EUMETSAT Satellite Application Facility on Climate Monitoring

Product User Manual

CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3)

Cloud Products

DOI: 10.5676/EUM_SAF_CM/CLARA_AVHRR/V003

	CDR	ICDR
Fractional Cloud Cover	**CM-11012**	**CM-6011**
Joint Cloud property Histogram	**CM-11022**	**CM-6021**
Cloud Top Level	**CM-11032**	**CM-6031**
Cloud Phase	**CM-11042**	**CM-6041**
Liquid Water Path	**CM-11052**	**CM-6051**
Ice Water Path	**CM-11062**	**CM-6061**

Reference Number: SAF/CM/DWD/PUM/CLARA/CLD
Issue/Revision Index: 3.1
Date: 06.02.2023

Document Signature Table

	Name	Function	Signature	Date
Author	Karl-Göran Karlsson Abhay Devasthale Martin Stengel Irina Solodovnik Jan Fokke Meirink Nikos Benas	CM SAF scientist (SMHI) CM SAF scientist (SMHI) CM SAF scientist (DWD) CM SAF scientist (DWD) CM SAF scientist (KNMI) CM SAF scientist (KNMI)		06.02.2023
Editor	Marc Schröder	CM SAF Science Coordinator		06.02.2023
Approval	CM SAF Steering Group			
Release	Rainer Hollmann	CM SAF Project Manager		

Distribution List

Internal Distribution	
Name	**No. Copies**
DWD Archive	1
CM SAF Team	1

External Distribution		
Company	**Name**	**No. Copies**
PUBLIC		1

Document Change Record

Issue/ Revision	Date	DCN No.	Changed Pages/Paragraphs
3.0	02/10/2022	SAF/CM/DWD/PUM/CLARA/CLD	Version for CLARA-A3 DRR
3.1	25/11/2022	SAF/CM/DWD/PUM/CLARA/CLD	Updates following discussions at DRR3.2/ORR Adapted ICDR processing

Applicable documents

Reference	Title	Code
AD 1	EUMETSAT CM SAF CDOP 4 Product Requirements Document (PRD)	SAF/CM/DWD/PRD/ Issue 4.1
AD 2	SYSTEMATIC OBSERVATION REQUIREMENTS FOR SATELLITE-BASED DATA PRODUCTS FOR CLIMATE 2011 Update	GCOS-154

Reference Documents

Reference	Title	Code
RD 1	Validation Report CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3) Cloud Products	SAF/CM/SMHI/VAL/CLARA/ CLD, 3.1
RD 2	Algorithm Theoretical Basis Document CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3) Cloud Products (level-1 to level-3)	SAF/CM/DWD/ATBD/CLAR A/CLD, 3.3
RD 3	Algorithm Theoretical Basis Document for the Cloud Probability of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SC I/ATBD/CloudProbability Issue: 2.0, 26.04.2021
RD 4	Algorithm Theoretical Basis Document for Cloud Top Temperature, Pressure and Height of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SC I/ATBD/CTTH Issue 3.0, 26.04.2021
RD 5	Algorithm Theoretical Basis Document for Cloud Micro Physics of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SC I/ATBD/CMIC Issue 3, Rev. 0e, 26.04.2021
RD 6	Algorithm Theoretical Basis Document for the Cloud Mask of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SC I/ATBD/CloudMask, Issue 3, Rev. 0 26.04.2021

Table of Contents

List of Tables

List of Figures

1 Executive summary

This CM SAF Product User Manual provides information on the cloud products of the CLARA-A3 data record (CLARA-A3: CM SAF CLoud, Albedo and Radiation data record – AVHRR-based, Edition 3). The CLARA-A3 record is based on Advanced Very High Resolution Radiometer (AVHRR) observations onboard the NOAA and EUMETSAT Metop satellites. CLARA-A3 consists of a Climate Data Record (CDR), spanning the period from 1979 to 2020, and an Interim Climate Data Record (ICDR), starting in 2021 and extended operationally with low latency to the present (see Figure 1-1). The ICDR is produced with the same algorithms as the CDR but with a few differences in input data.

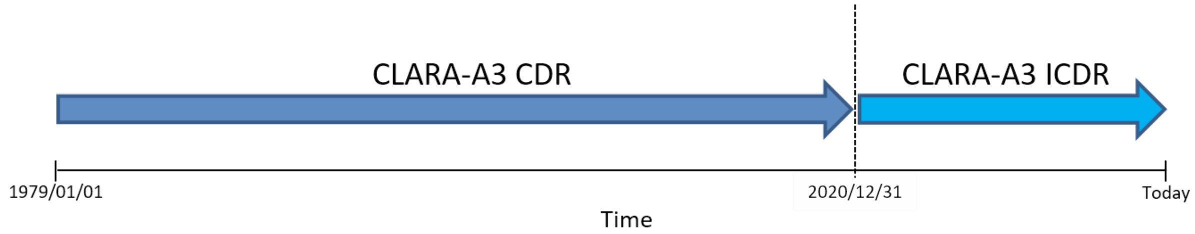

Figure 1-1: Illustration of the CLARA-3 CDR and ICDR temporal coverage.

This report presents user information on the following cloud products:

	CDR	ICDR
Fractional Cloud Cover (CFC)	**CM-11012**	**CM-6011**
Joint Cloud property Histogram (JCH)	**CM-11012**	**CM-6021**
Cloud Top Level (CTO)	**CM-11032**	**CM-6031**
Cloud Phase (CPH)	**CM-11042**	**CM-6041**
Liquid Water Path (LWP)	**CM-11052**	**CM-6051**
Ice Water Path (IWP)	**CM-11062**	**CM-6061**

Some attractive features of the CLARA-A3 data record are:

- More than 4 decades spanning global data record of geophysical variables based on one single satellite sensor family of AVHRR giving homogeneous coverage of conditions along specific latitude bands
- Multiple processing levels of the retrieved cloud properties are available ranging from pixel-based cloud retrievals, which are sampled onto a global 0.05° grid (close to the original sensor resolution), to monthly averages and histograms on a global 0.25° latitude-longitude grid.
- Comprehensive documentation available including a description of the algorithms used, the validation efforts performed, and important aspects for user guidance.

- Freely available data without restrictions. Easy access: (in netcdf format, CF standard) via an interactive download facility (http://wui.cmsaf.eu/).

This manual briefly describes the CLARA-A3 cloud products. A technical description of the data records including information on the file format as well as on the data access is provided. Furthermore, details on the implementation of the retrieval processing chain, and individual algorithm descriptions are available in the algorithm theoretical basis documents [RD 2] - [RD 7]. Basic accuracy requirements are defined in the product requirements document [AD 1]. A detailed validation of the CLARA-A3 cloud parameters is available in the validation report [RD 1].

2 The EUMETSAT SAF on Climate Monitoring (CM SAF)

The importance of climate monitoring with satellites was recognized in 2000 by EUMETSAT Member States when they amended the EUMETSAT Convention to affirm that the EUMETSAT mandate is also to "contribute to the operational monitoring of the climate and the detection of global climatic changes". Following this, EUMETSAT established within its Satellite Application Facility (SAF) network a dedicated centre, the SAF on Climate Monitoring (CM SAF, http://www.cmsaf.eu).

The consortium of CM SAF currently comprises the Deutscher Wetterdienst (DWD) as host institute, and the partners from the Royal Meteorological Institute of Belgium (RMIB), the Finnish Meteorological Institute (FMI), the Royal Meteorological Institute of the Netherlands (KNMI), the Swedish Meteorological and Hydrological Institute (SMHI), the Meteorological Service of Switzerland (MeteoSwiss), the Meteorological Service of the United Kingdom (UK MetOffice), and the Centre National de la Recherché Scientifique (CNRS). Since the beginning in 1999, the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) has developed and will continue to develop capabilities for a sustained generation and provision of Climate Data Records (CDR's) derived from operational meteorological satellites.

In particular, the generation of long-term data records is pursued. The ultimate aim is to make the resulting data records suitable for the analysis of climate variability and potentially the detection of climate trends. CM SAF works in close collaboration with the EUMETSAT Central Facility and liaises with other satellite operators to advance the availability, quality and usability of Fundamental Climate Data Records (FCDRs) as defined by the Global Climate Observing System (GCOS). As a major task the CM SAF utilizes FCDRs to produce records of Essential Climate Variables (ECVs) as defined by GCOS. Thematically, the focus of CM SAF is on ECVs associated with the global energy and water cycle.

Another essential task of CM SAF is to produce data records that can serve applications related to the Global Framework of Climate Services initiated by the WMO World Climate Conference-3 in 2009. CM SAF is supporting climate services at national meteorological and hydrological services (NMHSs) with long-term data records but also with data records produced close to real time that can be used to prepare monthly/annual updates of the state of the climate. Both types of products together allow for a consistent description of mean values, anomalies, variability and potential trends for the chosen ECVs. CM SAF ECV data records also serve the improvement of climate models both at global and regional scale.

As an essential partner in the related international frameworks, in particular WMO SCOPE-CM (Sustained COordinated Processing of Environmental satellite data for Climate Monitoring), the CM SAF - together with the EUMETSAT Central Facility, assumes the role as main implementer of EUMETSAT's commitments in support to global climate monitoring. This is achieved through:

- Application of highest standards and guidelines as lined out by GCOS for the satellite data processing,

- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,
- Intensive validation and improvement of the CM SAF climate data records,
- Taking a major role in data record assessments performed by research organisations such as WCRP (World Climate Research Program). This role provides the CM SAF with deep contacts to research organizations that form a substantial user group for the CM SAF CDRs,
- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

A catalogue of all available CM SAF products is accessible via the CM SAF webpage, http:/www.cmsaf.eu/. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

3 Compilation of the CLARA-A3 cloud data record

Measurements made by the Advanced Very High Resolution Radiometer (AVHRR) onboard the polar orbiting NOAA satellites and the EUMETSAT Metop satellites have been performed since 1978. Figure 3-1 gives an overview over all satellites carrying the AVHRR instrument that were used for CLARA-A3 CDR (covering the period 1979-2020). The AVHRR instrument only measured in four spectral bands in the beginning (AVHRR/1) but from 1982 a fifth channel was added (AVHRR/2) and in 1998 even a sixth channel was made available (AVHRR/3), although only accessible if switched with the previous third channel at 3.7 micron. Table 3-1 describes the AVHRR instrument, its various versions and the satellites carrying them. The AVHRR instrument measures at a horizontal resolution close to 1 km at nadir but only data at a reduced resolution of approximately 4 km are permanently archived and available with global coverage since the beginning of measurements. This data record is denoted Global Area Coverage (GAC) AVHRR data.

3.1 Basic characteristics of satellite observations

The retrieval of cloud physical properties (in particular particle effective radius and liquid/ice water path) is sensitive to the shortwave infrared channel being used. Table 3.2 summarizes when either of the channels 3a and 3b have been active on the AVHRR/3 instruments.

Table 3-1: Spectral channels of the Advanced Very High Resolution Radiometer (AVHRR). The three different versions of the instrument are described as well as the corresponding satellites. Notice that channel 3A was only used continuously on NOAA-17, Metop-A/B/C. For the other satellites with AVHRR/3 it was used only for shorter periods. The wavelength ranges given represent the full width of the spectral response function (SRF) at half maximum of the SRF (FWHM).

Channel Number	Wavelength (micrometers) AVHRR/1 TIROS-N, NOAA-6,8,10	Wavelength (micrometers) AVHRR/2 NOAA-7,9,11,12,14	Wavelength (micrometers) AVHRR/3 NOAA-15,16,17,18,19 Metop-A/B/C
1	0.58 - 0.68	0.58 - 0.68	0.58 - 0.68
2	0.725 - 1.10	0.725 - 1.10	0.725 - 1.10
3A	-	-	1.58 - 1.64
3B	3.55 - 3.93	3.55 - 3.93	3.55 - 3.93
4	10.50 - 11.50	10.50 - 11.50	10.50 - 11.50
5	Channel 4 repeated	11.5 – 12.5	11.5 – 12.5

Table 3-2: Channel 3A and 3B activity for the AVHRR/3 instruments during daytime. Notice that the given time periods show the availability in the CLARA-A3 data record and not the true lifetime of the individual sensor/satellite.

Satellite	Channel 3a active	Channel 3b active
NOAA-15	03/1999 – 04/1999 (3a/3b switching test period) Channel 3a has been active exclusively over the Alaska region from September to May for some years after 2016	10/1998 – 03/1999 04/1999 – present
NOAA-16	01/2001 – 04/2003	05/2003 – 12/2011
NOAA-17	07/2002 – 02/2010	-
NOAA-18	06/2005 – 08/2005	08/2005 present
NOAA-19	03/2009 – 05/2009 Channel 3a has been active exclusively over the Alaska region from September to May for some years after 2016	05/2009 present
Metop-A	06/2007 – 11/2021	-
Metop-B	01/2013 – present	-
Metop-C	07/2019 – present	-

Figure 3-1 describes the actual coverage of observations in CLARA-A3 from each individual satellite in the CDR period 1979 to 2020. Notice the different numbers of available AVHRRs in different sub-periods, starting with only one AVHRR in 1979, followed by two AVHRRs until 2002 and up to 5 AVHRRs in the remaining part of the period shown. At the end of the CDR period in 2020, NOAA-15/18/19 and Metop-A/B/C were used. This was also the baseline set of satellites at the beginning of the ICDR processing in 2021 with the exception of Metop-C where calibration issues did not allow an extension of the products for this particular satellite (see section 3.2).

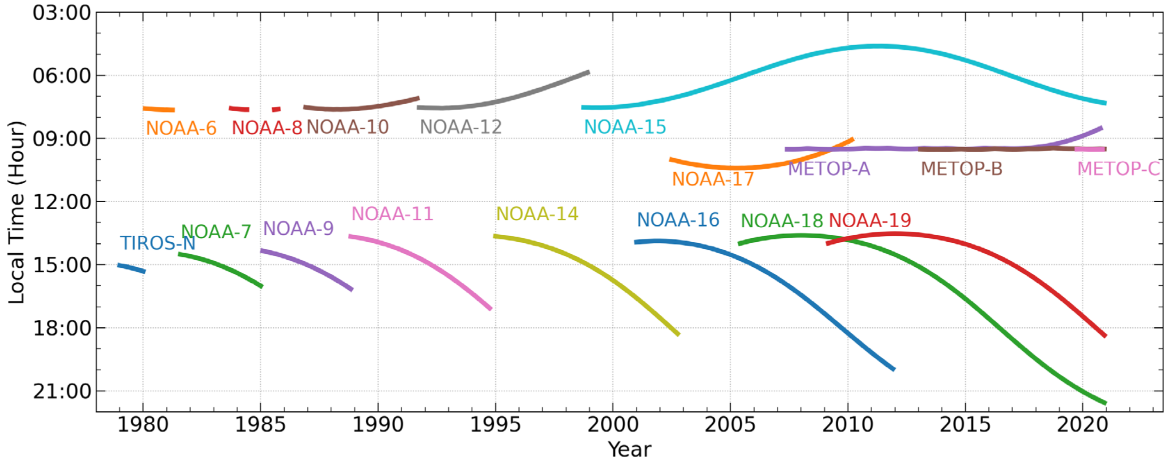

Figure 3-1: Local solar times at equator observations for all AVHRR-carrying NOAA satellites from TIROS-N to NOAA-19 and EUMETSAT's Metop A/B/C satellites. Shown are all data that are used for the CLARA-A3 CDR processing. Satellites shown for 2020 are also the baseline for the ICDR processing beyond 2020 (except Metop-C, explained further down in the text). The figure shows ascending (northbound) equator crossing times for afternoon satellites (NOAA-7 to NOAA-19) and descending (southbound) equator crossing times for morning satellites (NOAA-12 to NOAA-17 and Metop A/B/C). Corresponding night-time observations take place 12 hours earlier/later. Some data gaps are present but only for some isolated dates.

Observations from polar orbiting sun synchronous satellites are made at the same local solar time at each latitude band. Normally, satellites are classified into observation nodes according to the local solar time when crossing the equator during daytime (illuminated conditions). For the NOAA satellite observations, a system with one morning observation node and one afternoon observation node has been utilised as the fundamental polar orbiting observation system. This guarantees four equally distributed observations per day (if including the complementary observation times at night and in the evening when the satellite passes again 12 hours later). Equator crossing times have varied slightly between satellites. Morning satellites have generally been confined to the local solar time interval 07:00-08:00 and afternoon satellites to the interval 13:30-14:30. However, a change was introduced for the morning satellites NOAA-17, Metop-A, Metop-B and Metop-C, now being defined in a so-called mid-morning orbit with equator crossing times close to 10:00. A specific problem with the observation nodes for the NOAA satellites has been the difficulty to keep observation times stable for each individual satellite due to orbital drift (Figure 3-1; described in more detail by Ignatov et al., 2004). For regions exposed to strong diurnal cycles in cloudiness (or other cloud properties) this could cause artificial trends in time series analyses. This may also occur due to decreasing cloud detection efficiency at higher solar zenith angles. No compensation for this has been attempted in the CLARA-A3 data record but corrections are considered for future CLARA versions. A compact study on the effect of the (potentially) imperfect sampled diurnal cycle of clouds and on the related effect of orbital drift is given in Annex A.

3.2 Calibration aspects

An important aspect for any product-based climate data record (CDR) is that retrieved products have been derived from accurately calibrated and harmonized radiances (formally denoted

Fundamental Climate Data Records - FCDRs). So far, only the visible channels of the AVHRR have received FCDR status. Attempts were made in the FIDUCEO project to create an FCDR also for the infrared channels but this work could not be finished within the framework of that project. As a consequence, and awaiting an FCDR solution later on, the EUMETSAT Secretariat defined a project that in cooperation with CM SAF would compile an AVHRR FDR that could serve the needs for the CLARA-A3 CDR and which also could be released in parallel with CLARA-A3 for external use. The FDR is labelled the EUM AVHRR FDR. The FDR is basically compiled with the PyGAC tool, which was used also for CLARA-A2, but with some changes to output format and quality flag information. Thus, the AVHRR input to CLARA-A3 is the EUM AVHRR FDR with the following DOI number: 10.15770/EUM_SEC_CLM_0060.

The calibration of the AVHRR shortwave reflectances, which is applied in the generation of the FDR by EUMETSAT, is based on work initially performed by NOAA (updated version of Heidinger et al. (2010)). A similar FDR is prepared for the compilation of the "NOAA Pathfinder Atmospheres - Extended" (PATMOS-x) data record (for full description, see http://cimss.ssec.wisc.edu/patmosx/overview.html). The AVHRR reflectances are calibrated based on collocations with MODIS. This information is propagated back in time to all AVHRR-carrying satellites. A complete documentation of the calibration method is given by Heidinger (2018). It should be noted that uncertainties in the calibration of visible channels for Metop-C led to that Metop-C products are currently not included in the ICDR dataset. The reason is an anticipated decay with time of the calibration accuracy for the visible radiances for this satellite. The NOAA calibration method is based on a vicarious calibration technique for the visible channels (i.e., no on-board calibration is available) and this requires data from several years to estimate sensor degradation with time and potential biases compared to previous AVHRR instruments. With only 1 ½ years of data available from Metop-C in 2020, the calibration information (more clearly, the time-dependent corrections) was not considered reliable for this satellite for allowing further processing after 2020. Data from Metop-C will therefore be added in 2023 or later to the ICDR as soon as a new calibration update is received. Results will then also improve for the other active satellites due to an enhanced quality of the time-dependent calibration corrections.

The calibration of infrared AVHRR channels is based on the onboard blackbody calibration. This has been found to provide stable and reliable results already. However, future upgrades of the AVHRR FCDR need to address remaining issues here also for the infrared channels (e.g., recognizing the work of Mittaz et al., 2009).

4 Product definitions

The CLARA-A3 data record based on AVHRR provides global coverage of a number of cloud parameters. Instantaneous, pixel-based cloud retrievals at original AVHRR GAC swath level are used to derive the spatio-temporally sampled (level-2b and level-3 histograms) and averaged (level-3 daily and monthly mean) products. The products are available as daily and monthly composites for each satellite on a regular latitude/longitude grid with a spatial resolution of 0.05° × 0.05° degrees (level-2b), 0.25° × 0.25° (level-3 averages and 1d histograms), and 1.0° × 1.0° (JCH 2d-histograms).

4.1 CLARA-A3 level-2b products

The level-2b format is motivated by the inhomogeneous global coverage of polar sun-synchronous satellite data. Each polar satellite offers 14 observations per day near the pole (evenly distributed over the day) while when passing the equator each location is only observed twice (approximately 12 hours apart). The idea with the introduction of the level-2b data representation is to form a more homogeneous data record having only two observations per day per satellite for each location globally. The alternative to use all observations for level-3 products results in a very skew distribution of the observations because of the inhomogeneous observation frequency (increasing with latitude). By selecting only the observations which are made closest to zenith (the NADIR condition) we ensure that observations are made at almost the same viewing conditions and, most importantly, observations are made at nearly the same local time globally for each level-2b product. In this way, the restricted level-2b products are easier to deal with for certain applications compared to the full set of observations. This concerns in particular the use in COSP simulators aiming at reproducing satellite datasets from Climate Model data.

The level-2b approach means obviously a significant reduction of the amount of used observations. However, the high observation frequency near the poles is undoubtedly very valuable. Therefore, cloud fraction data based on all available observations are also available on two equal-area polar grids at 25 km resolution for the Arctic and Antarctic regions, respectively. These grids are centred at the poles and cover areas of 1000km x 1000km.

All mentioned products exist separately for each satellite. The daily and monthly averages as well as the histograms, however, are also available in aggregated form (i.e., merging all satellites). Acknowledging the different observation capabilities during night and during day and also taking into account existing diurnal variations in cloudiness, a further separation of results into daytime and night-time portions has also been done for some products. Here, all observations made under twilight conditions (solar zenith angles between 75-95 degrees) have been excluded in order to avoid being affected by specific cloud detection problems occurring in the twilight zone.

The temporal coverage of all products ranges from January 1979 to the present. Notice (as visualised in Figure 3-1) that compared to earlier CLARA editions, the afternoon satellite TIROS-N as well as the morning satellites NOAA-6, NOAA-8 and NOAA-10 are now included. A special challenge for these satellites is that they carry AVHRR-1 which has only 4 instead of 5 spectral channels.

Table 4-1 gives an overview of the CLARA-A3 product suite, containing both daily sampled (level-2b) as well as daily and monthly averaged cloud properties and additionally also monthly histograms.

A complete description of the retrieval methods for each individual product is given in the Algorithm Theoretical Basis Documents [RD 2] to [RD 6]. The general methods for AVHRR calibration and compilation of level-2b and level-3 products are described in [RD 2].

4.2 CLARA-A3 level-3 products

Regarding level-3 products it should be emphasized that the method for defining level-3 products means that high resolution level-2b data are first averaged into daily means and then all daily means are averaged into monthly means. This follows the GEWEX Cloud Assessment approach (Stubenrauch et al., 2012). Furthermore, for liquid and ice water path (LWP and IWP) the daily mean is calculated as both in-cloud and all-sky values which leads to two different versions of the monthly mean. Since the monthly mean is calculated from the daily means it means that the in-cloud version will give a relatively larger weight to days with low cloud amounts while the all-sky version will give equal weight to all days in a month.

The level-2b approach and the current level-3 definition leads obviously to a significant reduction of the amount of used observations. However, the high observation frequency near the poles is undoubtedly very valuable and consequently there are also some polar products added which use all available observations.

All products have been evaluated with respect to requirement goals defined in [AD 1]. Of specific interest here are requirements in [AD 2] as outlined by the Global Climate Observing System (GCOS) community and issued by the United Nations World Meteorological Organisation (WMO) in 2012. The finally achieved product accuracies are described in [RD 1]. All products in the CLARA-A3 cloud data record fulfil GCOS requirements regarding the horizontal resolution which is mainly explained by the desire to serve also applications in regional climate modelling and in regional climate monitoring. The GCOS requirement on a temporal resolution of 3 hours is not reachable globally for a data record based on polar orbiting satellite data. This resolution is only achieved for high-latitudes and for the Polar Regions. In the tropical region the temporal resolution is close to 6 hours. All GCOS accuracy requirements are generally fulfilled for all cloud products although the requirements on stability have been assessed only to a limited extent since this is inherently difficult without having a true reference data record. Summaries on validation results are given in each parameter subsection in Section 5, and in Section 6 the overall compliance to the requirements is presented.

Table 4-1: The CLARA-A3 cloud products feature a wide range of global composites data (level-2b), daily and monthly averages as well, as monthly histograms, with liquid and ice cloud separation done for some products. Day and night separation are indicated for products having such a distinction but it should be noted that LWP and IWP products (and as a consequence also JCH) are only retrieved during daytime. Level-2b refers to the non-averaged, pixel-based cloud retrievals sampled onto a global lat/lon grid. All products listed exist separately for each satellite, but also aggregated as so-called 'AVPOS' products. The AVPOS products combine all available satellites, while for CPH day, LWP, IWP and JCH only the satellites having channel 3b active during daytime are included.

	level 2b global[1] 0.05° lat/lon grid	**Daily mean global**[3] 0.25° lat/lon grid	**Monthly mean global**[2] 0.25° lat/lon grid	**Monthly histograms global**[2] 0.25° lat/lon grid	**Daily mean Polar**[3] 25 km EASE grid	**Monthly mean Polar**[3] 25 km EASE grid
CFC	✓ as CMA*	✓ day/night high/mid/low	✓ day/night high/mid/low	-	✓ day/night high/mid/low	✓ day
CTO (CTH, CTP, CTT)	✓	✓ day/night liquid/ice	✓ day/night liquid/ice	✓ liquid/ice	✓ day/night liquid/ice	✓ day liquid/ice
CPH	✓	✓ day/night	✓ day/night	-	-	-
LWP (+τ, r_e, CDNC, CGT)	✓ as CWP	✓	✓	✓ as CWP	-	-
IWP (+τ, r_e)	✓ as CWP	✓	✓	✓ as CWP	-	-
JCH	N/A	N/A	N/A	✓ liquid/ice (1.0° grid)	N/A	N/A

** CMA is a level-2b binary cloud mask that also contains a data layer the probabilistic cloud mask.*
[1] Products exist for individual satellites only
[2] Products exist for individual satellites and as AVPOS (all satellites combined) product
[3] Products exist as AVPOS (all satellites combined) product only

5 Product description

In this section, each cloud product is shortly described regarding its definition, retrieval method, information content and limitations. Validation results are also described shortly for each cloud product. A summary statement on the validation results can also be found in Section 6. More details on achieved validation results are given in [RD 1]. At the end of each product description a short statement on recommended applications areas is given.

5.1 Fractional cloud cover [CM-11012, CM-6011, CFC]

This product is derived directly from results of a cloud detection method provided as a binary cloud mask in level-2b (daily sampled on a global 0.05° grid) and mean cloud fraction in level-3 (on a global 0.25° lat/lon grid). The cloud fractional cover is defined as the fraction of cloudy pixels per grid square compared to the total number of analysed pixels in the grid square. Fractional cloud cover is expressed in percent. The probabilistic cloud mask (named CMAPROB), which is the basis for the binary cloud mask, is included in the level-2b product. The probabilistic version gives cloud probabilities from 0 to 100 % making it possible for a user to flexibly select the confidence level of the cloud mask. Figure 5-1 shows example maps of level-2b standard cloud mask and the probabilistic cloud mask as well as level-3 cloud fraction and uncertainty.

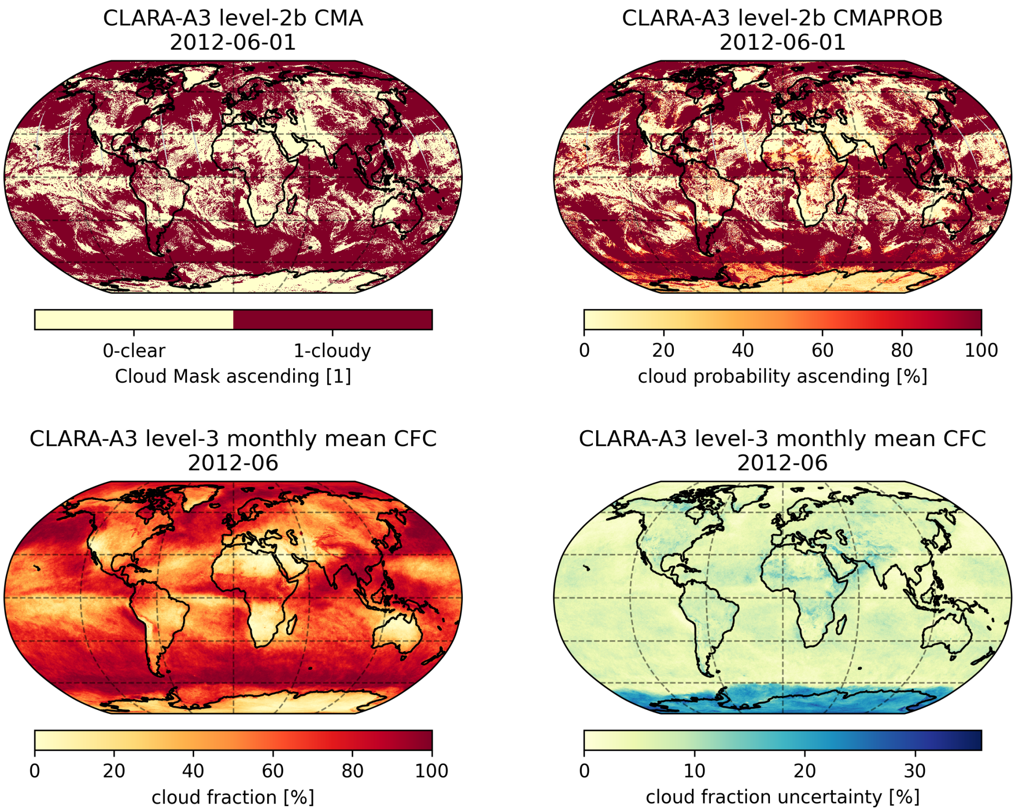

Figure 5-1: Top row: Global map of CLARA-A3 level-2b cloud mask and cloud probability for 2012/06/01 NOAA-19. Bottom row: Global map of CLARA-A3 monthly mean total cloud fraction and uncertainty for 2012/06/01 NOAA-19.

5.1.1 Short Algorithm description

This product is a result of the PPS probabilistic cloud masking procedure based on Bayesian theory. Individual measurement features from all available AVHRR channels (e.g., reflectance at 0.63 µm), are linked to the likelihood of being cloudy, which is inferred from a reliable reference, i.e. the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) in this case. The impact of differences in observation geometry between AVHRR and CALIOP has been thoroughly examined and is discussed in detail by Karlsson and Håkansson (2018). The probabilities of being cloudy for all individual features are multiplied, assuming no correlations, which enables an analytic solution to infer the total probability of a pixel being cloudy. The assumption of no correlation between features is not always fulfilled but this approach has nevertheless been shown to give improved results compared to the previous method applied for CLARA-A2 (as reported by Karlsson et al., 2020). The changed methodology has also led to that the CFC product is now extended to include also an estimation of uncertainty.

All downstream cloud products (including the binary cloud mask) are based on applying a 50% threshold to distinguish between cloudy and clear-sky pixels.

A detailed description is given in [RD 2] and RD 3].

5.1.2 Highlights

- Cloud screening is based on information from all available AVHRR channels
- Thin Cirrus clouds are identified at night using split-window infrared channels and during daytime (and for early AVHRR/1 measurements) from radiances in the 3.7 micron channel.
- Water clouds with cloud temperatures close to surface temperatures are identified at night using 3.7 micron channel information (often not identified by traditional IR methods)
- Observation frequency is high in polar areas – 14 observations per day per satellite
- Cloud detection in polar summer in the Arctic and Antarctic regions shown to be comparable with MODIS cloud data records (Karlsson and Dybbroe, 2010) - thus, the advantage here is the very long observation period
- Efficient polar summer cloud screening allows high-quality estimations of surface albedo (e.g., at the time of minimum ice extent in the Arctic region)
- Daytime conditions with good illumination (i.e., conditions enabling access to information in all spectral channels) provide best cloud screening results
- The probabilistic formulation of the cloud mask in L2b products makes it possible for a user to adjust the criteria for the cloud mask, either making it more clear-conservative or more cloud-conservative.

5.1.3 Limitations

- Not all clouds will be detected due to inherent limitations of the AVHRR imager as being a passive radiometer with a rather coarse field of view (here about 5 km in size). Comparisons with the actively probing lidar instrument CALIOP show that clouds with optical thicknesses below 0.22 are more likely being missed than being detected from AVHRR data.
- Some thin clouds (particularly ice clouds) over cold ground surfaces may remain undetected even if having cloud optical thicknesses higher than the above mentioned detection limit
- Twilight conditions are especially challenging to AVHRR cloud screening methods (especially for morning-evening satellites). On the dark side of twilight some underestimation of cloudiness can be seen whereas on the illuminated side some overestimation of cloudiness has been noted.
- CFC results over the Arctic and Antarctic regions have larger uncertainties during the polar winter when also sampling issues are more frequent (e.g., solar contamination effects). Problems with the calibration accuracy in the 3.7 micron channel for very cold temperatures (especially for satellites NOAA-15, NOAA-16 and Metop-B) have also led to some exclusion (blacklisting) of results during the polar winter.
- The inclusion of data from the early four-channel AVHRR/1 instrument (without the 12 micron channel) has led to a larger sensitivity of cloud fraction results to noise in AVHRR channel 3b. This has given some overestimation of CFC, especially over cold land areas in winter. Cloud products after 2017 (i.e., the year with the latest calibration update) in the CDR and in the ICDR might be affected by increasing uncertainties in the time-dependent corrections of visible radiances. The CFC product is, however, quite robust and not very dependent on the accuracy of visible radiances. Noticeable errors will nevertheless possibly appear in the time frame 2023 or later. A calibration update is expected to solve these potential problems.

5.1.4 Validation

Validation studies based on CALIPSO-CALIOP observations shows that for clouds having a vertically integrated cloud optical thickness greater than 0.22 the overall global bias is -0.1 %. The largest underestimation (10-20 %) occurs over the Polar Regions during the Polar winter season. These results agree very well with corresponding validation studies based on surface stations (SYNOP) showing a global bias of 2 %.

5.1.5 Recommended applications

Despite efforts for inter-calibration and homogenisation, the CM SAF CFC data record may still not be of sufficient quality for offering global climate trend analysis. More work is needed before this status can be reached (see also concluding discussion in section 5). However, for most regions, except the Polar Regions in the Polar winter, validation efforts have shown very good results. It concerns in particular mid-latitude and high-latitude regions but also Tropical land regions. For the Polar Regions, the treatment of polar winter conditions still remains as a very big challenge for AVHRR cloud retrievals. Nevertheless, results during the Polar summer appear to be of good quality and studies related to cloud-sea ice interaction in the Polar summer are encouraged.

5.2 Cloud Top level [CM-11032, CM-6031, CTO]

Three versions of the CM SAF Cloud Top product exist: 1. Cloud Top Temperature (CTT), expressed in Kelvin; 2. Cloud Top Height (CTH), expressed as altitude over ground topography (m); 3. Cloud Top Pressure (CTP), expressed in pressure co-ordinates (hPa). The CTO product is derived in all cloudy pixels as identified by cloud mask product. The data is provided as pixel-based values in level-2b (daily sampled on a global 0.05° grid), as daily and monthly averages (on a global 0.25° lat/lon grid), and monthly sampled histograms (on a global 0.25° lat/lon grid). For the daily and monthly averages, the data are averaged arithmetically (linearly) but for the Cloud Top Pressure product also a geometric mean (i.e., average of the logarithm of cloud top pressure) is available. Figure 5-2 shows example maps of level-2b and level-3 cloud top pressure and associated uncertainties.

Notice that the cloud altitude given by CTH is defined as the mean cloud altitude relative to the topography (GTOPO30) and not relative to the mean sea level. The used topography data set is available to users as an ancillary dataset.

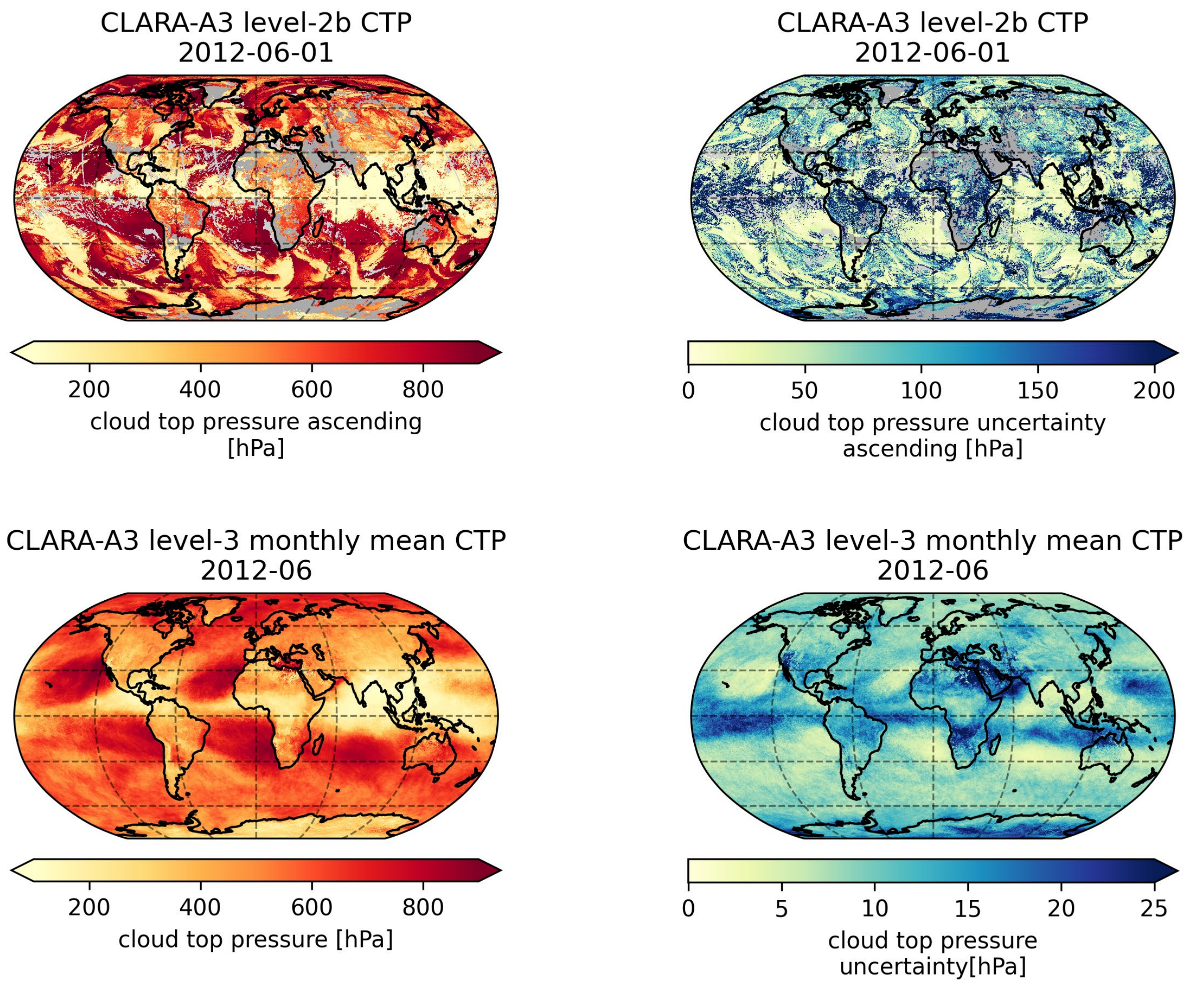

Figure 5-2: Left: Global map of CLARA-A3 level-2b cloud top pressure for 2012/06/01 NOAA-19. Right: Global map of CLARA-A3 monthly mean cloud top pressure for 2012/06/01 NOAA-19

5.2.1 Short Algorithm description.

The CTO product, based on a different methodology compared to the one previously used for CLARA-A2, gives significantly higher cloud top altitudes for high-level clouds and a reduction of altitudes for low-level boundary layer clouds trapped in inversions. The CTP retrieval is now based on the PPS neural network (of the type multilayer perceptron) trained offline with collocations of passive imager measurements (i.e., the three infrared channels of AVHRR) and CALIOP (version 4 data) cloud top pressure observations. Training was performed using collocations over the full globe, thus covering all variations caused by varying viewing angles. Collocated Reanalysis profiles from ERA5/ERA5T are used to convert CTP to CTT and CTH.

In addition to the retrieved values (direct output of the neural network), an estimation of the upper and lower retrieval boundaries (at the 16^{th} and 84^{th} percentile level) are produced using a special version of Quantile Regression Neural Networks (see Pfreundschuh et al. 2018). With this the 68% confidence interval around the retrieved value can be determined which are used to determine the uncertainties for all three CTO variables.

A detailed algorithm description is given in [RD 2] and [RD 4].

5.2.2 Highlights

- Cloud top heights are determined using the closest reference profiles available from the ERA5/ERA5T dataset
- The neural network approach includes using information from surrounding pixels and this leads to improved treatment of both optically thick and thin clouds.
- The largest improvement of results compared to CLARA-A2 is seen for thin clouds in the tropical and subtropical regions.

5.2.3 Limitations

- Cloud top level estimations for boundary layer clouds have often a positive bias, sometimes larger than 500 m, due to problems with resolving boundary layer inversions in reference profiles from ancillary datasets (ERA5/ERA5T)
- As mentioned for the CFC product, optically very thin clouds may not be detected at all. Even if being detected, it is still very difficult to assign a correct cloud top level for the thinnest clouds and this leads to some remaining underestimation of cloud top levels despite the improvement since CLARA-A2.
- The current method is only applied to the uppermost cloud layer.
- A suspected cold bias for infrared channels for satellites Tiros-N and NOAA-6 has also been observed leading to anomalous cold cloud top temperatures in the beginning of the CLARA-A3 time series. The causes of this deviation will be investigated prior to next CLARA edition.

- Cloud products after 2017 (i.e., the year with the latest calibration update) in the CDR and in the ICDR might be affected by increasing uncertainties in the time-dependent corrections of visible radiances. However, the CTO product is not using visible information and will only be affected indirectly by potential problems for the cloud mask (i.e., the CFC product – see section 5.1.3 above).

5.2.4 Validation

Validation studies (detailed in [RD 1]) show globally a mean negative bias (underestimation) of -898 m, dominated by large underestimations of the altitude of high, semi-transparent clouds, although partly compensated by an overestimation of low-level boundary layer clouds. Stratified by cloud level, the biases amount to +277 m (low-level clouds), +117 m (mid-level clouds) and -1781 m (high-level clouds), respectively.

5.2.5 Recommended applications

CTO is an important cloud feature for many climate applications since it largely determines the effective height of thermal emission to space from the cloudy portions of the atmosphere. However, it is clear that some cloud tops of boundary layer clouds are still slightly overestimated in height in the CLARA-A3 data record, mainly because of problems in having good enough background reference profiles available from NWP reanalysis datasets. We also still see some underestimation of the highest cloud tops for thin clouds. However, the improvements of the CTO product are substantial and CLARA-A3 is now likely to provide the best 3-dimensional cloud distribution over four decades compared to other datasets.

5.3 Cloud Phase [CM-11042, CM-6041, CPH]

The cloud phase product is meant to represent the thermodynamic phase of the particles near the cloud top. This product is provided as a binary cloud phase (liquid or ice) as well as extended cloud phase/type information in level-2b (daily sampled on a global 0.05° grid), and mean liquid cloud fraction in level-3 (on a global 0.25° lat/lon grid). The liquid cloud fractional cover is defined as the fraction of liquid cloud pixels per grid square compared to the total number of cloudy pixels in the grid square. Liquid cloud fraction is expressed in percent. Figure 5-3 shows example maps of level-2b binary and extended cloud phase and level-3 monthly mean and standard deviation of the liquid cloud fraction.

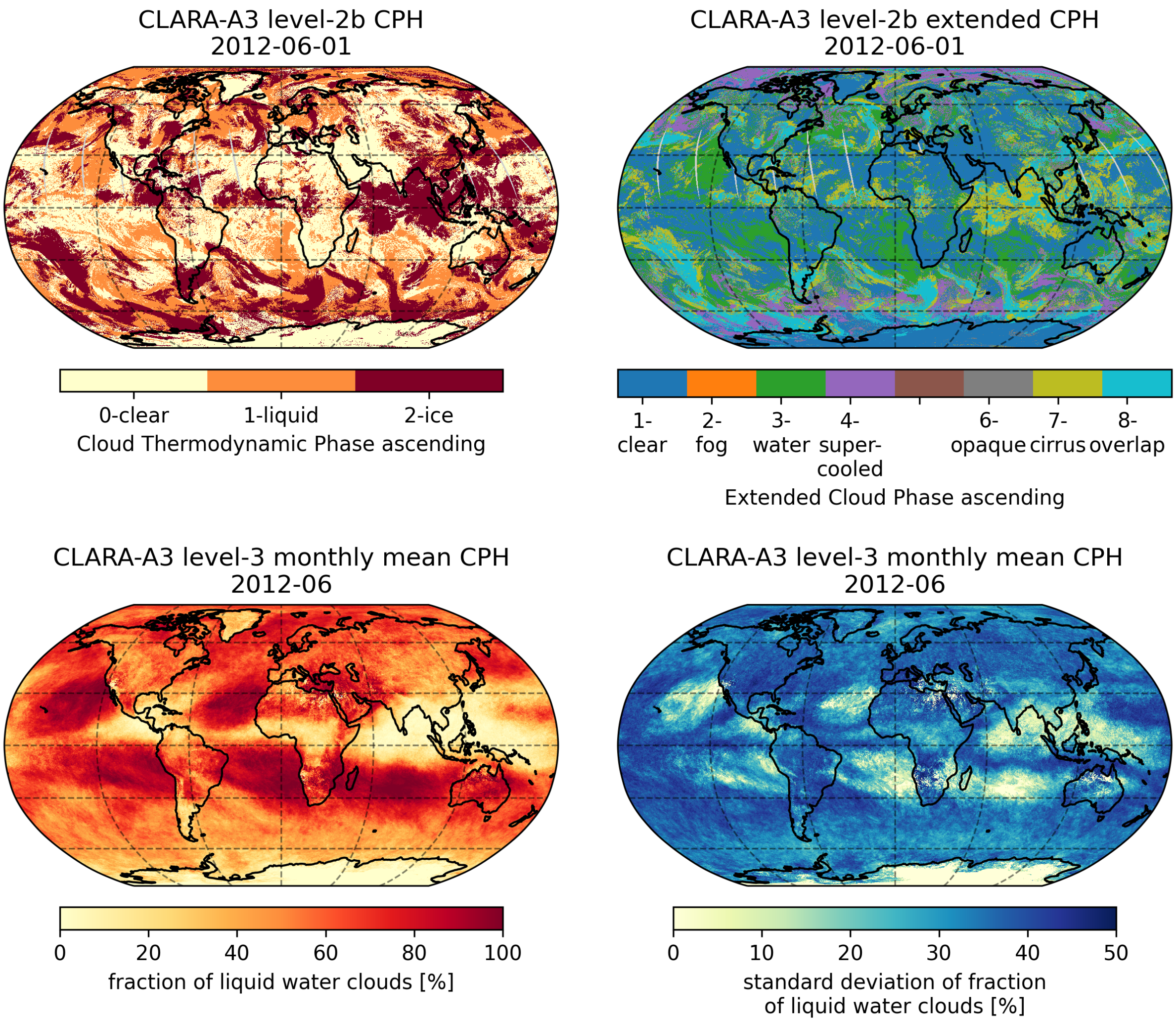

Figure 5-3: Top row: global maps of CLARA-A3 level-2b cloud phase (left) and extended cloud phase (right) for 2012/06/01 NOAA-19. Bottom row: global maps of CLARA-A3 monthly mean (left) and standard deviation (right) of liquid cloud fraction for 2012/06/01 NOAA-19.

5.3.1 Short Algorithm description

The cloud-top thermodynamic phase (CPH) is determined from a cloud typing approach following Pavolonis and Heidinger (2004) and Pavolonis et al. (2005). This algorithm consists of a series of spectral tests applied to infrared brightness temperatures. It has a night-time branch as well as a daytime branch in which shortwave reflectances are considered in addition. The algorithm initially yields one of the following cloud types: liquid, supercooled, opaque ice, cirrus, and overlap. These are then further condensed to liquid (former two) and ice (latter three) phase.

5.3.2 Highlights

- The algorithm provides cloud phase both during daytime and night-time.
- In addition to liquid/ice discrimination a further breakdown into cloud types is provided.
- The phase discrimination shows good agreement with active CALIOP observations.

5.3.3 Limitations

The main limitations of the CPH retrieval are:

- The identification of thin cirrus over water clouds is challenging.
- Due to the nature of passive satellite observations, the phase near the top of the clouds is retrieved with limited sensitivity to lower cloud layers.
- Cloud products after 2017 (i.e., the year with the latest calibration update) in the CDR and in the ICDR might be affected by increasing uncertainties in the time-dependent corrections of visible radiances. The daytime CPH product could be affected by this and by the indirect dependence on the cloud mask (i.e., the CFC product). A calibration update is expected to solve these potential problems.

5.3.4 Validation

The CPH product was extensively compared with other VIS-IR imager-based data records, showing an overall reasonable degree of consistency with for example MODIS, with a global multi-year mean difference in liquid cloud fraction of about -4.9 % (absolute units). Validation was performed against CALIPSO observations. This yielded a liquid cloud fraction bias of -2% and a KSS 0.67. The stability with respect to MODIS was evaluated to be 2.5 %/decade. For more detailed validation results, see [RD 1].

5.3.5 Recommended applications

The CPH data record is specifically useful for studies of cloud glaciation, e.g. in the life cycle of convective activity characterized by a transition from liquid to ice phase as the clouds grow vertically. In general, wherever multi annual, stable and spatially highly resolved information is needed, CPH data from CLARA-A3 can be applied.

Care should be taken when using the data record for long-term trend analysis. Despite extensive calibration efforts, inter-satellite discontinuities still exist, and trend analysis is further complicated by the orbital drift of the satellites.

5.4 Liquid Water Path [CM-11052, CM-6051, LWP]

This product is the vertical mass integral of liquid cloud particles per area. LWP is provided as pixel-based cloud water path (CWP) in level-2b (daily sampled on a global 0.05° grid) grid cells that are assigned a liquid cloud, as monthly averages, and monthly sampled histograms (on a global 0.25° lat/lon grid). The daily and monthly LWP averages are given as so-called in-cloud values (only averaging over the liquid cloud pixels) and as so-called 'allsky' or 'grid-mean' values (for which the non-liquid cloud pixels are accounted for with an LWP of 0). The latter usually correspond to the output from atmospheric models. Figure 5-4 shows example maps of level-2b and level-3 liquid water path with corresponding uncertainty estimates.

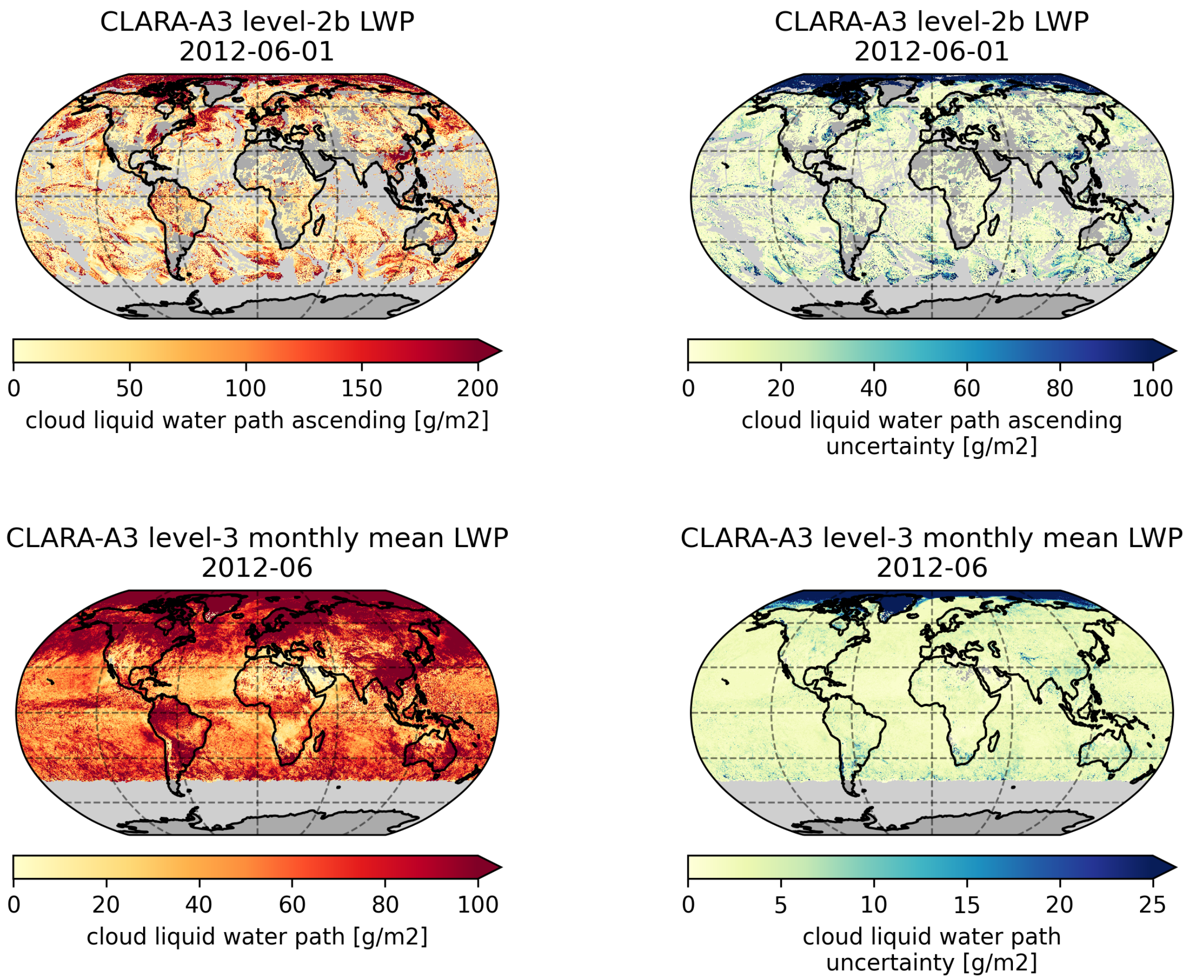

Figure 5-4: Top row: global maps of CLARA-A3 level-2b cloud liquid water path (left) and its uncertainty (right) for 2012/06/01 NOAA-19. Bottom row: global maps of CLARA-A3 monthly mean in-cloud liquid water path (left) and its uncertainty (right) for 2012/06/01 NOAA-19.

5.4.1 Short Algorithm description

The central principle of the method to retrieve cloud optical and microphysical properties is that the reflectance of clouds at a (for cloud particles) non-absorbing wavelength in the visible region (e.g., 0.6 or 0.8 μm) is strongly related to the optical thickness (τ) and has little dependence on particle effective radius (r_e), whereas the reflectance of clouds at an absorbing wavelength in the shortwave-infrared region (e.g., 1.6 or 3.7 μm) is strongly dependent on effective radius (Nakajima and King, 1992).

In the CPP algorithm (Benas et al., 2017; Roebeling et al. 2006), the Doubling-Adding KNMI (DAK) radiative transfer model (De Haan et al. 1987 and Stammes 2001) is used to simulate 0.6-, 1.6, and 3.7-μm top-of-atmosphere reflectances of homogeneous, plane-parallel clouds as a function of viewing and illumination geometry, cloud optical thickness, effective radius, and cloud phase. These simulated reflectances are stored in a look-up table (LUT).

τ and r_e are retrieved for cloudy pixels in an iterative manner by matching satellite-observed reflectances to the LUT of RTM-simulated reflectances. For water clouds effective radii between 3 and 34 µm are retrieved, while τ is limited at 150. From these two properties, the cloud water path (CWP) of water clouds (or liquid water path, LWP) can be computed using the following relation assuming vertically homogeneous liquid water content (Stephens, 1978):

$$\mathrm{CWP} = 2/3\ \rho_l\ \tau\ \mathrm{r_e}, \qquad (1)$$

where ρ_l is the density of liquid water.

Dependent on the available AVHRR channels, τ, r_e and LWP are derived from the 0.6-1.6 µm channel combination or from the 0.6-3.7 µm channel combination. In the latter case, and now assuming semi-adiabatically stratified clouds, the (vertically constant) cloud droplet number concentration (CDNC) and cloud geometrical thickness (CGT) are derived, following Bennartz and Rausch (2017).

5.4.2 Highlights

- Together with LWP, τ, r_e, CDNC and CGT are included as additional layers in the level-2b and level-3 products.
- Estimates of the retrieval errors are provided taking into account a range of error sources.
- The careful calibration of the shortwave AVHRR channels has a pronounced effect on the quality of the retrieved τ and r_e, and thus LWP.

5.4.3 Limitations

The main limitations of the LWP retrieval are:

- The derivation of cloud physical properties from reflected solar radiation is dependent on the availability of daylight. Thus, no retrievals can be done during night time. Even if pixel-level retrievals are performed and reported up to solar zenith angles of 84°, for level-3 aggregation a maximum solar zenith angle of 70° is applied.

- Sunglint can affect the cloud property retrievals considerably, in particular for broken cloudy scenes over ocean. Therefore, possibly sunglint-affected pixels (defined by a scattering angle differing less than 27 degrees from the direct glint angle) are flagged.

- The retrieval is highly problematic over very bright surfaces, particularly ice and snow, as the visible reflectance from clouds is similar to that from the surface. This is reflected by large estimated uncertainties (see polar regions in Figure 5-4).

- Cloud property retrievals are performed assuming that clouds are plane parallel. Two prominent examples of cases for which this assumption is violated are: (1) three-dimensional radiative effects become important if large sub-pixel variations in cloud-top height occur, and particularly if the solar zenith angle is large; (2) retrievals for broken clouds are affected by a reflectance contribution from the surface.

- Aerosols are not considered in the CPP retrieval. This assumption is usually justified because aerosols reside below or within the cloud and their optical thickness is small compared to that of the cloud. However, if the aerosols reside above the cloud and if they are sufficiently absorbing, they can significantly lower the visible reflectance. This typically leads to underestimations of both τ and r_e, and thus LWP (Haywood et al. 2004).

- Retrievals from passive satellite instruments are limited by the fact that the obtained signal emanates from the integrated profile. Since near-infrared radiation is only penetrating into the cloud to a certain depth (due to absorption by cloud particles), the retrieved effective radius is representative for the upper part of the cloud (Platnick, 2001). The penetration depth depends on the amount of absorption by cloud particles, which is increasing with wavelength. This means that the retrieved r_e depends on which NIR spectral channel is used (1.6 μm or 3.7 μm). To mitigate inhomogeneities in the all-satellite (AVPOS) level-3 cloud optical and microphysical products, these only include retrievals based on the 3.7 μm channel.

- Cloud products after 2017 (i.e., the year with the latest calibration update) in the CDR and in the ICDR might be affected by increasing uncertainties in the time-dependent corrections of visible radiances. The LWP product is particularly sensitive here due to the dependence on both AVHRR channel 1 and channel 3a. Noticeable deviations or trends have been observed for several of the active satellites after 2017. A calibration update is expected to solve these potential problems.

5.4.4 Validation

The LWP product was extensively compared with other VIS-IR imager-based data records, showing an overall reasonable degree of consistency with for example MODIS, with a global multi-year mean all-sky LWP difference of -5.7 g m^{-2}. Differences are largest in high-latitude regions. Comparison with the passive microwave-based MAC-LWP data record for marine stratocumulus areas yielded a mean difference of -0.7 to +8.1 g m^{-2}. The stability with respect to MODIS was evaluated to be 1.8 g m^{-2} $decade^{-1}$. For more validation results, see [RD 1].

5.4.5 Recommended applications

The LWP data record can be used for climate studies (e.g., Benas et al., 2020) and model evaluation. The LWP retrievals are most reliable at lower latitudes, or (better said) lower solar zenith angles (below about 65 degrees). In addition, high latitudes are frequently affected by snow and ice cover, making it difficult to estimate cloud optical properties against the bright background. Care should be taken when using the data record for very-long-term trend analysis. Despite extensive calibration efforts, inter-satellite discontinuities do occur, and trend analysis is further complicated by the orbital drift of the satellites. Significant differences are observed between the parts of the data record relying on the 1.6-µm and 3.7-µm channels. It is recommended that those parts are not combined but used separately for analyses.

5.5 Ice Water Path [CM-11062, CM-6061, IWP]

This product is the vertical mass integral of ice cloud particles per area. IWP is provided as pixel-based cloud water path (CWP) in level-2b (daily sampled on a global 0.05° grid) grid cells that are assigned an ice cloud, as monthly averages, and monthly sampled histograms (on a global 0.25° lat/lon grid). The daily and monthly IWP averages are given as so-called in-cloud values (only averaging over the ice cloud pixels) and as so-called ‘allsky’ or ‘grid-mean’ values (for which the non-ice cloud pixels are accounted for with an IWP of 0). The latter usually correspond to the output of atmospheric models. Figure 5-5 shows example maps of level-2b cloud ice water path and level-3 monthly mean ice water path with corresponding uncertainty estimates.

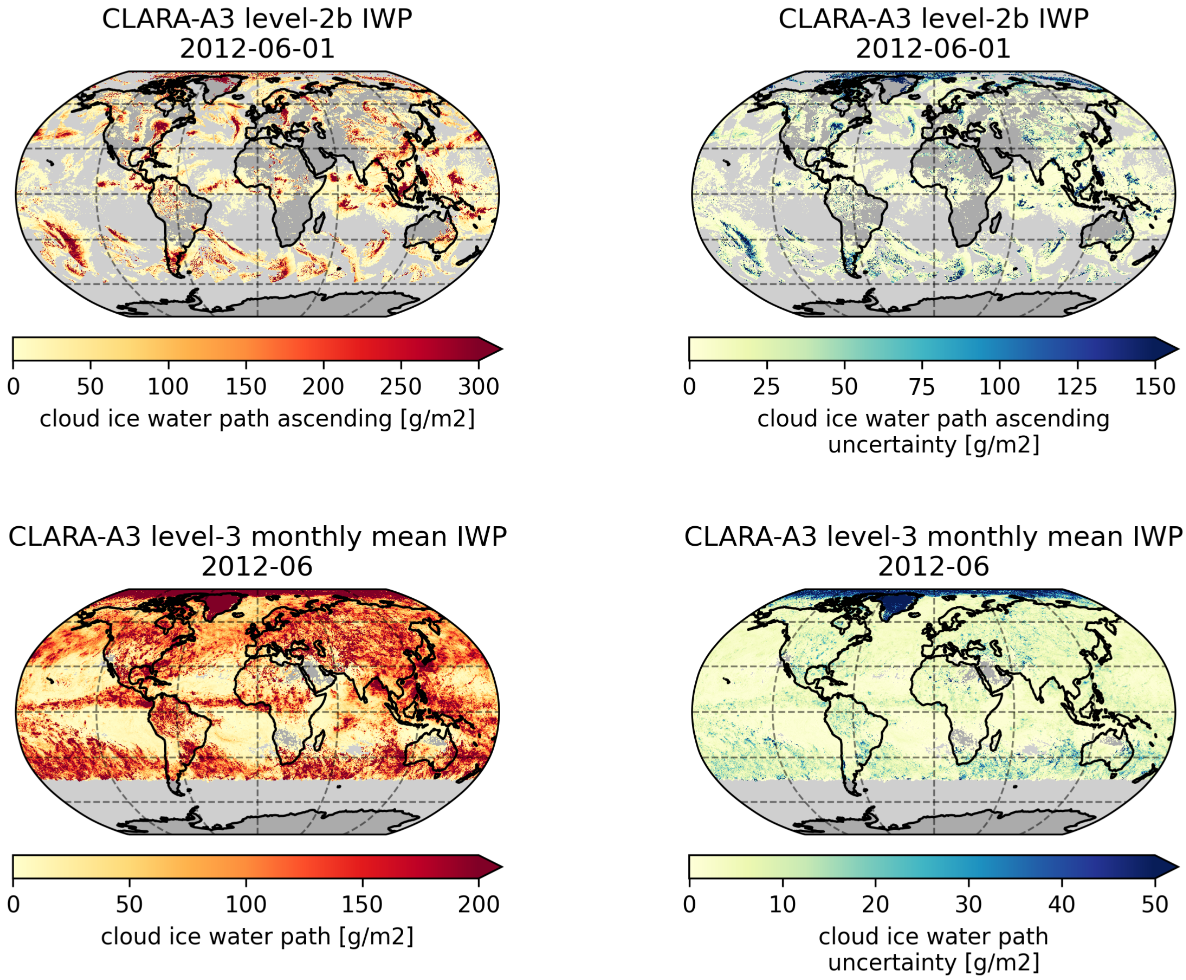

Figure 5-5: Top row: global maps of CLARA-A3 level-2b cloud ice water path (left) and its uncertainty (right) for 2012/06/01 NOAA-19. Bottom row: global maps of CLARA-A3 monthly mean in-cloud ice water path (left) and its uncertainty (right) for 2012/06/01 NOAA-19.

5.5.1 Short Algorithm description

Ice water path is retrieved in the same way as LWP but with τ and r_e retrievals based on RTM simulations collections of severely roughened aggregated solid columns (Yang et al., 2013 and Baum et al., 2011) with effective radii ranging between 5 and 60 μm.

5.5.2 Highlights

The highlights mentioned for LWP also hold for IWP.

5.5.3 Limitations

The same limitations as for LWP hold also for IWP. In addition, the r_e retrieval for ice clouds is considerably more uncertain than for water clouds, because particle shapes and roughness vary widely and are not well known. The assumptions on ice crystal habits used to generate the LUTs have a profound impact on the retrieved r_e and IWP. Retrievals of multi-layer clouds are problematic. Typically, a scene with cirrus overlying low liquid water clouds will be assigned the ice phase, and the retrieved COT and IWP will include contributions from both cloud layers.

5.5.4 Validation

The IWP product was extensively compared with other VIS-IR imager-based data records, showing a global multi-year mean all-sky IWP difference of -0.4 g m^{-2} and stability of -0.8 g m^{-2} $decade^{-1}$ with respect to MODIS.

5.5.5 Recommended applications

The IWP data record is just like LWP most reliable at lower latitudes, or lower solar zenith angles below about 65 degrees. In addition, high latitudes are frequently affected by snow and ice cover. In these cases, retrievals are problematic. IWP can like LWP be applied where multi annual and spatially highly resolved information is needed. Care should be taken when using the data record for very-long-term trend analysis because satellite switches and orbital drift compromise its homogeneity.

5.6 Joint Cloud property Histograms [CM-11022, CM-6021, JCH]

The JCH product is a combined histogram of CTP and COT covering the solution space of both parameters. This two-dimensional histogram gives the absolute numbers of occurrences for specific COT and CTP combinations defined by specific bins. It is further separated in liquid and ice clouds. An example of JCH is shown in Figure 5-6.

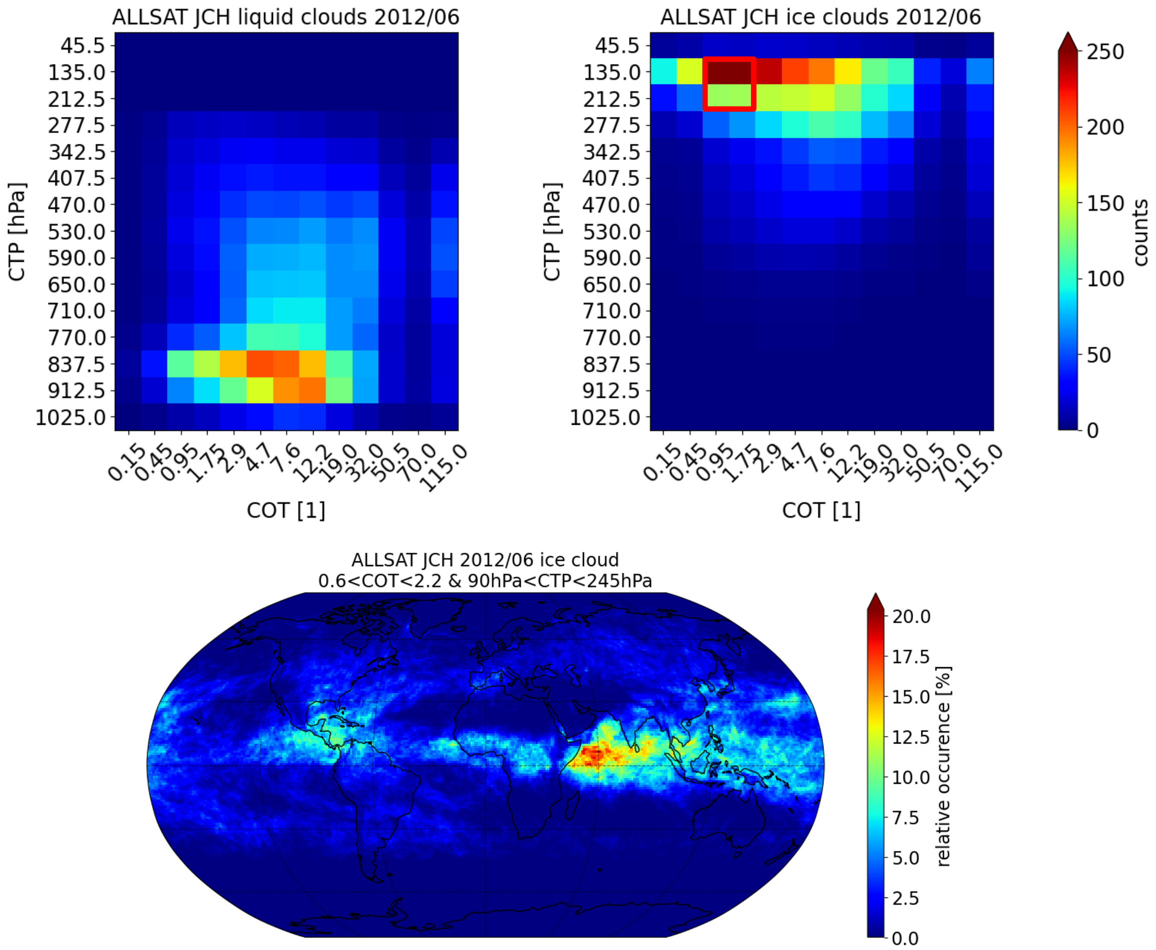

Figure 5-6: COT-CTP histogram for 2012/06. Top: Global average for liquid clouds (top-left) and ice clouds (top-right). Bottom: Global map of relative occurrence of ice clouds with CTP between 90 hPa and 245 hPa and COT between 0.6 and 2.2 (indicated by the red rectangle in the top-right plot).

5.6.1 Short Algorithm description

No specific retrieval algorithm is needed for this product. It is composed by collecting pixel-level COT and CTP retrieval values as counts in pre-defined bins ([RD 2]) on monthly basis. No averaging is done in the course of this process.

5.6.2 Highlights

- The product adds value to the single standard level-3 products of CTP and COT by showing how the two parameters vary together.
- The histograms are given for each grid point which means that a user can aggregate results over any local or regional domain in order to analyse typical cloud regimes (or types).
- The use of joint histograms is common in applications for evaluating climate models.
- The CLARA-A3 COT and CTP bins for the JCH product have been chosen such that they are a superset of the traditional ISCCP bins. Thus, the CLARA-A3 product can be aggregated to exactly the same bins as used for the ISCCP product.

5.6.3 Limitations

- The product is only available during daytime since the COT is not retrieved at night.

5.6.4 Recommended applications

With the JCH diagrams the discrimination between different cloud regimes is supported. COT and CTP information can readily be considered in a more statistical way. Since the CM SAF JCH product is compiled as distributions for each grid point, it is possible to compose cloud distribution statistics for any region size on the globe by simple aggregation of grid point values. Two-dimensional COT-CTP histograms, such as JCH, are nowadays standard outputs of the COSP satellite simulator packages based on climate modelling data. Thus, the evaluation of climate models is a suited application of the JCH product.

5.7 ICDR specifics

For the operational continuation of the CDR, the ICDR, the exact same algorithms are used. The ICDR, however, is run in a mode with only a few days latency, which requires compromises with respect to input data. While the by far most significant source of CDR and ICDR inconsistencies in the past, the use of different NWP input data, has been removed by using ERA5 for the CDR and ERA5T for the ICDR, remaining auxiliary data slightly differ between CDR and ICDR processing (see Table 3.2 in RD 2]). Furthermore, the level-1 data was provided by EUMETSAT for the CDR, while CM SAF is using the same framework for ICDR but implemented at its processing facilities. The level-1 processing includes the application of latest AVHRR calibration information from 2017 (see Section 3.2 and Section 2.3.1 in [RD 2]), which however is not further updated on a regular basis in the course of the ICDR, which could create calibration inaccuracies in case of abrupt or long-term AVHRR sensor degradations. Hence, it was decided to exclude results from Metop-C among the ICDR products. However, an update of the calibration is expected in 2023 or later to allow inclusion of Metop-C data in the ICDR processing chain. This will also improve the calibration for other active satellites at that time as a consequence of enhanced quality of time-dependent calibration corrections.

In Section 6.3 of [RD 1] the ICDR has been evaluated against the CDR for a 6-month overlap period. In could be demonstrated that the ICDR has a high agreement with the CDR with differences in CFC, CPH and CTP, caused mainly by the exclusion of Metop-C from the ICDR, and with small deviations found for cloud water path over sea-ice which are primarily caused by slightly deviating sea-ice input data for CDR and ICDR.

Beware that possible changes in the ICDR input data streams and a decay with time of calibration corrections (before the next update) can lead to reduced quality in the CM SAF ICDR products, with partly distinct regional imprints and exceeding the effective ICDR Service Specifications. Please use the data and associated plots with care and check the Service Messages and the Annual Validation Reports (AQARep) of the ICDR (https://www.cmsaf.eu).

6 Summary table of validation results regarding product accuracy

Table 6-1 shows the evaluation results (bias and bc-RMSE) and their compliance with the accuracy and precision requirements. This table is a copy of Table 1-1 from [RD 1].

Table 6-1: Summary of CLARA-A3 validation results compared to requirements for each cloud product. Required and achieved accuracies are formulated in terms of bias, precisions in terms of bc-RMSE (except for CFC and CPH L2, for which the metric is the Hanssen-Kuipers Skill Score, KSS), and stabilities in terms of decadal trend in bias. All numbers, except KSS, are in the units indicated for the respective cloud products. Validation results are color-coded as follows: worse than threshold, fulfils threshold, fulfils target, and fulfils optimal requirement. CALIOP results are reported for different values of (ICOT). Evaluations against similar datasets are indicated in grey since they are viewed as consistency checks rather than true validation.

L2 / L3	Reference	Accuracy (bias)		Precision (bc-RMSE)		Stability (decadal trend in bias)	
		Req'd (target)	Ach'd	Req'd (target)	Ach'd	Req'd (target)	Ach'd
Cloud Fractional Cover (CFC) [%]							
L2	CALIOP (COT > 0)	5	-11.1	0.6 (KSS)	0.67	n.a.	n.a.
	CALIOP (COT > 0.217)		-0.1		0.69		
L3	SYNOP	5	2.04	10	8.77	2	-1.27
	CALIOP (COT>0)		-5.01		8.9		-
	CALIOP (COT > 0.3)		2.43		6.9		
	MODIS		-		-		0.0
Cloud Top Height (CTH) [m]							
L2	CALIOP (ICOT > 0)	800	-898	2400	3236	n.a.	n.a.
	CALIOP (ICOT > 0.4)		250		1244		
Cloud Top Pressure (CTP) [hPa]							
L3	CALIOP (ICOT > 0)	45	27	85	41	15	-
	CALIOP (ICOT > 0.3)		-50		66		
	MODIS		-		-		-4.7
Cloud Phase (CPH), defined as fraction of liquid clouds [%]							
L2	CALIOP (ICOT > 0)	5	-2	0.6 (KSS)	0.67	n.a.	n.a.
L3	MODIS	5	-4.9	10	13.1	2	2.5
	Cloud_cci		-4.3		10.1		-0.6
Liquid Water Path (LWP) [$g\ m^{-2}$]							
L3	MAC-LWP	10	(-0.7)-8.1	20	11.7-16.1	3	(-6.9)-(-4.8)
	MODIS		-5.7		13.4		1.8
	Cloud_cci		-9.3		16.4		-0.1
Ice Water Path (IWP) [$g\ m^{-2}$]							
L3	MODIS	20	-0.4	40	22.5	6	-0.8
	Cloud_cci		-15.4		33.1		1.2
Joint Cloud property Histogram (JCH)							
L3		n/a		n/a		n/a	

7 Outlook

The CM SAF CLARA-A3 record is the result of the third reprocessing of global AVHRR data in the CM SAF project. The generation of the CLARA data records is a sustained effort. In a few years time a revision of CLARA-A3 is planned (CLARA-A3.5). While all parts of the retrieval and processing components undergo continuous enhancements for future CLARA editions, CLARA-A3.5 will potentially feature (a) improvements for the AVHRR-1 instrument and (b) include VIIRS GAC data to strengthen the afternoon component of the CLARA record for the years 2012 and beyond for which the NOAA afternoon satellites have significantly drifted away from the initial orbits.

8 Data format description

CLARA-A3 products are provided as NetCDF (Network Common Data Format) files (http://www.unidata.ucar.edu/software/netcdf/). The data files are created following NetCDF Climate and Forecast (CF) Metadata Convention version 1.7 (http://cf-pcmdi.llnl.gov/) and NetCDF Attribute Convention for Dataset Discovery version 1.3.

A common NetCDF file consists of dimensions, variables, and attributes. These components can be used together to capture the meaning of data and relations among data.

For data processing and conversion to various graphical packages input format, CM SAF recommends the usage of the climate data operators (CDO), available under GNU Public License (GPL) from MPI-M (http://www.mpimet.mpg.de/~cdo).

8.1 Common variables and attributes

In all CLARA-A3 level-2b and level-3 products the same set of general variables are used. They are listed in Table 8.1. Level-2b products contain an additional field *scanline_time* which is only applicable for non-averaged products.

Table 8-1: Common variables in level-2b and level-3 products.

Name	Description
time	observation timestamp in days since 1970-01-01 00:00 UTC
lon	geographical longitude of grid-box centre [degree_east]
lat	geographical latitude of grid-box centre [degree_north]
time_bnds	two-dimensional array defining the first and last point in time of the observation
lon_bnds	two-dimensional array defining geographical longitude of grid-box boundaries
lat_bnds	two-dimensional array defining geographical latitude of grid-box boundaries
record_status	overall status of each record in the file: 0: “ok”, 1: “void”, 2: “bad_quality”
scanline_time[1]	actual time of the pixel observation as fractional hours of the day

[1] *in level-2b products only*

Table 8-2 provides a list of general attributes from which only first four attributes are obligatory for each field. The other attributes are included if applicable, for example byte fields don’t need a scaling factor. Flag attributes are only applicable for non-averaged products.

Table 8-2: Common attributes of each variable in level-2 and level-3 products.

Name	Description
_FillValue	*This number represents missing or undefined data. Missing values are to be filtered before scaling*
valid_min	*Lowest valid value*
valid_max	*Largest valid value*
long_name	*Long descriptive name*
standard_name	*Standard name that references a description of a variable's content in the CF standard name table*
scale_factor	*The data are to be multiplied by this factor after it is read*
add_offset	*This number is to be added to the data after it is read. If scale_factor is present, the data are first scaled before the offset is added.*
units	*Physical unit [udunits standards]*[1]
flag_values[2], *flag_masks*[2], *flag_meanings*[2]	These attributes describe variables containing status flags. The *flag_values/flag_masks* attributes describe all possible values of the status flag, whereas *flag_meanings* indicates the meaning of a certain flag value. See section 3.5 in the CF standard for details.

[1] *All cloud properties are provided in SI units wherever possible and appropriate.*

[2] *in level-2b products only*

The absolute value of the respective variable Y is composed of the value in the data field x together with scale_factor s and the offset b:

$$Y = s * x + b \tag{2}$$

8.2 Data format description of non-averaged products

Level-2b data are stored in on a regular lat-lon grid with a spatial resolution of 0.05°. Each specific variable is stored twice, once for the ascending node and once for the descending node.

8.2.1 Product specific data fields

8.2.1.1 Cloud mask (CMA)

cma(time, lat, lon)	field contains the binary cloud mask
cma_prob(time, lat, lon)	probability that a cloud is observed in this pixel

8.2.1.2 Cloud top (CTO)

ctt(time, lat, lon)	cloud top temperature
ctp(time, lat, lon)	cloud top pressure
cth(time, lat, lon)	cloud top height
ctt_unc(time, lat, lon)	cloud top temperature uncertainty
ctp_unc(time, lat, lon)	cloud top pressure uncertainty
cth_unc(time, lat, lon)	cloud top height uncertainty

8.2.1.3 Cloud type (CPH)

cph(time, lat, lon)	cloud phase (clear, liquid, ice)
cph_extended(time,lat, lon)	cloud phase which is separated between clear, fog, water, super-cooled, opaque, cirrus, and overlap

8.2.1.4 Cloud physical properties (CTO)

cwp(time, lat, lon)	cloud water path
cwp_unc(time, lat, lon)	retrieval error in cloud water path
cot(time, lat, lon)	cloud optical thickness
cot_unc(time, lat, lon)	retrieval error in cloud optical thickness
cre(time, lat, lon)	effective droplet radius / effective ice particle size
cre_unc(time, lat, lon)	retrieval error in effective droplet radius / effective ice particle size
cdnc(time, lat, lon)	cloud droplet number concentration
cdnc_unc(time, lat, lon)	retrieval error in cloud droplet number concentration
cgt(time, lat, lon)	geometrical thickness of a liquid water cloud
cgt_unc(time, lat, lon)	retrieval error in geometrical thickness of a liquid water cloud

8.2.2 Global attributes

Each level-2b netCDF file also possesses general attributes that are valid for all variables contained in this file, the attributes are described in Table 8.3.

Table 8-3: General attributes of a non-averaged netCDF file.

Name	Description
title	CM SAF cLoud, Albedo and RAdiation dataset, AVHRR-based, edition 3 (CLARA-A3)
summary	This file contains AVHRR-based Thematic Climate Data Records (TCDR) produced by the Satellite Application Facility on Climate Monitoring (CM SAF)
id	DOI:10.5676/EUM_SAF_CM/CLARA_AVHRR/V003
product_version	3.0
creator_name	DE/DWD
creator_email	contact.cmsaf@dwd.de
creator_url	http://www.cmsaf.eu/
institution	EUMETSAT/CMSAF
project	Satellite Application Facility on Climate Monitoring (CM SAF)
references	https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003
platform	Satellite name of the platform where AVHRR operates on
instrument	AVHRR>Advanced Very High Resolution Radiometer
time_coverage_start	Temporal coverage start of the data [ISO8601 date]
time_coverage_end	Temporal coverage end of the data [ISO8601 date]
time_coverage_duration	P1D
Time_coverage_resolution	P1D
geospatial_lat_min	-90
geospatial_lat_max	90
geospatial_lat_units	degrees_north
geospatial_lon_min	-180
geospatial_lon_max	180
geospatial_lon_units	degrees_east

Name	Description
geospatial_lat_resolution	0.05 degree
geospatial_lon_resolution	0.05 degree
keywords	EARTH SCIENCE>ATMOSPHERE>CLOUDS>"variable group"
Conventions	convention tables for metadata and attributes (CF-1.7, ACDD-1.3)
keywords_vocabulary	Vocabulary for keywords in the global attributes (GCMD Science Keywords, Version 9.1.5)
standard_name_vocabulary	Vocabulary for standard names in the parameter attributes (Standard Name Table (v70, 10 December 2019))
instrument_vocabulary	Vocabulary for the instrument in the global attributes (GCMD Instruments, Version 9.1.5)
platform_vocabulary	Vocabulary for the platform in the global attributes (GCMD Platforms, Version 9.1.5)
date_created	Point in time, when the file was created [ISO8601 date]
CMSAF_platform_and_orbit	Satellite name and number of orbits, that are included in this file
CMSAF_L1_processor	PyGAC, level1c4pps
CMSAF_L2_processor	PPSv2018-patch5
CMSAF_L3_processor	CMSAFGACL3_V3.0
variable_id	names of the main variables in the product
license	The CM SAF data are owned by EUMETSAT and are available to all users free of charge and with no conditions to use. If you wish to use these products, EUMETSAT\'s copyright credit must be shown by displaying the words \"Copyright (c) (2022) EUMETSAT\" under/in each of these SAF Products used in a project or shown in a publication or website. Please follow the citation guidelines given at https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003 and also register as a user at http://cm-saf.eumetsat.int/ to receive latest information on CM SAF services and to get access to the CM SAF User Help Desk.

8.3 Data format description of monthly and daily mean products

The averaged CLARA-A3 product are provided as daily and monthly means. The daily means are based on the level-2b data files, averaged in time over the ascending and descending node and in space from the 0.05° grid to the 0.25° grid. The monthly mean is a temporal average of all daily means. The product files contain general variables, which are common for all files, and product specific variables. The latter are three-dimensional, except for the histograms. Dimension of all three-dimensional fields are *time, lon*, *lat*. For the JCHs, additional two dimensions for COT and CTP bins are included. All other provided histograms have one additional dimension for variable bin.

All level-3 products contain the same set of common variables provided in Table 8.1. Data fields included in level-3 products are listed and described in section 8.3.1. Each of them also contains specific attributes as given in Table 8-2. Global attributes of each file are reported in section 8.3.2. Additionally to the global products, polar data is provided on an equal-area grid with 25km resolution as daily and monthly means, which are described in section 8.3.3.

8.3.1 Product specific data fields

8.3.1.1 Fractional cloud coverage (CFC)

nobs(time, lon, lat)	number of observations used to create mean CFC
nobs_cloud_day(time, lon, lat)	number of cloudy classified observations used to create mean daytime CFC
nobs_cloud_night(time, lon, lat)	number of cloudy classified observations used to create mean nighttime CFC
cfc(time, lon, lat)	mean CFC value
cfc_std(time, lon, lat)	standard deviation over all CFC data points
cfc_unc_mean(lon, lat, time)	mean of CFC uncertainty
cfc_day(time, lon, lat)	mean daytime CFC value
cfc_night(time, lon, lat)	mean nighttime CFC value
cfc_low(time, lon, lat)	mean CFC of all clouds below 680 hPa
cfc_middle(time, lon, lat)	mean CFC of all clouds between 440hPa and 680 hPa
cfc_high(time, lon, lat)	mean CFC of all clouds above 440 hPa
cma_prob(lon, lat, time)	mean probabilistic CFC value
cma_prob_day(lon, lat, time)	mean daytime probabilistic CFC value
cma_prob_night(lon, lat, time)	mean nighttime probabilistic CFC value

8.3.1.2 Cloud phase (CPH)

nobs(time, lon, lat)	number of observations used to create mean CPH
nobs_cloud_liq*(time, lon, lat)*	number of observations classified as liquid cloud
nobs_cloud_ice(time, lon, lat)	number of observations classified as ice cloud
nobs_cloud_liq_day(time, lon, lat)	number of observations classified as liquid cloud at day
nobs_cloud_ice_day(time, lon, lat)	number of observations classified as ice cloud at night
nobs_cloud_liq_night(time, lon, lat)	number of observations classified as liquid cloud at day
nobs_cloud_ice_night(time, lon, lat)	number of observations classified as ice cloud at night
cph(time, lon, lat)	fraction of liquid water clouds
cph_std(time, lon, lat)	standard deviation over all CPH data points
cph_day(time, lon, lat)	mean daytime liquid cloud fraction
cph_day_std(time, lon, lat)	standard deviation over all daytime CPH data points
cph_night(time, lon, lat)	mean nighttime liquid cloud fraction
cph_night_std(time, lon, lat)	standard deviation over all nighttime CPH data points

8.3.1.3 Cloud top level (CTO)

nobs(time, lon, lat)	number of valid observations available for CTO
nobs_cloud(time, lon, lat)	number of cloudy classified observations used to create mean cloud top products
nobs_day_liq(time, lon, lat)	number of observations classified as liquid cloud used to create mean daytime cloud top products
nobs_day_ice(time, lon, lat)	number of observations classified as ice cloud used to create mean daytime cloud top products
nobs_night_liq(time, lon, lat)	number of observations classified as liquid cloud used to create mean night-time cloud top products
nobs_night_ice(time, lon, lat)	number of observations classified as ice cloud used to create mean night-time cloud top products
ctt(time, lon, lat)	arithmetical mean cloud top temperature (CTT)
cth(time, lon, lat)	arithmetical mean cloud top height (CTH)
ctp(time, lon, lat)	arithmetical mean cloud top pressure (CTP)

ctp_log(time, lon, lat)	logarithmic mean CTP
*ctX_std(time, lon, lat)**	standard deviation over CTX* data points
ctX_liq_day(time, lon, lat)	arithmetical mean CTX for liquid clouds at day
ctX_liq_night(time, lon, lat)	arithmetical mean CTX for liquid at night-time
ctX ice_day(time, lon, lat)	arithmetical mean CTX for ice clouds at day
ctX_ice_night(time, lon, lat)	arithmetical mean CTX for ice clouds at night-time
ctX_unc_mean(time, lon, lat)	CTX uncertainty mean
ctX_unc_std(time, lon, lat)	CTX uncertainty standard deviation
ctX_unc2_mean(time, lon, lat)	CTX squared uncertainty mean
ctX_unc01 (time, lon, lat)	CTX uncertainty assuming correlation of 0.1
ctX_unc10(time, lon, lat)	CTX uncertainty assuming correlation of 1.0

**CTX (CTT or CTH or CTP); additional fields containing standard deviation and uncertainties are the same for CTT, CTH and CTP*

8.3.1.4 Liquid water path (LWP)

nobs(time, lon, lat)	number of valid observations available for LWP
nobs_cloud_liq_cot(time, lon, lat)	number of observations classified as liquid clouds for LWP and COT averaging
nobs_cloud_liq_cre(time, lon, lat)	number of observations classified as liquid clouds for CRE, CDNC and CGT averaging
SZA(time, lon, lat)	mean solar zenith angle of successful retrieval results and liquid phase results
SZA_std(time, lon, lat)	standard deviation of the solar zenith angle of successful retrieval results and liquid phase results
lwp(time, lon, lat)	mean liquid water path (LWP)
lwp_allsky(time, lon, lat)	grid box mean LWP, weighted by the grid box cloud fraction
cot_liq(time, lon, lat)	athimetic mean liquid cloud optical thickness (COT)
cot_liq_log(time, lon, lat)	logarithmic mean liquid COT
cre_liq(time, lon, lat)	mean effective radius (CRE) of water droplets
cdnc_liq(time, lon, lat)	mean cloud droplet number concentration of liquid clouds

cgt_liq(time, lon, lat)	mean geometrical thickness of liquid clouds
*lwp_std(time, lon, lat)**	standard deviation over LWP* data points
lwp_unc_mean(time, lon, lat)	LWP uncertainty mean
lwp_unc_std(time, lon, lat)	LWP uncertainty standard deviation
lwp_unc2_mean(time, lon, lat)	LWP squared uncertainty mean
lwp_unc01 (time, lon, lat)	LWP uncertainty assuming correlation of 0.1
lwp_unc10(time, lon, lat)	LWP uncertainty assuming correlation of 1.0

**additional fields containing standard deviation and uncertainties are the same for lwp, cot_liq, cre_liq, cdnc_liq, cgt_liq*

8.3.1.5 Ice water path (IWP)

nobs(time, lon, lat)	number of valid observations available for IWP
nobs_cloud_ice_cot(time, lon, lat)	number of observations classified as ice clouds for IWP and COT averaging
nobs_cloud_ice_cre(time, lon, lat)	number of observations classified as ice clouds for CRE averaging
SZA(time, lon, lat)	mean solar zenith angle of successful retrieval results and ice phase results
SZA_std(time, lon, lat)	standard deviation of the solar zenith angle of successful retrieval results and ice phase results
iwp(time, lon, lat)	mean ice water path (IWP)
iwp_allsky(time, lon, lat)	grid box mean IWP, weighted by the grid box cloud fraction
cot_ice(time, lon, lat)	athimetic mean ice cloud optical thickness (COT)
cot_ice_log(time, lon, lat)	logarithmic mean ice COT
cre_ice(time, lon, lat)	mean effective radius (CRE) of ice particles
*iwp_std(time, lon, lat)**	standard deviation over IWP* data points
iwp_unc_mean(time, lon, lat)	IWP uncertainty mean
iwp_unc_std(time, lon, lat)	IWP uncertainty standard deviation
iwp_unc2_mean(time, lon, lat)	IWP squared uncertainty mean

iwp_unc01 (time, lon, lat)	IWP uncertainty assuming correlation of 0.1
iwp_unc10(time, lon, lat)	IWP uncertainty assuming correlation of 1.0

**additional fields containing standard deviation and uncertainties are the same for iwp, cot_ice, cre_ice*

8.3.1.6 Joint Cloud property Histograms (JCH)

hist_phase(hist_phase)	two-elements vector containing liquid and ice phase
hist2d_cot_bin_border(hist_cot_bin_border)	vector contains outer limits of the COT bins
hist2d_cot_bin_centre(hist_cot_bin_centre)	vector contains centre of the COT bins
hist2d_ctp_bin_border(hist_ctp_bin_border)	vector contains outer limits of the CTP bins
hist2d_ctp_bin_centre(hist_ctp_bin_centre)	vector contains centre of the CTP bins
cfc(time, lat, lon)	mean fractional cloud cover, used to the JCH
hist2d_cot_ctp(time,hist_phase, hist2d_ctp_bin_centre, hist2d_cot_bin_centre,lat, lon)	number of occurrences of specific combinations of COT and CTP ranges at given spatial location. The Joint Cloud property Histograms are defined on coarser spatial resolution (1°) compared to all other products.

8.3.1.7 One-dimensional histograms

Cloud top histograms (CTO)

hist_phase(hist_phase)	two-elements vector containing liquid and ice phase
hist1d_ctt_bin_border(hist_ctt_bin_border)	vector contains outer limits of the CTT bins
hist1d_ctt_bin_centre(hist_ctt_bin_centre)	vector contains centre of the CTT bins
hist1d_ctp_bin_border(hist_ctp_bin_border)	vector contains outer limits of the CTP bins
hist1d_ctp_bin_centre(hist_ctp_bin_centre)	vector contains centre of the CTP bins
hist1d_ctt(lon, lat, hist1d_ctt_bin_centre,hist_phase,time)	field contains the number of occurrences of specific CTT ranges at given spatial location.
hist1d_ctp(lon, lat, hist1d_ctp_bin_centre,hist_phase,time)	field contains the number of occurrences of specific CTP ranges at given spatial location.

Cloud water path histograms (CWP)

hist_phase(hist_phase)	two-elements vector containing liquid and ice phase
hist1d_cwp_bin_border(hist_cwp_bin_border)	vector contains outer limits of the cloud water path (CWP) bins
hist1d_cwp_bin_centre(hist_cwp_bin_centre)	vector contains centre of the CWP bins
hist1d_cot_bin_border(hist_cot_bin_border)	vector contains outer limits of the cloud optical thickness (COT) bins
hist1d_cot_bin_centre(hist_cot_bin_centre)	vector contains centre of the COT bins
hist1d_ref_bin_border(hist_ref_bin_border)	vector contains outer limits of the cloud particle effective radius (REF) bins
hist1d_ref_bin_centre(hist_ref_bin_centre)	vector contains centre of the REF bins
hist1d_cdnc_bin_border(hist_cdnc_bin_border)	vector contains outer limits of the cloud droplet number concentration (CDNC) bins
hist1d_cdnc_bin_centre(hist_cdnc_bin_centre)	vector contains centre of the CDNC bins
hist1d_cgt_bin_border(hist_cgt_bin_border)	vector contains outer limits of the cloud geometrical thickness (CGT) bins
hist1d_cgt_bin_centre(hist_cgt_bin_centre)	vector contains centre of the CGT bins
hist1d_cwp(lon, lat, hist1d_cwp_bin_centre,hist_phase,time)	field contains the number of occurrences of specific CWP ranges at given spatial location
hist1d_cot(lon, lat, hist1d_cot_bin_centre,hist_phase,time)	field contains the number of occurrences of specific COT ranges at given spatial location
hist1d_cdnc(lon, lat, hist1d_cdnc_bin_centre,hist_phase,time)	field contains the number of occurrences of specific CDNC ranges at given spatial location (for liquid clouds only)
hist1d_cgt (lon, lat, hist1d_cgt_bin_centre,hist_phase,time)	field contains the number of occurrences of specific CGT ranges at given spatial location (for liquid clouds only)

8.3.2 Global attributes

Table 8-4 provides the global attributes of averaged CLARA-A3 level-3 final product files. Possible values of the attributes are also given as well as explanations.

Table 8-4: Overview of global attributes of NetCDF files of CLARA-3 products and possible corresponding values.

Name	**Description**
title	CM SAF cLoud, Albedo and RAdiation data record, AVHRR-based, edition 3 (CLARA-A3)
summary	This file contains AVHRR-based Thematic Climate Data Records (TCDR) produced by the Satellite Application Facility on Climate Monitoring (CM SAF)
id	DOI:10.5676/EUM_SAF_CM/CLARA_AVHRR/V003
product_version	3.0
creator_name	DE/DWD
creator_email	contact.cmsaf@dwd.de
creator_url	http://www.cmsaf.eu/
institution	EUMETSAT/CMSAF
project	Satellite Application Facility on Climate Monitoring (CM SAF)
references	http://dx.doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003
keywords_vocabulary	Vocabulary for keywords in the global attributes (GCMD Science Keywords, Version 9.1.5)
keywords	EARTH SCIENCE>ATMOSPHERE>CLOUDS>"variable group"
Conventions	convention tables for metadata and attributes (CF-1.7, ACDD-1.3)
standard_name_vocabulary	Vocabulary for standard names in the parameter attributes (Standard Name Table (v70, 10 December 2019))
date_created	Point in time, when the file was created [ISO8601 date]
geospatial_lat_max	90
geospatial_lat_min	-90
geospatial_lat_units	degrees_north

Name	**Description**
geospatial_lon_max	180
geospatial_lon_min	-180
geospatial_lon_units	degrees_east
Geospational_lat_resolution	0.25 degrees
Geospational_lon_resolution	0.25 degrees
time_coverage_start	Temporal coverage start of the data [ISO8601 date]
time_coverage_end	Temporal coverage end of the data [ISO8601 date]
time_coverage_duration	P1M or P1D (period 1 month, day)
time_coverage_resolution	P1M or P1D (period 1 month, day)
platform	Satellite name of the platform where AVHRR operates on
platform_vocabulary	Vocabulary for platform in the global attributes (GCMD Platforms, Version 9.1.5)
instrument_vocabulary	Vocabulary for instrument in the global attributes (GCMD Instruments, Version 9.1.5)
instrument	AVHRR>Advanced Very High Resolution Radiometer
CMSAF_platform_and_orbits	For daily means, the attribute counts the number of included orbits, that are used to build the daily mean
CMSAF_included_Daily_Means	For monthly means, this attribute counts the number of daily means, that are used to build the monthly mean
CMSAF_L1_processor	PyGAC, level1c4pps
CMSAF_L2_processor	PPSv2018-patch5
CMSAF_L3_processor	CMSAFGACL3_V3.0
variable_id	names of the main variables in the product
license	The CM SAF data are owned by EUMETSAT and are available to all users free of charge and with no conditions to use. If you

Name	**Description**
	wish to use these products, EUMETSAT\'s copyright credit must be shown by displaying the words \"Copyright (c) (2022) EUMETSAT\" under/in each of these SAF Products used in a project or shown in a publication or website. Please follow the citation guidelines given at https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003 and also register as a user at http://cm-saf.eumetsat.int/ to receive latest information on CM SAF services and to get access to the CM SAF User Help Desk.

8.3.3 Polar data

Additionally to the global data, CLARA-A3 also provides the CFC and CTO L3 products on two polar grids with a spatial resolution of 25 km. For each daily or monthly mean one file for each polar area, the Arctic and the Antarctic. These data are provided on an equal area ease grid. Therefore, the latitude and longitude variable are changed from vectors to arrays. They show the latitude and longitude for each grid box. Otherwise, the data structure is identical to the global data files described above.

8.4 ICDR specific adaptions

CLARA-A3 ICDR products contain the same set of variables and have the same format.

9 Data ordering via the Web User Interface (WUI)

User services are provided through the CM SAF homepage www.cmsaf.eu. The user service includes information and documentation about the CM SAF and the CM SAF products, information on how to contact the user help desk and allows to search the product catalogue and to order products.

On the main webpage, a detailed description how to use the web interface for product search and ordering is given. We refer the user to this description since it is the central and most up to date documentation. However, some of the key features and services are briefly described in the following sections.

9.1 Product ordering process

You need to be registered and logged in to order products. A login is provided upon registration, all products are delivered free of charge. After the selection of the product, the desired way of data transfer can be chosen. This is either via a temporary ftp account (the default setting), or by CD/DVD or email. Each order will be confirmed via email, and the user will get another email once the data have been prepared. If the ftp data transfer was selected, this second email will provide the information on how to access the ftp server.

9.2 Contact User Help Desk staff

In case of questions the contact information of the User Help Desk (e-mail address contact.cmsaf@dwd.de, telephone and fax number) are available via the CM SAF main webpage (http://www.cmsaf.eu) or the main page of the Web User Interface.

9.3 Feedback/User Problem Report

Users of CM SAF products and services are encouraged to provide feedback on the CM SAF product and services to the CM SAF team. Users can either contact the User Help Desk (see chapter 9.2) or use the “User Problem Report” page. A link to the “User Problem Report” is available either from the CM SAF main page (www.cmsaf.eu) or the Web User Interface main page.

9.4 Service Messages / log of changes

Service messages and a log of changes are also accessible from the CM SAF main webpage (www.cmsaf.eu) and provide useful information on product status, versioning and known deficiencies.

10 Copyright and Disclaimer

The user of CM SAF data agrees to respect the following regulations:

Copyright

All intellectual property rights of the CM SAF products belong to EUMETSAT. The use of these products is granted to every interested user, free of charge. If you wish to use these products in publications, presentations, web pages etc., EUMETSAT's copyright credit must be shown by displaying the words "copyright (year) EUMETSAT" on each of the products used.

Acknowledgement and Identification

When exploiting EUMETSAT/CM SAF data you are kindly requested to acknowledge this contribution accordingly and refer to the CM SAF, e.g. by stating "The work performed was done (i.a.) by using data from EUMETSAT's Satellite Application Facility on Climate Monitoring (CM SAF)". It is highly recommended to clearly identify the product version used. An effective way to do this is the citation of CM SAF data records via the digital object identifier (doi). The doi of the data record can be retrieved through (http://www.cmsaf.eu/DOI).

Re-distribution of CM SAF data

Please do not re-distribute CM SAF data to 3rd parties. The use of the CM SAF products is granted free of charge to every interested user, but we have an essential interest to know how many and what users the CM SAF has. This helps to ensure of the CM SAF operational services as well as its evolution according to user needs and requirements. Each new user shall register at CM SAF in order to retrieve the data.

Feedback

We are keen to learn of what use the CM SAF data are. So please feedback your experiences and your application area of the CM SAF data. EUMETSAT CM SAF is user driven service and is committed to consider the needs and requirements of its users in the planning for product improvements and additions. Users are invited to provide their specific requirements on future products for their applications.

11 References

Baum, B. A., P. Yang, A. J. Heymsfield, C. G. Schmitt, Y. Xie, A. Bansemer, Y.-X. Hu, and Z. Zhang, 2011: Improvements in shortwave bulk scattering and absorption models for the remote sensing of ice clouds.

Benas, N., Finkensieper, S., Stengel, M., van Zadelhoff, G.-J., Hanschmann, T., Hollmann, R., and Meirink, J. F., 2017: The MSG-SEVIRI-based cloud property data record CLAAS-2, Earth System Science Data, 9, 415–434, doi:10.5194/essd-9-415-2017.

Benas, N., Meirink, J. F., Karlsson, K.-G., Stengel, M., and Stammes, P.: Satellite observations of aerosols and clouds over southern China from 2006 to 2015: analysis of changes and possible interaction mechanisms, Atmospheric Chemistry and Physics, 20, 457–474, doi:10.5194/acp-20-457-2020, 2020.

Berk, A., G. P. Anderson, P. K. Acharya, J. H. Chetwynd, L.S. Bernstein, E. P. Shettle, M. W. Matthew, and S. M. Adler-Golden, 2000: MODTRAN4 Version 2 Users Manual. Technical report, Air Force Materiel Command, Air Force Research Laboratory, Space Vehicles Directorate, Hanscom AFB, MA 01731, USA.

Dee et al. (35 named authors), 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quart. J. Royal Met. Soc., 137 (656), 553 - 597. Online version: http://onlinelibrary.wiley.com/doi/10.1002/qj.828/abstract

De Haan, J. F., P. Bosma, and J. W. Hovenier, 1987: The adding method for multiple scattering calculations of polarized light, *Astron. Astrophys.*, **183**, 371-391.

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

Dybbroe, A., A. Thoss and K.-G. Karlsson, 2005: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling - Part I: Algorithm description, *J. Appl. Meteor*, **44**, 39-54

Haywood, J.M., S.R. Osborne, S.J. Abel, 2004: The effect of overlying absorbing aerosol layers on remote sensing retrievals of cloud effective radius and cloud optical depth, *Quart. J. Roy. Meteorol. Soc.*, **130**, 779-800, doi: 10.1256/qj.03.100.

Heidinger, A.K., W.C. Straka, C.C. Molling, J.T. Sullivan and X.Q. Wu, 2010: Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record. *Int. J. Rem. Sens.,***31**(24), 6493-6517.

Heidinger, A. K., M. J Foster, A. Walther and X. Zhao, 2013: The Pathfinder Atmospheres Extended (PATMOS-x) AVHRR Climate Data Set., Bull. Amer. Meteor. Soc, doi: http://dx.doi.org/10.1175/BAMS-D-12-00246.1.

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

Ignatov, I.L., E.D. Harrod, K.B. Kidwell and G.P. Goodrum, 2004: Equator crossing times for NOAA, ERS and EOS sun-synchronous Satellites, *Int. J. Rem. Sens.,***25** (23), 5255–5266, DOI: 10.1080/0143116041000171298

Karlsson, K.-G.and Dybbroe, A., 2009. Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations. Atmos. Chem. Phys., 10, 1789-1807, 2010 www.atmos-chem-phys.net/10/1789/2010/
doi:10.5194/acp-10-1789-2010

Knap, W. H., L. C. Labonnote, G. Brogniez, and P. Stammes, 2005: Modeling total and polarized reflectances of ice clouds: evaluation by means of POLDER and ATSR-2 measurements. *Appl. Optics*, **44**, 4060-4073.

Meirink, J.F., R.A. Roebeling and P. Stammes, 2009: Atmospheric correction for the KNMI Cloud Physical Properties retrieval algorithm, KNMI publication: TR-304, 17/2/2009, pp22.

Minnis, P., K. N. Liou, and Y. Takano, 1993: Inference of Cirrus Cloud Properties Using Satellite-observed Visible and Infrared Radiances. Part I: Parameterization of Radiance Fields. *J. Atmos. Sci.*, **50**, 1279–1304.

Mittaz, P.D. and R. Harris, 2009: A Physical Method for the Calibration of the AVHRR/3 Thermal IR Channels 1: The Prelaunch Calibration Data. *J. Atmos. Ocean. Tech.,* **26,** 996-1019, doi: 10.1175/2008JTECHO636.1.

Nakajima, T., and M. D. King, 1990: Determination of the Optical Thickness and Effective Particle Radius of Clouds from Reflected Solar Radiation Measurements. Part 1: Theory. *J. Atmos. Sci .*, **47**, 1878-1893.

Pavolonis, M. J. and A. K. Heidinger, 2004: Daytime cloud overlap detection from AVHRR and VIIRS, *J. Appl. Meteorol*., **43**, 762-778.

Pavolonis, M. J., A. K. Heidinger, and T. Uttal, 2005: Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparison, *J. Appl. Meteorol*., **44**, 804-826.

Platnick, S., 2001: A superposition technique for deriving mean photon scattering statistics in plane-parallel cloudy atmospheres, *J. Quant. Spectrosc. Radiat. Transfer*, 68, 57-73.

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C., Frey, R. A., 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Trans. Geosci. Remote Sens*., **41**, 459-473.

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

Rossow, W.B., and R.A. Schiffer, 1999: Advances in understanding clouds from ISCCP. B. Am. Meteorol. Soc., 80, 2261-2287.

Stammes, P., 2001: Spectral radiance modeling in the UV-Visible range. *IRS 2000: Current problems in Atmospheric Radiation, edited by W.L. Smith and Y.M. Timofeyev*, pp 385-388, A. Deepak Publ., Hampton, VA.

Stephens, G.., 1978: Radiation profiles in extended water clouds. II: Parameterization schemes. *J. Atmos. Sci.*, **35**, 2123-2132.

Yang, P., L. Bi, B. A. Baum, K.-N. Liou, G. W. Kattawar, M. I. Mishchenko, and B. Cole, 2013: Spectrally consistent scattering, absorption, and polarization properties of atmospheric ice crystals at wavelengths from 0.2 to 100 μm, J. Atmos. Sci., 70, 330-347, doi:10.1175/JAS-D-12-039.1.

12 Acronyms

AMSR-E	Advanced Microwave Scanning Radiometer for EOS
ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
AVPOS	AVHRR Polar Satellites (L3 products based on all available AVHRR-carrying satellites)
BC-RMS	Bias-Corrected RMS
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarisation
CDNC	Cloud Droplet Number Concentration
CDO	Climate Data Operators
CDOP	Continuous Development and Operations Phase
CF	Climate and Forecast Metadata Convention ((http://cf-pcmdi.llnl.gov/))
CFC	Fractional Cloud Cover
CGT	Cloud Geometrical Thickness
CFOT	Cloud Feature Optical Depth
CLARA-A	CM SAF cLoud, Albedo and Radiation products, AVHRR-based
CLAAS	CM SAF cLoud dAtAset using SEVIRI
CM SAF	Satellite Application Facility on Climate Monitoring
COT	Cloud Optical Thickness
CPH	Cloud Phase
CPR	Cloud Profiling Radar
CRE	Cloud Particle Effective Radius
CTH	Cloud Top Height
CTO	Cloud Top product
CTP	Cloud Top Pressure
CTT	Cloud Top Temperature
CPP	Cloud Physical Properties

DAK	Doubling Adding KNMI (radiative transfer model)
DRR	Delivery Readiness Review
DWD	Deutscher Wetterdienst (German MetService)
ECMWF	European Centre for Medium Range Forecast
ECV	Essential Climate Variable
ERA5	ECMWF Re-Analysis dataset 5
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	Fundamental Climate Data Record
FCI	Flexible Combined Imager
GAC	Global Area Coverage (AVHRR)
GCOS	Global Climate Observing System
GSICS	Global Space-Based Inter-Calibration System
HIRS	High resolution Infrared Radiation Sounder
IR	InfraRed
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Inter Tropical Convergence Zone
IWP	Ice Water Path
JCH	Joint Cloud properties Histogram
KNMI	Koninklijk Nederlands Meteorologisch Instituut
KSS	Hanssen-Kuipers Skill Score
LWP	Liquid Water Path
MAC-LWP	Multisensor Advanced Climatology of LWP
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
NIR	Near InfraRed
NOAA	National Oceanic & Atmospheric Administration
NWC SAF	SAF on Nowcasting and Very Short Range Forecasting
NWP	Numerical Weather Prediction

PATMOS-x	Pathfinder Atmospheres-Extended dataset (NOAA)
PPS	Polar Platform System (NWC SAF polar cloud software package)
PRD	Product Requirement Document
PUM	Product User Manual
RMS	Root Mean Square (Error)
RTTOV	Radiative Transfer model for TOVS
SEVIRI	Spinning Enhanced Visible and InfraRed Imager
SAF	Satellite Application Facility
SMHI	Swedish Meteorological and Hydrological Institute
SYNOP	Synoptic observations
SZA	Solar Zenith Angle
VIS	Visible
VZA	Viewing Zenith Angle

Annex A – Impact of imperfect sampling of the clouds' diurnal cycle and orbital drift

For analysing the potential systematic errors introduced by reduced, and thus potentially imperfect diurnal sampling of AVHRR-carrying polar-orbiting satellites, temporally highly resolved, geostationary observations of the CLAAS-3 dataset, https://doi.org/10.5676/EUM_SAF_CM/CLAAS/V003) have been utilized. It is worth noting that this analysis benefits from the fact that basically the same CM SAF cloud retrieval schemes have been used for CLAAS-3 and CLARA-A3, which reduced the risk for falsely attributed results in this study.

The reduced temporal sampling was emulated by temporally sub-sampling the CLAAS-3 data (at specific local solar times) and comparison to the fully sampled CLAAS-3 data. Figure A-1 shows cloud fraction maps for reduced and full sampling and their difference. Although systematic deviations might exist locally, on average over a larger domain these systematic deviations seem to average out. Figures A-2 (morning orbits) and A-3 (afternoon orbits) show frequency distributions over the mean differences of all pixels in the domain indicated in Figure A-1 for all main cloud variables for varying sampling scenarios (observations times). Relatively speaking, the reduced temporal sampling, twice per day for CFC, CTP and CPH and once per day for COT, LWP and IWP, does not introduce systematic uncertainties when considered over a larger domain. Furthermore, the small biases remain small with varying observations times. A slight exception is CTP, for which an increasing positive bias towards lunchtime and decreasing bias thereafter is found.

In the second part of this study, time series of global mean cloud properties were analysed per satellite and examples are plotted in Figures A-4 to A-6. The key results can be summarized as follows:

- Seven satellites are exposed to an orbital drift of more than 2 hours: NOAA-11, NOAA-12, NOAA-14, NOAA-15, NOAA-16, NOAA-18, NOAA-19
- Small effects are seen on global means for CFC and CTH
- If isolating daytime results for CFC an impact is seen when the observation time of individual sensors enter or leave twilight conditions: Increasing CFC values with maximum values in NH autumn/winter. Increased values appear to be explained by a combination of diurnal changes in cloudiness (best seen in NH) and cloud misclassifications in twilight conditions
- For CTP, almost no impact is seen (very weak trends, if any)
- For cloud optical and microphysical properties time series look qualitatively similar to daytime CFC in terms, thus time series are relative stable but variability is increased when observations are close to twilight conditions.

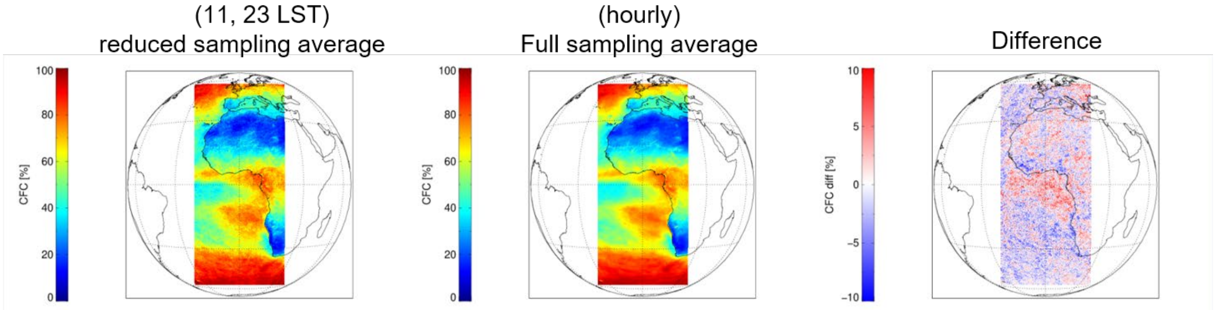

Figure A-1 Annual mean cloud fraction for reduced temporal sampling at 11AM and 11PM (left) compared to a mean inferred from full temporal sampling (middle) and the difference (right).

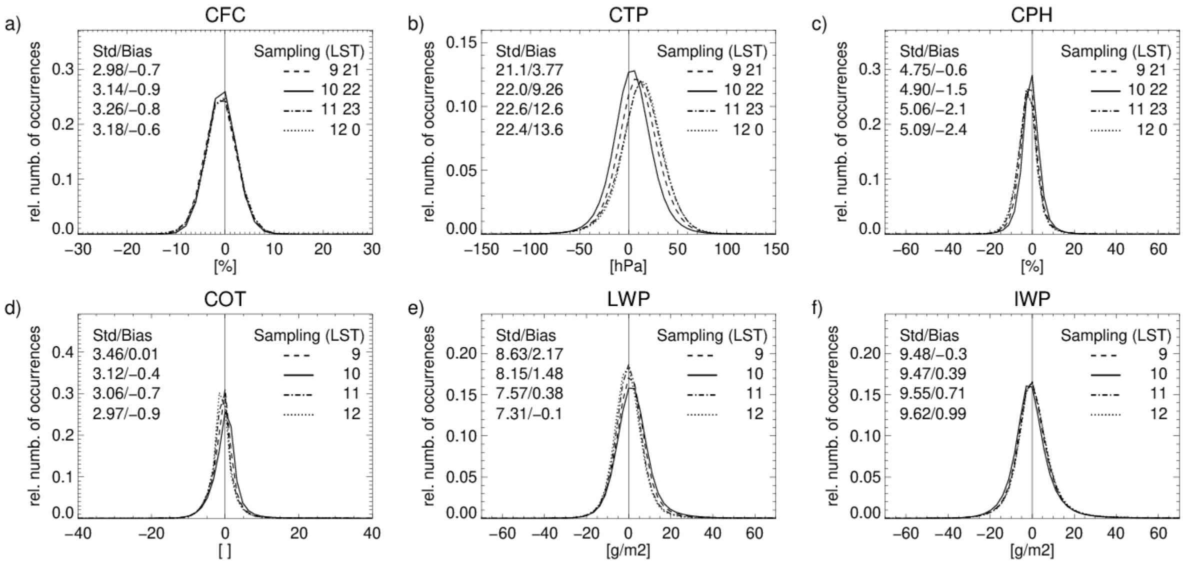

Figure A-2 Frequency distribution over annual mean differences over all pixels in the domain indicated in Figure A-1 depending on sampling time (morning orbits).

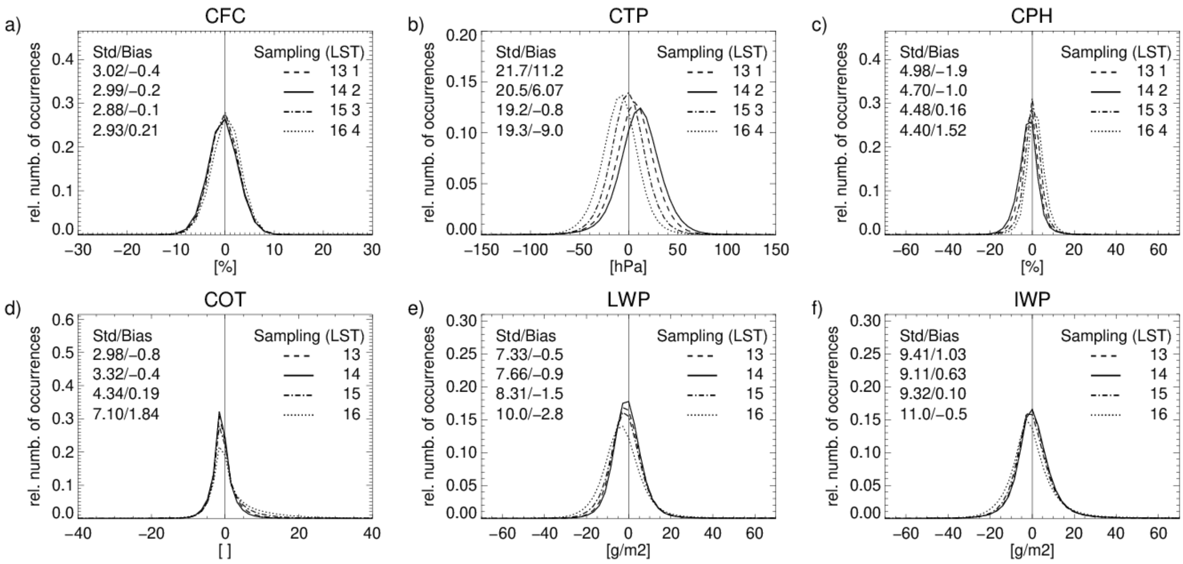

Figure A-3 As Figure A-2 but for afternoon orbits.

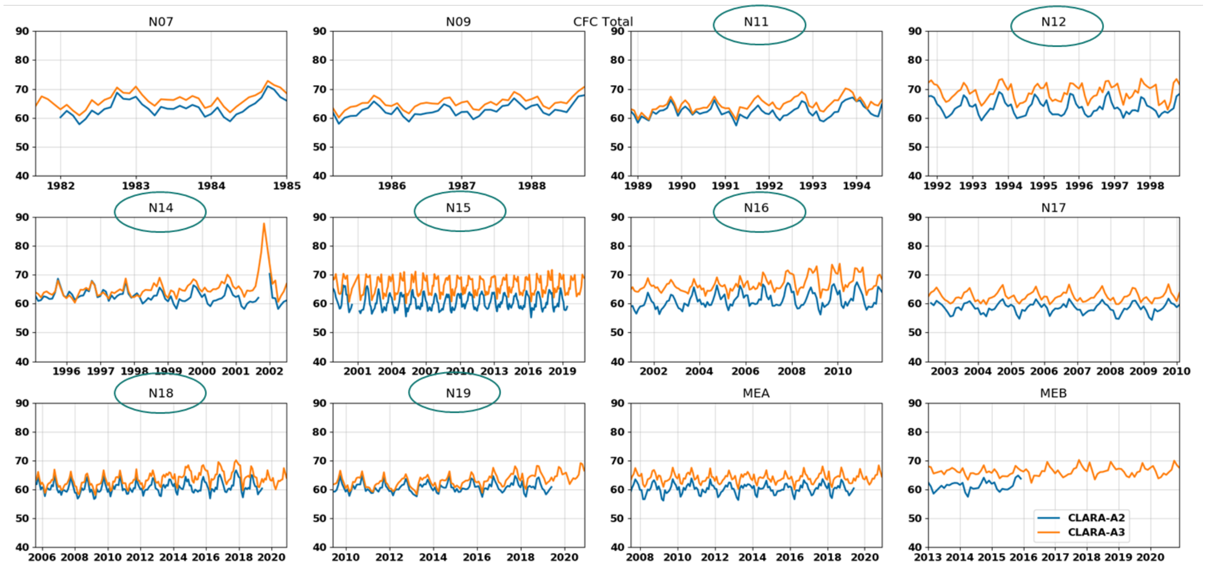

Figure A-4 Time series of global mean total cloud fraction shown individually for a selected set of satellites.

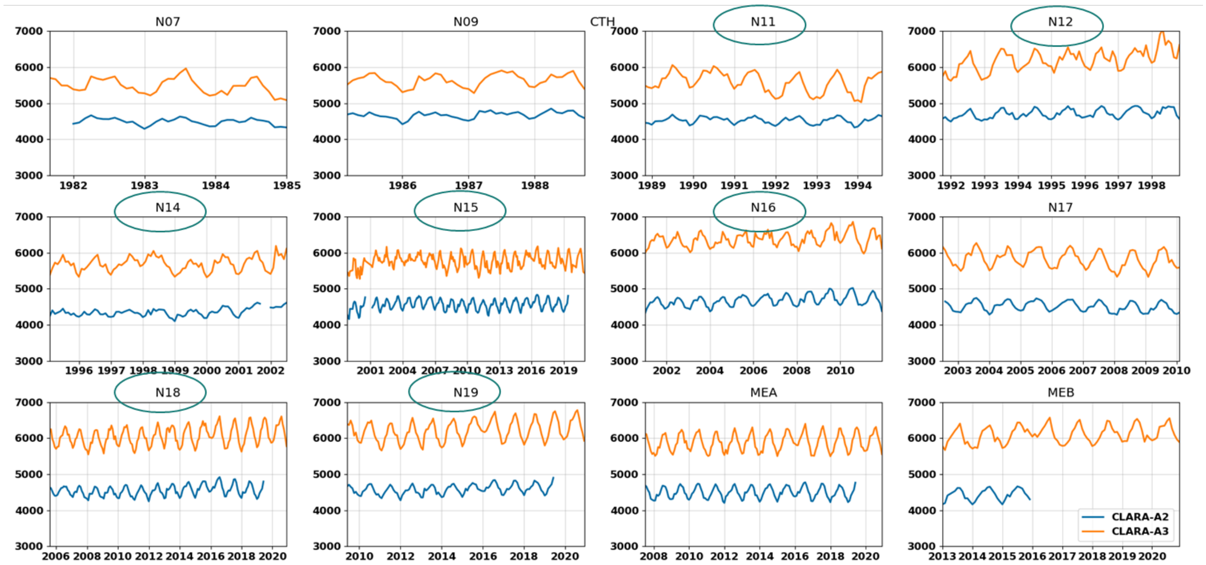

Figure A-5 As Figure A-4 but for global mean cloud top height.

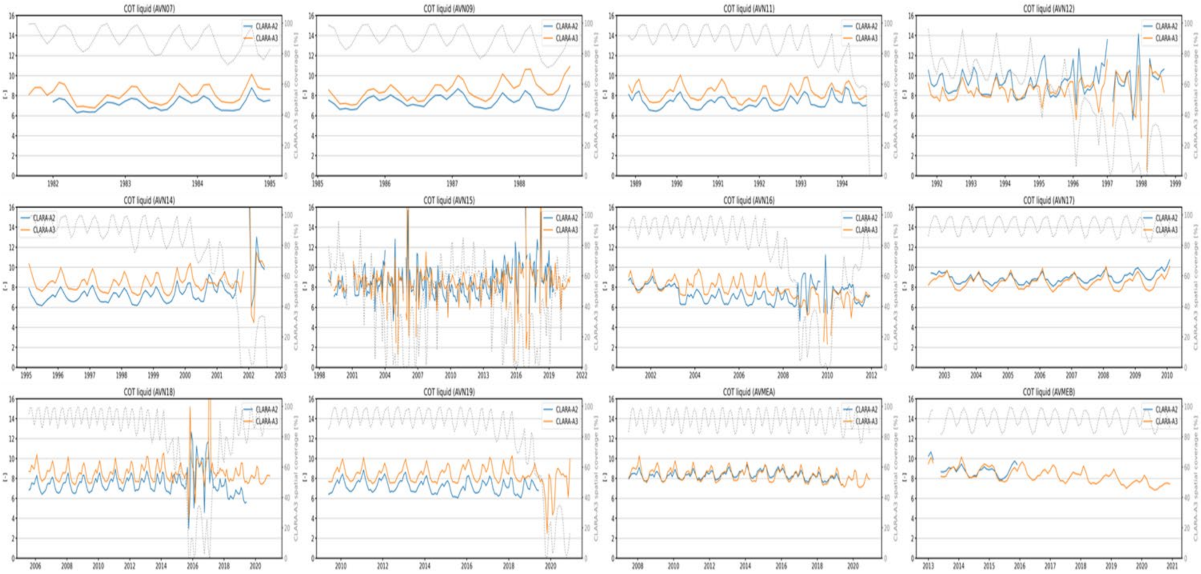

Figure A-6 As Figure A-4 but for cloud optical thickness for liquid clouds.