. Atmos Chem Phys 12:503–525. https://doi.org/10.5194/acp-12-503-2012,2012

Hahn CJ, Rossow WB, Warren SG (2001) ISCCP cloud properties associated with standard cloud types identified in individual surface observations. J Clim 14:11–28. https://doi.org/10.1175/1520-0442(2001)014%3c0011:ICPAWS%3e2.0.CO;2

Ham S-H, Kato S, Rose FG, Loeb NG, Xu K-M, Thorsen T, Basilovich MG, Sun-Mack S, Chen Y, Miller WF (2021) Examining cloud macrophysical changes over the Pacific for 2007–17 using CALIPSO, CloudSat, and MODIS observations. J Appl Meteor Climatol 60:1105–1126. https://doi.org/10.1175/JAMC-D-20-0226.1

Hamann U, Walther A, Baum B, Bennartz R, Bugliaro L, Derrien M, Francis PN, Heidinger A, Joro S, Kniffka A, Le Gléau H, Lockhoff M, Lutz H-J, Meirink JF, Minnis P, Palikonda R, Roebeling R, Thoss A, Platnick S, Watts P, Wind G (2014) Remote sensing of cloud top pressure/height from SEVIRI: analysis of ten current retrieval algorithms. Atmos Meas Tech 7:2839–2867. https://doi.org/10.5194/amt-7-2839-2014

Heidinger AK, Pavolonis MJ (2009) Gazing at cirrus clouds for 25 years through a split window. Part I: methodology. J Appl Meteor Climatol 48:1100–1116. https://doi.org/10.1175/2008JAMC1882.1

Heidinger AK, Evan AT, Foster MJ, Walther A (2012) A naive Bayesian cloud detection scheme derived from CALIPSO and applied within PATMOS-x. J Appl Meteorol Climatol 51:1129–1144. https://doi.org/10.1175/JAMC-D-11-02.1

Heidinger AK, Bearson N, Foster MJ, Li Y, Wanzong S, Ackerman S, Holz RE, Platnick S, Meyer K (2019) Using sounder data to improve cirrus cloud height estimation from satellite imagers. J Atmos Ocean Technol 36:1331–1342. https://doi.org/10.1175/JTECH-D-18-0079.1

Hersbach H, Bell B, Berrisford P et al (2020) The ERA5 global reanalysis. Q J R Meteorol Soc 146:1999–2049. https://doi.org/10.1002/qj.3803

Holz RE, Ackerman SA, Nagle FW, Frey R, Dutcher S, Kuehn RE, Vaughan MA, Baum B (2008) Global MODIS cloud detection and height evaluation using CALIOP. J Geophys Res 113:D00A19. https://doi.org/10.1029/2008JD009837

Hu X, Ge J, Li Y, Marchand R, Huang J, Fu Q (2020) Improved hydrometeor detection method: an application to CloudSat. Earth Space Sci 7:e2019EA000900. https://doi.org/10.1029/2019EA000900

Jensen EJ, van den Heever SC, Grant LD (2018) The life cycles of ice crystals detrained from the tops of deep convection. J Geophys Res Atmos 123:9624–9634. https://doi.org/10.1029/2018JD028832

Jin Y, Rossow WB (1997) Detection of cirrus overlapping low-level clouds. J Geophys Res 102:1727–1737. https://doi.org/10.1029/96jd02996

Jing X, Zhang H, Peng J, Li J, Barker H (2016) Cloud overlapping parameter obtained from CloudSat/CALIPSO dataset and its application in AGCM with McICA scheme. Atmos Res 170:52–65. https://doi.org/10.1016/j.atmosres.2015.11.007

Johnson RH, Rickenbach TM, Rutledge SA, Ciesielski PE, Schubert WH (1999) Trimodal characteristics of tropical convection. J Clim 12:2397–2418. https://doi.org/10.1175/1520-0442(1999)012%3c2397:TCOTC%3e2.0.CO;2

Jones AL, Di Girolamo L, Zhao G (2012) Reducing the resolution bias in cloud fraction from satellite derived clear-conservative cloud masks. J Geophys Res 117:D12201. https://doi.org/10.1029/2011JD017195

Karlsson K-G, Devasthale A (2018) Inter-comparison and evaluation of the four longest satellite-derived cloud climate data records: CLARA-A2, ESA Cloud CCI V3, ISCCP-HGM, and PATMOS-x. Remote Sens 10:1567. https://doi.org/10.3390/rs10101567

Karlsson K-G, Anttila K, Trentmann J, Stengel M, Meirink J-F, Devasthale A, Hanschmann T, Kothe S, Jääskeläinen E, Sedlar J, Benas N, van Zadelhoff G-J, Schlundt C, Stein D, Finkensieper S, Håkansson N, Hollmann R, Fuchs P, Werscheck M (2017) CLARA-A2: CM SAF cLoud, Albedo and surface RAdiation dataset from AVHRR data-edition 2, satellite application facility on climate monitoring. https://doi.org/10.5676/EUM_SAF_CM/CLARA_ AVHRR/V002

Karlsson K-G, Stengel M, Meirink JF, Riihelä A, Trentmann J, Akkermans T, Stein D, Devasthale A, Eliasson S, Johansson E, Håkansson N, Solodovnik I, Benas, N, Clerbaux N, Selbach N, Schröder M, Hollmann R (2023) CLARA-A3: the third edition of the AVHRR-based CM SAF climate data record on clouds, radiation and surface albedo covering the period 1979 to 2023. Earth Syst Sci Data Discuss [preprint]. https://doi.org/10.5194/essd-2023-133 (in review)

Koren I, Oreopoulos L, Feingold G, Remer LA, Altaratz O (2008) How small is a small cloud? Atmos Chem Phys 8:3855–3864. https://doi.org/10.5194/acp-8-3855-2008

Kotarba AZ (2022) Errors in global cloud climatology due to transect sampling with the CALIPSO satellite lidar mission. Atmos Res 279:106379. https://doi.org/10.1016/j.atmosres.2022.106379

Lau N-C, Crane MW (1995) A satellite view of the synoptic-scale organization of cloud properties in mid-latitude and tropical circulation systems. Mon Weather Rev 123:1984–2006. https://doi.org/10.1175/1520-0493(1995)123%3c1984:asvots%3e2.0.co;2

Lau N-C, Crane MW (1997) Comparing satellite and surface observations of cloud patterns in synoptic-scale circulations. Mon Weather Rev 125:3172–3189. https://doi.org/10.1175/1520-0493(1997)125%3c3172:csasoo%3e2.0.c0;2

Lenssen N, Schmidt G, Hansen J, Menne M, Persin A, Ruedy R, Zyss D (2019) Improvements in the GISTEMP uncertainty model. J Geophys Res Atmos 124:6307–6326. https://doi.org/10.1029/2018JD029522

Li J, Huang J, Stamnes K, Wang T, Lv Q, Jin H (2015) A global survey of cloud overlap based on CALIPSO and CloudSat measurements. Atmos Chem Phys 15:519–536. https://doi.org/10.5194/acp-15-519-2015

Liao X, Rossow WB, Rind D (1995a) Comparison between SAGE II and ISCCP high-level clouds, part I: global and zonal mean cloud amounts. J Geophys Res 100:1121–1135. https://doi.org/10.1029/94JD02429

Liao X, Rossow WB, Rind D (1995b) Comparison between SAGE II and ISCCP high-level clouds, part II: locating cloud tops. J Geophys Res 100:1137–1147. https://doi.org/10.1029/94JD02430

Liu Z, Kar J, Zeng S, Tackett J, Vaughan M, Avery M, Pelon J, Getzewich B, Lee K-P, Magill B, Omar A, Lucker P, Trepte C, Winker D (2019) Discriminating between clouds and aerosols in the CALIOP version 4.1 data products. Atmos Meas Tech 12:703–734. https://doi.org/10.5194/amt-12-703-2019

Liu H, Koren I, Altaratz O, Chekroun MD (2023) Opposing trends of cloud coverage over land and ocean under global warming. Atmos Chem Phys 23:6559–6569. https://doi.org/10.5194/acp-23-6559-2023

Luo Z, Anderson RC, Rossow WB, Takahashi H (2017) Tropical cloud and precipitation regimes as seen from near-simultaneous TRMM, CloudSat and CALIPSO observations and comparison with ISCCP. J Geophys Res Atmos 122:5988–6003. https://doi.org/10.1002/2017JD026569

Mace GG, Zhang Q (2014) The CloudSat radar–lidar geometrical profile product (RL-GeoProf): updates, improvements, and selected results. J Geophys Res Atmos 119:9441–9462. https://doi.org/10.1002/2013JD021374

Mace GG, Zhang Q, Vaughan M, Marchand R, Stephens G, Trepte C, Winker D (2009) A description of hydrometeor layer occurrence statistics derived from the first year of merged Cloudsat and CALIPSO data. J Geophys Res 114:D00A26. https://doi.org/10.1029/2007JD009755

Marvel K, Zelinka M, Klein SA, Bonfils C, Caldwell P, Doutriaux C, Santer BD, Taylor KE (2015) External influences on modeled and observed cloud trends. J Clim 28:4820–4840. https://doi.org/10.1175/JCLI-D-14-00734.1

McGarragh GR, Poulsen CA, Thomas GE, Povey AC, Sus O, Stapelberg S, Schlundt C, Proud S, Christensen MW, Stengel M, Hollmann R, Grainger RG (2018) The Community Cloud retrieval for CLimate (CC4CL): part 2: the optimal estimation approach. Atmos Meas Tech 11:3397–3431. https://doi.org/10.5194/amt-11-3397-2018

Menzel WP, Frey RA, Borbas EE, Baum BA, Cureton G, Bearson N (2016) Reprocessing of HIRS satellite measurements from 1980 to 2015: development toward a consistent decadal cloud record. J Appl Meteor Climatol 55:2397–2410. https://doi.org/10.1175/JAMC-D-16-0129.1

Meyer K, Platnick S, Holz R, Dutcher S, Quinn G, Nagle F (2020) Derivation of shortwave radiometric adjustments for SNPP and NOAA-20 VIIRS for the NASA MODIS-VIIRS continuity cloud products. Remote Sens 12:4096. https://doi.org/10.3390/rs12244096

Minnis P, Yost CR, Sun-Mack S, Chen Y (2008) Estimating the physical top altitude of optically thick ice clouds from thermal infrared satellite observations using CALIPSO data. Geophys Res Lett 35:L12801. https://doi.org/10.1029/2008GL033947

Minnis P, Sun-Mack S, Yost CR, Chen Y, Smith WL Jr, Chang F-L, Heck PW, Arduini RF, Trepte QZ, Ayers K, Bedka K, Bedka S, Brown RR, Heckert E, Hong G, Jin Z, Palikonda R, Smith R, Scarino B, Spangenberg DA, Yang P, Xie Y, Yi Y (2021) CERES MODIS cloud product retrievals for edition 4, part I: algorithm changes to CERES MODIS. IEEE Trans Geosci Remote Sens 59:2744–2780

Minnis P, Sun-Mack S, Smith WL Jr, Hong G, Chen Y (2019) Advances in neural network detection and retrieval of multilayer clouds for CERES using multispectral satellite data. In: Proceedings of the SPIE conference remote sensing clouds and the atmosphere. XXIV, Strasbourg, France, Sept 9–12, 11152, p 12. https://doi.org/10.1117/12.2532931

Mitra A, Di Girolamo L, Hong Y, Zhan Y, Mueller KJ (2021) Assessment and error analysis of Terra-MODIS and MISR cloud-top heights through comparison with ISS-CATS lidar. J Geophys Res Atmos 126:e2020JD034281. https://doi.org/10.1029/2020JD034281

Mitra A, Loveridge JR, Di Girolamo L (2023) Fusion of MISR stereo cloud heights and terra-MODIS thermal infrared radiances to estimate two-layered cloud properties. J Geophys Res Atmos 128:e2022JD038135. https://doi.org/10.1029/2022JD038135

Mueller K, Maroney C, Jovanovic V, Garay MJ, Muller J-P, Di Girolamo L, Davies R (2013) MISR level 2 cloud product algorithm theoretical basis, JPL D-73327, Pasadena, CA, p 61. Available at https://eospso.gsfc.nasa.gov/sites/default/files/atbd/MISR_L2_CLOUD_ATBD-1.pdf

Naud CM, Del Genio A, Mace GG, Benson S, Clothiaux EE, Kollias P (2008) Impact of dynamics and atmospheric state on cloud vertical overlap. J Clim 21:1758–1770. https://doi.org/10.1175/2007JCLI1828.1

Noël V, Chepfer H, Chiriaco M, Yorks J (2018) The diurnal cycle of cloud profiles over land and ocean between 51° S and 51° N, seen by the CATS spaceborne lidar from the International Space Station. Atmos Chem Phys 18:9457–9473. https://doi.org/10.5194/acp-18-9457-2018

Norris JR (2005) Multidecadal changes in near-global cloud cover and estimated cloud cover radiative forcing. J Geophys Res 110:D08206. https://doi.org/10.1029/2004JD005600

Norris JR, Allen RJ, Evan AT, Zelinka MD, O'Dell CW, Klein SA (2016) Evidence for climate change in the satellite cloud record. Nature 536:72–75. https://doi.org/10.1038/nature18273

Oreopoulos L, Cho N, Lee D (2017) New insights about cloud vertical structure from CloudSat and CALIPSO observations. J Geophys Res Atmos 12:9280–9300. https://doi.org/10.1002/2017JD026629

Philipp D, Stengel M, Ahrens B (2020) Analyzing the arctic feedback mechanism between sea ice and low-level clouds using 34 years of satellite observations. J Clim 33:7479–7501. https://doi.org/10.1175/JCLI-D-19-0895.1

Pincus R, Platnick S, Ackerman SA, Hemler RS, Hofmann RJP (2012) Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators. J Clim 25:4699–4720. https://doi.org/10.1175/JCLI-D-11-00267.1

Platnick S, Meyer KG, Yang P, Ridgway WL, Riedi JC, King MD, Wind G, Amarasinghe N, Marchant B, Arnold GT et al (2017) The MODIS cloud optical and microphysical products: collection 6 updates and examples from Terra and Aqua. IEEE Trans Geosci Remote Sens 55:502–525. https://doi.org/10.1109/TGRS.2016.2610522

Platnick S, Meyer K, Wind G, Holz RE, Amarasinghe N, Hubanks PA, Marchant B, Dutcher S, Veglio P (2021) The NASA MODIS-VIIRS continuity cloud optical properties products. Remote Sens 13:2. https://doi.org/10.3390/rs13010002

Protopapadaki ES, Stubenrauch CJ, Feofilov AG (2017) Upper tropospheric cloud systems derived from IR sounders: properties of Cirrus Anvils in the tropics. Atmosph Chem Phys 17:3845–3859. https://doi.org/10.5194/acp-17-3845-2017

Rossow WB (1989) Measuring cloud properties from space: a review. J Clim 2:201–213. https://doi.org/10.1175/1520-0442(1989)002%3c0201:mcpfsa%3e2.0.co;2

Rossow WB, Schiffer RA (1991) ISCCP cloud data products. Bull Am Meteorl Soc 72:2–20. https://doi.org/10.1175/1520-0477(1991)072%3c0002:icdp%3e2.0.co;2

Rossow WB, Schiffer RA (1999) Advances in understanding clouds from ISCCP. Bull Am Meteorl Soc 80:2261–2287. https://doi.org/10.1175/1520-0477(1999)080%3c2261:aiucfi%3e2.0.co;2

Rossow WB, Zhang Y-C (1995) Calculation of surface and top-of-atmosphere radiative fluxes from physical quantities based on ISCCP datasets: 2. Validation and first results. J Geophys Res 100:1167–1197. https://doi.org/10.1029/94JD02746

Rossow WB, Zhang Y-C (2010) Evaluation of a statistical model of cloud vertical structure using combined CloudSat and CALIPSO cloud layer profiles. J Clim 23:6641–6653. https://doi.org/10.1175/2010JCLI3734.1

Rossow WB, Tselioudis G, Polak A, Jakob C (2005a) Tropical climate described as a distribution of weather states indicated by distinct mesoscale cloud property mixtures. Geophys Res Lett 32:L21812. https://doi.org/10.1029/2005GL024584

Rossow WB, Zhang Y-C, Wang J (2005b) A statistical model of cloud vertical structure based on reconciling cloud layer amounts inferred from satellites and radiosonde humidity profiles. J Clim 18:3587–3605. https://doi.org/10.1175/JCLI3479.1

Rossow WB, Knapp KR, Young AH (2022) International satellite cloud climatology project: extending the record. J Climate 35:141–158. https://doi.org/10.1175/jcli-d-21-0157.1

Rossow WB (2017) Climate data record (CDR) Program climate algorithm theoretical basis document of international satellite cloud climatology project (ISCCP) cloud properties. CDRP-ATBD-0872.

Available at https://www.ncei.noaa.gov/pub/data/sds/cdr/CDRs/Cloud_Properties-ISCCP/AlgorithmDescription_01B-29.pdf

Sassen K, Wang Z (2012) The clouds of the middle troposphere: composition, radiative impact, and global distribution. Surv Geophys 33:677–691. https://doi.org/10.1007/s10712-011-9163-x

Slingo JM, Slingo A (1991) The response of a general circulation model to cloud longwave radiative forcing. II: further studies. Q J R Meteorol Soc 117:333–364

Stachnik JP, Schumacher C, Ciesielski PE (2013) Total heating characteristics of the ISCCP tropical and subtropical cloud regimes. J Clim 26:7097–7116. https://doi.org/10.1175/JCLI-D-12-00673.1

Staten PW, Lu J, Grise KM, Davis SM, Birner T (2018) Re-examining tropical expansion. Nat Clim Change 8:768–775. https://doi.org/10.1038/s41558-018-0246-2

Stengel M, Stapelberg S, Sus O, Finkensieper S, Würzler B, Philipp D, Hollmann R, Poulsen C, Christensen M, McGarragh G (2020) Cloud_cci advanced very high resolution radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties. Earth Syst Sci Data 12:41–60. https://doi.org/10.5194/essd-12-41-2020

Stephens GL, Winker D, Pelon J, Trepte C, Vane D, Yuhas C, L'Ecuyer T, Lebsock M (2018a) CloudSat and CALIPSO within the a-train: ten years of actively observing the earth system. Bull Am Meteorol Soc 99:569–581. https://doi.org/10.1175/BAMS-D-16-0324.1

Stephens GL, Hakuba MZ, Webb M, Lebsock M, Yue Q, Kahn BH, Hristova-Veleva S, Rapp A, Stubenrauch C, Elsaesser GS, Slingo J (2018b) Regional intensification of the tropical hydrological cycle during ENSO. Geophys Res Lett 45:4361–4370. https://doi.org/10.1029/2018GL077598

Stubenrauch CJ, Chédin A, Armante R, Scott NA (1999) Clouds as seen by infrared sounders (3I) and imagers (ISCCP): part II) a new approach for cloud parameter determination in the 3I algorithms. J Clim 12:2214–2223. https://doi.org/10.1175/1520-0442(1999)012%3c2214:CASBSS%3e2.0.CO;2

Stubenrauch CJ, Chédin A, Rädel G, Scott NA, Serrar S (2006) Cloud properties and their seasonal and diurnal variability from TOVS Path-B. J Clim 19:5531–5553. https://doi.org/10.1175/JCLI3929.1

Stubenrauch CJ, Cros S, Guignard A, Lamquin N (2010) A six-year global cloud climatology from the atmospheric InfraRed sounder aboard the AQUA satellite: statistical analysis in synergy with CALIPSO and CloudSat. Atmos Chem Phys 10:7197–7214. https://doi.org/10.5194/acp-10-7197-2010

Stubenrauch CJ, Rossow WB, Kinne S, Ackerman S, Cesana G, Chepfer H, Di Girolamo L, Getzewich B, Guignard A, Heidinger A, Maddux BC (2013) Assessment of global cloud datasets from satellites: project and database initiated by the GEWEX radiation panel. B Am Meteorol Soc 94:1031–1049. https://doi.org/10.1175/BAMS-D-12-00117.1

Stubenrauch CJ, Feofilov AG, Protopapadaki SE, Armante R (2017) Cloud climatologies from the infrared sounders AIRS and IASI: strengths and applications. Atmos Chem Phys 17:13625–13644. https://doi.org/10.5194/acp-17-13625-2017

Stubenrauch CJ, Caria G, Protopapadaki SE, Hemmer F (2021) The effect of tropical upper tropospheric cloud systems on radiative heating rate fields derived from synergistic A-train satellite observations. Atmos Chem Phys 21:1015–1034. https://doi.org/10.5194/acp-21-1015-2021

Stubenrauch CJ, Mandorli G, Lemaitre E (2023) Convective organization and 3D structure of tropical cloud systems deduced from synergistic A-Train observations and machine learning. Atmos Chem Phys 23:5867–5884. https://doi.org/10.5194/acp-23-5867-2023

Stubenrauch CJ, Rossow WB, Kinne S, GEWEX Cloud Assessment Team (2012) Assessment of Global Cloud Datasets from Satellites, A Project of the World Climate Research Programme Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel, WCRP report, p 180. Available at chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.wcrp-climate.org/documents/GEWEX_Cloud_Assessment_2012.pdf and at http://climserv.ipsl.polytechnique.fr/gewexca/

Sun-Mack S, Minnis P, Chen Y, Hong G, Smith WL Jr (2023) Identification of ice-over-water multilayer clouds using an artificial neural network with multispectral satellite data. Atmos Meas Tech Disc. https://doi.org/10.5194/egusphere-2023-2804

Sus O, Stengel M, Stapelberg S, McGarragh G, Poulsen C, Povey AC, Schlundt C, Thomas G, Christensen M, Proud S, Jerg M, Grainger R, Hollmann R (2018) The Community Cloud retrieval for CLimate (CC4CL): part 1: a framework applied to multiple satellite imaging sensors. Atmos Meas Tech 11:3373–3396. https://doi.org/10.5194/amt-11-3373-2018

Tan J, Oreopoulos L (2019) Subgrid precipitation properties of mesoscale atmospheric systems represented by MODIS cloud regimes. J Clim 32:1797–1812. https://doi.org/10.1175/JCLI-D-18-0570.1

Trepte QZ, Minnis P, Sun-Mack S, Yost CR, Chen Y, Jin Z, Chang F-L, Smith WL Jr, Bedka KM, Chee TL (2019) Global cloud detection for CERES 4 using Terra and Aqua MODIS data. IEEE Trans Geosci Remote Sens 57:9410–9449. https://doi.org/10.1109/TGRS.2019.2926620

Tselioudis G, Rossow WB, Zhang Y, Konsta D (2013) Global weather states and their properties from passive and active satellite cloud retrievals. J Clim 26:7734–7746. https://doi.org/10.1175/JCLI-D-13-00024.1

Tselioudis G, Rossow WB, Jakob C, Remillard J, Tropf D, Zhang Y-C (2021) Evaluation of clouds, radiation, and precipitation in CMIP6 models using global weather states derived from ISCCP-H cloud property data. J Clim 34:7311–7324. https://doi.org/10.1175/jcli-d-21-0076.1

Walther A, Heidinger AK (2012) Implementation of the daytime cloud optical and microphysical properties algorithm (DCOMP) in PATMOS-x. J Appl Meteorol Climatol 51:1371–1390. https://doi.org/10.1175/JAMC-D-11-0108.1

Wang J, Rossow WB (1998) Effects of cloud vertical structure on atmospheric circulation in the GISS GCM. J Clim 11:3010–3029. https://doi.org/10.1175/1520-0442(1998)011%3c3010:eocvso%3e2.0.co;2

Wang C, Luo ZJ, Chen X, Zeng X, Tao W-K, Huang X (2014) A physically based algorithm for non-blackbody correction of cloud top temperature and application to convection study. J Appl Meteorol Climatol 53:1844–1857. https://doi.org/10.1175/JAMC-D-13-0331.1

Warren SG, Hahn CJ, London J (1985) Simultaneous occurrence of different cloud types. J Clim Appl Meteorol 24:658–667. https://doi.org/10.1175/1520-0450(1985)024%3c0658:SOODCT%3e2.0.CO;2

Wielicki BA, Parker L (1992) On the determination of cloud cover from satellite sensors: the effect of sensor spatial resolution. J Geophys Res 97:12799–12823. https://doi.org/10.1029/92jd01061

Winker DM, Pelon J, Coakley JA Jr et al (2010) The CALIPSO mission: a global 3D view of aerosols and clouds. Bull Am Meteorol Soc 91:1211–1229. https://doi.org/10.1175/2010BAMS3009.1

Wylie DP, Wang P-H (1997) Comparison of cloud frequency data from the high-resolution infrared radiometer sounder and the stratospheric aerosol and gas experiment II. J Geophys Res 102:29893–29900. https://doi.org/10.1029/97JD02360

Wylie DP, Jackson DL, Menzel WP, Bates JJ (2005) Trends in global cloud cover in two decades of HIRS observations. J Clim 18:3021–3031. https://doi.org/10.1175/JCLI3461.1

Young AH, Knapp RK, Inamdar A, Hankins W, Rossow WB (2018) The international cloud climatology project H-series climate data record product. Earth Syst Sci Data 10:583–593. https://doi.org/10.5194/essd-10-583-2018

Zhang Y, Rossow WB (2023) Global radiative flux profile data set: revised and extended. J Geophys Res Atmos 128:e2022JD037340. https://doi.org/10.1029/2022JD037340

Zhang Y, Rossow WB, Lacis AA, Oinas V, Mishchenko MI (2004) Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: refinements of the radiative transfer model and the input data. J Geophys Res 109:D19105. https://doi.org/10.1029/2003JD004457

Zhao G, Di Girolamo L (2006) Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. Geophys Res Lett 33:L20802. https://doi.org/10.1029/2006GL027088

## Authors and Affiliations

**Claudia J. Stubenrauch[1] · Stefan Kinne[2] · Giulio Mandorli[1] · William B. Rossow[3] · David M. Winker[4] · Steven A. Ackerman[5] · Helene Chepfer[1] · Larry Di Girolamo[6] · Anne Garnier[4,7] · Andrew Heidinger[8] · Karl-Göran Karlsson[9] · Kerry Meyer[10] · Patrick Minnis[4,7] · Steven Platnick[10] · Martin Stengel[11] · Szedung Sun-Mack[4,7] · Paolo Veglio[5] · Andi Walther[5] · Xia Cai[7] · Alisa H. Young[12] · Guangyu Zhao[6]**

✉ Stefan Kinne
Stefan.Kinne@mpimet.mpg.de

[1] Laboratoire de Météorologie Dynamique / Institut Pierre-Simon Laplace, (LMD/IPSL), Sorbonne Université, Ecole Polytechnique, CNRS, 4 Place Jussieu, 75252 Paris Cedex 05, France

[2] Max Planck Institute for Meteorology, Bundesstr 53, 20146 Hamburg, Germany

[3] Franklin, USA

[4] NASA Langley Research Center, Hampton, VA 23681, USA

[5] Space Science and Engineering Center, University of Wisconsin—Madison, 1225W Dayton, Madison, WI 53706, USA

[6] Department of Atmospheric Sciences, University of Illinois at Urbana–Champaign, 1301 W. Green Str., Urbana, IL 61820, USA

[7] Analytical Mechanics Associates, 21 Enterprise Parkway Suite 300, Hampton, VA 23666, USA

[8] NOAA/NESDIS/STAR, 1225W Dayton, Madison, WI 53706, USA

[9] Swedish Meteorological and Hydrological Institute, Folksborgsvaegen 17, 60176 Norrköping, Sweden

[10] Earth Sciences Division, NASA Goddard Space Flight Center, Bldg 33, Greenbelt, MD 20771, USA

[11] Deutscher Wetterdienst, Frankfurter Str 135, 63067 Offenbach am Main, Germany

[12] NOAA Great Lakes Environmental Research Lab, 4840S State Rd, Ann Arbor, MI 48108, USA