



Royal Netherlands
Meteorological
Institute



ESA Cloud_cci Product Validation and Intercomparison Report (PVIR)

(Applicable to Cloud_cci version 3.0 products)



Issue 6 Revision 0

03 February 2020

Deliverable No.: D-4.1


ESRIN/Contract No.: 4000109870/13/I-NB

Project Coordinator: Dr. Rainer Hollmann
Deutscher Wetterdienst
rainer.hollmann@dwd.de

Technical Officer: Dr. Simon Pinnock
European Space Agency
Simon.Pinnock@esa.int

DOIs of Cloud_cci datasets: DOI: [10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003)
DOI: [10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V003)
DOI: [10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V003)



	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 2

Document Change Record

Document, Version	Date	Changes	Originator
PVIR v6.0 submitted version	05/12/2019	Major document update now providing information about version 3 Cloud_cci datasets: AVHRR-AMv3, AVHRR-PMv3 and ATSR2-AATSRv3	Martin Stengel, Benjamin Würzler
PVIR v6.1 revised version	07.01.2019	Revision after ESA review: <ul style="list-style-type: none"> Copy/paste of Introduction section from approved User Guide v5 document Minor/editorial modifications Adding new Section (Sec. 6) presenting uncertainty validation 	Martin Stengel
PVIR v6.1	03.02.2020	Revised and issued version	Martin Stengel

Purpose

This document is the Product Validation and Intercomparison Report (PVIR) for ESA Cloud_cci. It covers the validation results for all generated version 3 data sets, namely Cloud_cci AVHRR-PMv3, AVHRR-AMv3, and ATSR2-AATSRv3, together with outcomes from a comprehensive inter-comparison exercise.




	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 3

Table of Contents

1.	Executive summary	5
2.	Introduction.....	7
2.1	The ESA Cloud_cci project	7
2.2	The Cloud_cci version 3 datasets	8
2.3	Cloud_cci cloud products	10
2.4	Validation strategy in this report	13
2.4.1	Evaluation measures	13
2.4.2	Datasets used for evaluation	14
3.	Validation of Cloud_cci products.....	16
3.1	Level-2/Level-3U cloud products	16
3.1.1	Cloud mask	16
3.1.2	Cloud phase	17
3.1.3	Cloud top height	17
3.1.4	Liquid water path.....	18
3.1.5	Ice water path	19
3.2	Level-3C cloud fraction (CFC) against SYNOP	20
3.3	Level-3C radiation products.....	24
3.3.1	TOA fluxes.....	24
3.3.2	BOA fluxes.....	26
4.	Intercomparison of Cloud_cci cloud products with other satellite-based datasets	29
4.1	Morning Satellites.....	29
4.1.1	Cloud Fraction.....	30
4.1.2	Cloud Top Pressure.....	34
4.1.3	Cloud Optical Thickness - Liquid clouds	38
4.1.4	Cloud Optical Thickness - Ice clouds	42
4.1.5	Cloud Effective Radius - Liquid clouds	46
4.1.6	Cloud Effective Radius - Ice clouds	50
4.1.7	Liquid Water Path	54
4.1.8	Ice Water Path.....	58
4.1.9	Liquid Cloud Fraction	62
4.2	Afternoon satellites	66
4.2.1	Cloud Fraction.....	67
4.2.2	Cloud Top Pressure.....	70
4.2.3	Cloud Optical Thickness - Liquid clouds	74
4.2.4	Cloud Optical Thickness - Ice clouds	78

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 4

4.2.5	Cloud Effective Radius - Liquid clouds	82
4.2.6	Cloud Effective Radius - Ice clouds	86
4.2.7	Liquid Water Path	89
4.2.8	Ice Water Path.....	93
4.2.9	Liquid Cloud Fraction	97
5.	Intercomparison of Cloud_cci radiation products with other radiation datasets	100
5.1	Top of the Atmosphere Upwelling Solar Radiation.....	101
5.2	Top of the Atmosphere Upwelling Thermal Radiation	105
5.3	Bottom of the Atmosphere Downwelling Solar Radiation.....	109
5.4	Bottom of the Atmosphere Downwelling Thermal Radiation	113
6.	Uncertainty validation	117
6.1	Cloud top pressure	117
6.1.1	Liquid clouds.....	118
6.1.2	Ice clouds	120
6.2	Liquid water path.....	122
6.3	Ice water path	125
6.4	Summary of the uncertainty validation.....	128
7.	Summary.....	129
8.	Glossary	137
9.	References	141
Annex A	Datasets for comparisons with Cloud_cci products	146
A.1	CALIPSO-CALIOP	146
A.2	AMSRE-E LWP	149
A.3	DARDAR IWP	149
A.4	SYNOP: manual cloud observations from surface stations.....	150
A.5	Baseline Surface Radiation Network.....	151
A.6	CERES EBAF Ed4.0 TOA & Surface fluxes	152
A.7	ELDONET	153
A.8	The CLARA-A2 dataset	154
A.9	PATMOS-x.....	155
A.10	MODIS Collection 6.1	156
A.11	ISCCP	157
A.12	NOAA (HIRS) Climate Data Record OLR	158
A.13	ERA-Interim reanalysis	159

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 5

1. Executive summary


In the ESA Cloud_cci project, comprehensive efforts have been undertaken to optimize the utilization of passive imaging sensors of European and non-European satellite mission for the generation of enhanced cloud property climate datasets. The Cloud_cci datasets are innovative in a variety of aspects, which are summarized in Section 2.2. In this report the main results of the evaluation process for Cloud_cci version 3 datasets are summarized. The evaluation is divided into 2 tasks:

- (1) Validation of Cloud_cci data against high quality and satellite-based reference observations, i.e. Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), passive microwave (MW) observations, combined radar+lidar (CloudSat+CALIOP) and CERES. In addition, reference observations taken at ground site were employed, i.e. SYNOP and BSRN.
- (2) Comparisons to well-established, satellite-based cloud and radiation property datasets of similar kind (partly also based on passive imager satellite sensors)


The results of Task 1 are considered to be an objective measure of the quality of the Cloud_cci data, since the reference observations used are assumed to be of higher accuracy and precision than the Cloud_cci data. Here we specifically validated pixel-based cloud mask (CMA), cloud phase (CPH) and cloud top height (CTH) against CALIOP, a space based lidar. For cloud detection, cloud phase assignment and cloud-top height we find reasonably good agreement of Cloud_cci datasets with the references, when removing the optically very thin clouds and cloud layers from the reference profiles. This is done to account for and quantify the limited sensitivity of passive imaging sensors to those thin clouds. Including thin clouds leads to underestimation in cloud detection, liquid bias in the phase detection and strong underestimation of the cloud top height for ice clouds. Validation of liquid water path (over ocean) highlights the good quality of Cloud_cci data in that respect. High correlations and small absolute and relative biases underline this. Ice water path validation reveal relatively strong negative biases in Cloud_cci datasets, although standard deviations and correlations prove that there is skill in Cloud_cci ice water path. Validations of monthly mean radiative flux properties in the Cloud_cci v3 datasets to CERES and BSRN reveal an excellent agreement with only small biases and standard deviation and very high correlations. Cloud_cci longwave fluxes are found to be in slightly better agreement to the references than the shortwave fluxes.

Results of Task 2 are considered to be a consistency analysis among all Cloud_cci datasets but also in comparisons to reference datasets, which already exist and are well established, but cannot necessarily serve as absolute references. For cloud properties we used for example CLARA-A2, PATMOS-x and MODIS Collection 6.1 datasets. The comparison includes maps of multi-annual mean values as well as zonal mean value plots both for a common period for which all datasets are available. Also time series plots of 60S-60N average are shown for the full records and discussed. For cloud fraction (CFC), CPH and cloud top pressure (CTP) we find reasonable agreement among all datasets of consideration, with usually only small difference, which can mainly be explained by different sensitivities caused by the different spectral bands used (e.g. MODIS Collection 6.1) as well as differences in algorithm approaches. However, in polar regions the spread in CFC is very large, probably due to difficulties to correctly account for snow and ice surface condition and a related limited signal to noise ratio under in these conditions, preventing any firm conclusion. For optical thickness (COT) we found a relatively good agreement among all datasets for both liquid and ice clouds. Usually, the datasets agree within 5 to 10 optical thicknesses wrt. the average values for the common period (2003-2011). Interpretations for the polar regions are not easy, because the spread is large as well as many datasets tend to have very high optical thicknesses in the polar regions which seem unlikely. For cloud effective radius (CER) we find all datasets agreeing on reflecting the same global pattern. On the other hand the absolute values differ partly by 5 to 10 μ m for liquid clouds. For ice clouds the spread is even larger with one reason being the high CER_{ice} values reported in MODIS Collection 6.1. Similar to CER, the cloud water path comparison show a reasonable agreement for all datasets for liquid clouds (LWP) but less agreement for ice clouds (IWP). The LWP and IWP agreement is better once all-sky values are considered, because here the impact of differences in cloud detection and phase determination can be reduced.

Another aspect of Task 2 was the comparisons of radiative flux properties to correspondingly, well-established datasets, such as CERES and HIRS OLR, but also including ERA-Interim. A very high degree of agreement is

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 6

found for all considered radiative flux properties highlighting the high quality of the radiative flux properties in Cloud_cci v3 datasets.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 7

2. Introduction

2.1 The ESA Cloud_cci project

The ESA Cloud_cci project covers the cloud component in the European Space Agency's (ESA) Climate Change Initiative (CCI) programme (Hollmann et al., 2013). In the ESA Cloud_cci project, long-term and coherent cloud property datasets have been generated exploiting the synergic capabilities of different Earth observation missions (European and non-European) allowing for improved accuracies and enhanced temporal and spatial sampling better than those provided by the single sources. The Cloud_cci datasets are the attempt to respond to GCOS requirements for the Cloud Properties Essential Climate Variable (ECV).

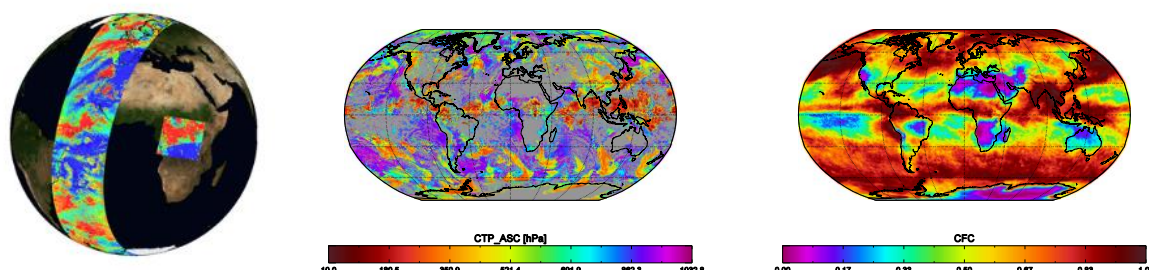


Figure 2-1 Examples of Cloud_cci cloud products. Left: Pixel-based (Level 2), middle: daily composite on a global grid (Level 3U), right: monthly averaged on a global grid (Level 3C)


To make the Cloud_cci datasets improved compared to existing ones, the following two essential steps were undertaken:

- 1) Revisit the measurement data (Level-1) and corresponding calibration performance and development of a carefully inter-calibrated and rigorously quality checked radiance data sets for AVHRR, so called Fundamental Climate Data Record (FCDR). Within this effort the calibration of AVHRR, MODIS and AATSR was compared and characterized. Please see the ATBDv5 for more information about all sensors used and their imaging characteristics. More information on the AVHRR FCDR produced and used is available in RAFCDRv1.0.
- 2) Development of two state-of-the-art physical retrieval systems that use the optimal estimation technique for a simultaneous, spectrally consistent retrieval of cloud properties including pixel-based uncertainty measures. The first retrieval framework is the Community Cloud retrieval for Climate (CC4CL; Sus et al., 2018; McGarragh et al., 2018) which is applied to AVHRR and AVHRR-heritage channels (i.e. channels which are available from all sensors) of MODIS and AATSR. The second retrieval framework is the Freie Universität Berlin AATSR MERIS Cloud retrieval (FAME-C; Carbajal Henken et al., 2014) and is applied to synergistic MERIS and AATSR measurements on-board of ENVISAT.

Based on these developments, six multi-annual, global datasets of cloud properties were generated using the passive imager satellite sensors AVHRR, MODIS, (A)ATSR and MERIS. These datasets were comprehensively evaluated (1) by using accurate reference observations of ground stations and space-based Lidar measurements and (2) by comparisons to existing and well-established global cloud property datasets. These datasets were published as **version 2** (version 1 being shorter, prototype datasets) and Digital Object Identifiers issued:

[DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V002)
[DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V002)
[DOI:10.5676/DWD/ESA_Cloud_cci/MODIS-Terra/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/MODIS-Terra/V002)
[DOI:10.5676/DWD/ESA_Cloud_cci/MODIS-Aqua/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/MODIS-Aqua/V002)
[DOI:10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V002)
[DOI:10.5676/DWD/ESA_Cloud_cci/MERIS+AATSR/V002](https://doi.org/10.5676/DWD/ESA_Cloud_cci/MERIS+AATSR/V002)

These datasets were comprehensively documented in Stengel et al. (2017), the Product Validation and Intercomparisons Report (PVIR; PVIRv5.1), the Product User Guide (PUGv4.0) and the overarching Algorithm

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 8

Theoretical Baseline Document (ATBD, [ATBDv5](#)) together with specific ATBDs for FAME-C ([ATBD-FAME-Cv5](#)) and CC4CL ([ATBD-CC4CLv5](#)).

To facilitate a suitable application of Cloud_cci datasets for model evaluation, satellite simulators have been developed, tested and published ([Eliasson et al., 2018](#) and [Stengel et al., 2018](#)).

In addition to the datasets mentioned above, AVHRR and AATSR based datasets were reprocessed again building the **version 3.0 datasets**, based on an updated CC4CL, partly covering longer periods (AVHRR) and including an extended product portfolio by including shortwave and longwave, all-sky and clear-sky radiative flux properties and top and bottom of the atmosphere (TOA and BOA).

The remaining part of this document will exclusively focus on the algorithm version used for generation of the version 3 datasets, for which the following DOI were issued:

[10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003)

[10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V003)

[10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V003](https://doi.org/10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V003).

In addition to this PUG, these v3 datasets are being documented in [ATBDv6.2](#), [PVIRv6.0](#) and [Stengel et al. \(2020\)](#).

2.2 The Cloud_cci version 3 datasets

Three Cloud_cci version 3 cloud property datasets have been generated, based on AVHRR and ATSR2+AATSR, utilizing the AVHRR-heritage channels (0.6, 0.8, 1.6/3.7, 10.8, 12.0 μm) only. The retrieval system used was an updated version of CC4CL. Since AVHRR sensors were separated into morning and afternoon orbits. Figure 2-2 shows the temporal coverage of the v3 datasets.

Table 2-1 summarizes the algorithms, sensors and satellites used for each dataset. The official versions of the datasets, as released under the issued Digital Object Identifies (DOIs, see Table 2-1), do not contain any diurnal cycle or satellite drift correction. Potential methods for such a drift correction were investigated for AVHRR and were documented in [RODCv1.0](#). In Figure 2-3 the local observation time of each individual sensor considered are visualized. This information is often essential for properly characterizing time series of cloud properties derived from the satellite-based climate datasets. Other important aspects are the imaging properties. The sensors differ in terms of native footprint resolution (1x1km² for ATSR2, AATSR; 5x1km² for AVHRR GAC). This, together with the sensor swath width, leads to very different observation frequency and spatial coverage. While AVHRRs have a complete global coverage within a day, the AATSR sensor needs about 3 days to accomplish this, however, with a higher spatial resolution compared to AVHRR.

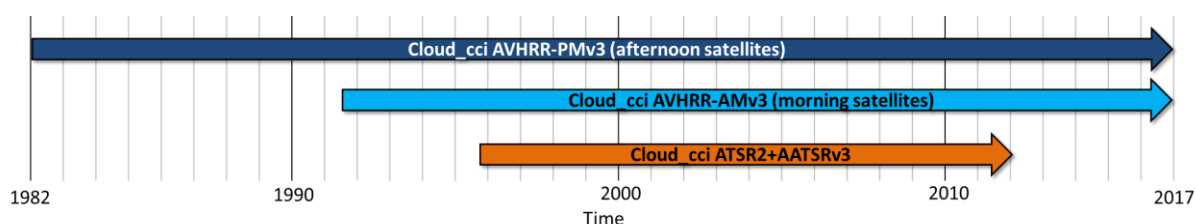



Figure 2-2 Overview of Cloud_cci v3 datasets and the time periods they cover.

All datasets contain identical sets of cloud properties: cloud mask/fraction (CMA/CFC), cloud phase/liquid cloud fraction (CPH), cloud top pressure/height/temperature (CTP/CTH/CTT), cloud effective radius (CER), cloud optical thickness (COT), spectral cloud albedo at two wave lengths (CLA) and liquid/ice water path (LWP/IWP). In addition a set of radiative broadband fluxes are contained: top of the atmosphere (TOA) and bottom of the atmosphere (BOA) radiative fluxes for shortwave and longwave, in clear-sky and all-sky conditions, upwelling and downwelling.

The data is presented at different processing levels ranging from pixel-based retrieval products (Level-2), which are additionally projected (sampling - no averaging) onto a global Latitude-Longitude grid of 0.05°

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 9

resolution (global composite, Level-3U), to monthly data summarizes including averages, standard deviation and histograms - all defined on a global Latitude-Longitude grid of 0.5° resolution (Level-3C). See Section 1.3 for more details.

All cloud properties are accompanied by uncertainty measures at all processing levels, which range from optimal estimation based uncertainty on pixel level (Level-2 and Level-3U) to propagated uncertainties in the monthly Level-3C products. See Section 1.3 for more information.

Key strengths of Cloud_cci version 3 datasets:

- The Cloud_cci datasets are based on a state-of-the art retrieval systems named CC4CL that uses the optimal estimation (OE) technique and are applied to passive imager sensors of current and past European and non-European satellite missions.
- All v3 datasets contain consistent sets of cloud and radiative flux properties.
- The measurement records of the utilized sensors have been revisited, re-characterized and, in case of AVHRR, re-calibrated.
- One special feature of CC4CL is, among others, its applicability to multiple sensors: ATSR2, AATSR, MODIS, AVHRR (and other passive imaging sensors) down to spatial footprint resolutions of 1km.
- Radiative consistency of derived cloud parameters is achieved by the OE-based, iterative fitting of a physically consistent cloud model (and radiative transfer simulations therefrom) to the sensor measurements in the visible and thermal infrared spectral range.
- Pixel-level uncertainty characterization is facilitated by the OE technique, which is physically consistent (1) with the uncertainties of the input data (e.g. measurements, a-priori) and (2) among the retrieved variables. These pixel-level uncertainties are further propagated into the monthly products using a developed sound mathematical framework.
- Potential to combine AVHRR-heritage datasets to achieve increased temporal resolution by including multiple polar-orbiting satellite instruments, which also allows for mature cloud property histograms on 0.5° resolution due to highly increased sampling rate.
- Comprehensive assessment and documentation of the retrieval schemes and the derived cloud property datasets, including possibilities of drift- and diurnal cycle corrections.
- Availability of a developed Cloud_cci satellite simulator facilitating the applicability of Cloud_cci data in regional and global climate models evaluation efforts.
- All datasets are available in netcdf (v4) format and fulfil high CCI-internal and external data standards (e.g. Climate and Forecast - CF conventions).

Table 2-1 *Cloud_cci v3 datasets with the algorithms, sensor(s) and satellite(s) used and the time periods they cover. The Digital Object Identifiers (DOI) of all datasets are also listed.*

Dataset name	Sensor(s)	Satellite(s)	Time period	Algorithm
Cloud_cci AVHRR-PMv3 DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V003	AVHRR-2/-3	NOAA-7,-9,-11,-14,-16,-18,-19	1982-2016	CC4CL
Cloud_cci AVHRR-AMv3 DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V003	AVHRR-2/-3	NOAA-12,-15,-17, Metop-A	1991-2016	CC4CL
Cloud_cci ATSR2-AATSRv3 DOI:10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V003	ATSR2, AATSR	ERS2, ENVISAT	1995-2012	CC4CL

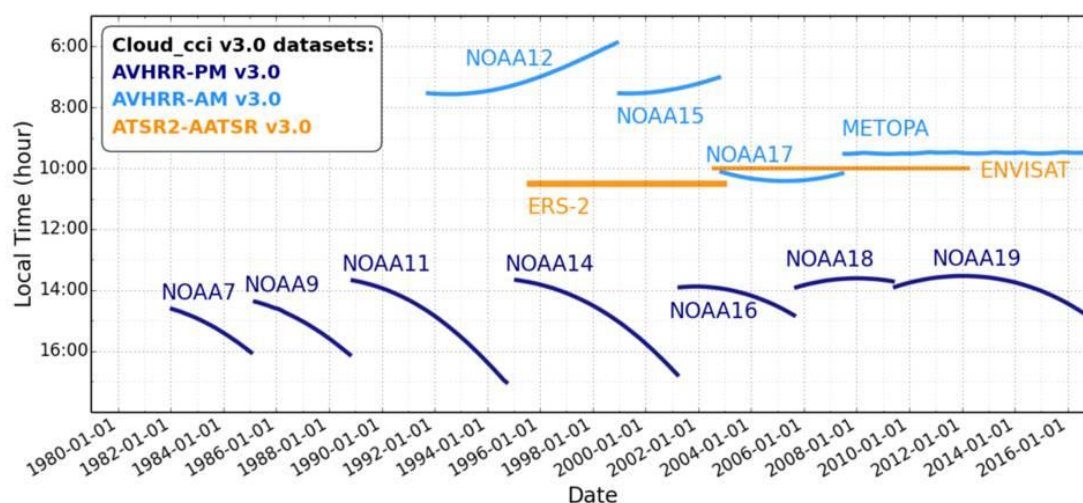


Figure 2-3 Time periods and local observation times (equator crossing times) of each satellite sensor considered in Cloud_cci.

2.3 Cloud_cci cloud products

The cloud properties derived on pixel level of each utilized sensor are listed in Table 2-2. It is important to note that the cloud properties CLA, LWP, IWP are not directly retrieved, but rather determined from retrieved COT and CER in a post processing step. The same applies to CTH and CTT, which are inferred from the retrieved CTP. In addition, it needs to be noted that for the determination of radiative fluxes a fair amount of ERA-Interim data was required.

Based on the pixel level retrievals the data is further processed into different processing levels as summarized in Table 2-3. Level-3U denotes a global composite on a global Latitude-Longitude grid (of 0.05° resolution) onto which the Level-2 data is sampled (see Section 4.1 for more details on Level-3U sampling). Level-3C products are also defined on Latitude-Longitude grid (here 0.5° resolution) onto which the properties are averaged or their frequency collected (histograms). Further separation of cloud properties in Level-3C in e.g. day/night, liquid/ice, were made wherever suitable (see Table 2-4).

Table 2-2 List of generated cloud properties. CMA/CFC and CPH are derived in a pre-processing step. In the next step, COT, CER and CTP are retrieved simultaneously by fitting a physically consistent cloud/atmosphere/surface model to the satellite observations using optimal estimation (OE). Moreover, LWP and IWP are obtained from COT and CER. In addition, spectral cloud albedo (CLA) for two visible channels are derived. In a post-processing step, derived cloud properties and ERA-Interim information are used to determine radiative broadband fluxes.

Variable	Abbrev.	Definition
Cloud mask / Cloud fraction	CMA/ CFC	A binary cloud mask per pixel (L2, L3U) and therefrom derived monthly total cloud fractional coverage (L3C, L3S) and separation into 3 vertical classes (high, mid-level, low clouds) following ISCCP classification (Rossow and Schiffer, 1999).
Cloud phase	CPH	The thermodynamic phase of the retrieved cloud (binary: liquid or ice; in L2, L3U) and the therefrom derived monthly liquid cloud fraction (L3C, L3S).



Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
Date:	03 February 2020			
Issue:	6	Revision:	1	Page 11

Variable	Abbrev.	Definition
Cloud optical thickness	COT	The line integral of the absorption coefficient and the scattering coefficient (at 0.55 μ m wavelength) along the vertical in cloudy pixels.
Cloud effective radius	CER	The area-weighted radius of the cloud drop and crystal particles, respectively.
Cloud top pressure/ height/ temperature	CTP/ CTH/ CTT	The air pressure [hPa] /height [m] /temperature [K] of the uppermost cloud layer that could be identified by the retrieval system.
Cloud liquid water path/ Ice water path	LWP/ IWP	The vertical integrated liquid/ice water content of existing cloud layers; derived from CER and COT. LWP and IWP together represent the cloud water path (CWP)
Joint cloud property histogram	JCH	This product is a spatially resolved two-dimensional histogram of combinations of COT and CTP for each spatial grid box.
Spectral cloud albedo	CLA	The blacksky cloud albedo derived for channel 1 (0.67 μ m) and 2 (0.87 μ m), respectively (experimental product)
Cloud effective emissivity	CEE	cloud radiative thickness in the infrared typically referred to as the “effective emissivity”
Top of atmosphere upwards/downwards flux	TOA	Shortwave (SW) and longwave (LW) fluxes at the Top of the atmosphere, upwelling and downwelling
Top of atmosphere upwards/downwards flux - clear-sky	TOA _{clear}	Shortwave (SW) and longwave (LW) fluxes at the Top of the atmosphere, upwelling and downwelling - for clear sky conditions
Bottom of atmosphere (surface) upwards/downwards flux	BOA	Shortwave (SW) and longwave (LW) fluxes at the Bottom of the atmosphere, upwelling and downwelling
Bottom of atmosphere (surface) upwards/downwards flux - clear-sky	BOA _{clear}	Shortwave (SW) and longwave (LW) fluxes at the Bottom of the atmosphere, upwelling and downwelling - for clear sky conditions
Photosynthetically active radiation	PAR	Bottom of atmosphere incoming shortwave radiation in the spectral range between 400 and 700nm



	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 12

Table 2-3 Processing levels of Cloud_cci data products. Level-3U, Level-3C and Level-3S are each directly derived from Level-2.

Processing level	Spatial resolution	Description
Level-2 (L2)	AATSR: 1km AVHRR: 5 km	Retrieved cloud variables at satellite sensor pixel level, thus with the same resolution and location as the sensor measurements (Level-1)
Level-3U (L3U)	Latitude-Longitude grid at 0.05° res.	Cloud properties of Level-2 orbits projected onto a global space grid without combining any observations of overlapping orbits. Only subsampling is done. Common notation for this processing level is also L2b. Temporal coverage is 24 hours (0-23:59 UTC).
Level-3C (L3C)	Latitude-Longitude grid at 0.5° res.	Cloud properties of Level-2 orbits of one single sensor combined (averaged / sampled for histograms) on a global space grid. Temporal coverage of this product is 1 month.

Table 2-4 Cloud_cci product features incl. day and night separation, liquid water and ice as well as histogram representation. Level-3U refers to the un-averaged, pixel-based cloud retrievals sampled onto a global Latitude-Longitude (lat/lon) grid. ¹CMA in Level-2 and Level-3U is a binary cloud mask. All products listed exist in each dataset listed above.

	Level 2 swath based 1km/5km	Level-3U daily sampled global 0.05° lat/lon grid	Level-3C monthly averages global 0.5° lat/lon grid	Level-3C monthly histograms global 0.5° lat/lon grid
CMA/CFC	✓ as CMA ¹	✓ as CMA ¹	✓ day/night/high/mid/low	-
CTP, CTH, CTT	✓	✓	✓	✓ liquid/ice
CPH	✓	✓	✓ day/night	-
COT	✓	✓	✓ liquid/ice	✓ liquid/ice
CER	✓	✓	✓ liquid/ice	✓ liquid/ice
LWP	✓ as CWP	✓ as CWP	✓	✓ as CWP
IWP			✓	
CLA	✓ 0.6/0.8µm	✓ 0.6/0.8µm	✓ 0.6/0.8µm	✓ 0.6/0.8µm/liquid/ice
JCH	-	-	-	✓ liquid/ice
TOA_{up,dn,sw,lw}	✓	✓	✓	-
BOA_{up,dn,sw,lw}	✓	✓	✓	-
PAR	✓	✓	✓	-

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 13

2.4 Validation strategy in this report

2.4.1 Evaluation measures

For geophysical quantities at Level-2 and Level-3U, such as cloud top height, and for aggregated products (Level-3C), we use the bias, i.e. mean difference between Cloud_cci and reference data as the metric for accuracy. In addition, the bias corrected root mean squared error (bc-RMSE) is used to express the precision of Cloud_cci compared to a reference data record.


Bias (accuracy):	Mean difference between Cloud_cci and reference data
bc-RMSE (precision):	Bias corrected root mean squared error to express the precision of Cloud_cci compared to a reference data record
Stability:	The variation of the bias over a multi-annual time period.

In case of discrete Level-2 and Level-3U variables with only two possible events, e.g. cloud mask (*clear* or *cloudy*) and cloud phase (*liquid* or *ice*), we use the following scores which can be derived from the contingency table (Table 2-5).

Score	Description	Example
POD:	Probability of Detection: The fraction of correct Cloud_cci reports of a particular category relative to all reference reports of this category.	POD for event 1: $\frac{n_{11}}{n_{11} + n_{12}}$
FAR:	False Alarm Rate: The fraction of incorrect Cloud_cci reports of a particular category relative to all Cloud_cci reports of this category.	FAR for event 1: $\frac{n_{12}}{n_{11} + n_{12}}$
Hit Rate:	The total fraction of all correct Cloud_cci reports (i.e., summing n_{11} and n_{22} in Table 2-5) relative to all reference reports.	Hit rate: $\frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}}$
KSS:	Hanssen-Kuipers Skill Score: This is a measure of correct Cloud_cci reports, with random correct and unbiased reports subtracted out.	KSS: $\frac{n_{11}n_{22} - n_{21}n_{12}}{(n_{11} + n_{21})(n_{12} + n_{22})} \in [-1, 1]$

Table 2-5: Contingency table for the 2x2 problem. n_{ij} is the number of cases where Cloud_cci reports event i and the reference reports event j . For example event 1 may be clear and event 2 may be cloudy.

	Reference reports 1	Reference reports 2
Cloud_cci reports 1	n_{11}	n_{12}
Cloud_cci reports 2	n_{21}	n_{22}


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 14

2.4.2 Datasets used for evaluation


In Table 2-6 all reference data used in this report are listed together with the cloud and/or radiation properties considered and the type of evaluation they are used for in this report. Data of CALIPSO-CALIOP, SYNOP, BSRN, ELDONET, AMSR-E, DARDAR and CERES are considered as validation reference with smaller errors than the difference to Cloud_cci product. ISCCP, CLARA-A2, PATMOS-x, MODIS C6.1, HIRS OLR and ERA-Interim are considered for comparison purposes.

Table 2-6: Reference data used in this report including a description of the type of data, the cloud properties available and the type of reference the data is considered as (validation or comparison).

Reference data	Type of data	Variables	Type of evaluation
CALIPSO-CALIOP	Retrieval based on space-based (active) Lidar measurements (See Section A.1 for details)	Level-2/Level-3U CMA, CTH, CPH	Validation
AMSR-E LWP	Space-based liquid water path over ocean based on passive microwave observations (See Section A.2 for details)	Level-2LWP	Validation
DARDAR	Space-based ice water path based on active lidar/radar observations (See Section A.3 for details)	Level-2 IWP	Validation
SYNOP	Ground-based, human observations of cloud cover (See Section A.4 for details)	Level-3C CFC	Validation
BSRN	Ground-based observations of incoming shortwave and downwelling longwave radiative fluxes (See Section A.5 for details)	Level-3 BOA SW and LW	Validation
CERES	Space-based observations of outgoing shortwave and longwave radiative fluxes (See Section A.6 for details)	Level-3 TOA SW and LWP	Validation
ELDONET	Ground-based measurement of Photosynthetically Active Radiation (See Section A.7 for details)	Level-3 PAR	Validation
CLARA-A2	Retrievals based on space-based (passive) visible and infrared measurements of AVHRR (See Section A.8 for details)	Level-3C CFC, CPH, CTP, CER, COT, LWP, IWP, JCH	Comparison

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 15

Reference data	Type of data	Variables	Type of evaluation
PATMOS-x	Retrievals based on space-based (passive) visible and infrared measurements of AVHRR (See Section A.9 for details)	Level-3C CFC, CPH, CTP, CER, COT, LWP, IWP, JCH	Comparison
MODIS Collection 6.1	Retrievals based on space-based (passive) visible and infrared measurements of MODIS (See Section A.10 for details)	Level-3C CFC, CPH, CTP, CER, COT, LWP, IWP, JCH	Comparison
ISCCP	Retrievals based on space-based (passive) visible and infrared measurements combined from geostationary and polar-orbiting satellites (See Section A.11 for details)	Level-3C CFC, CPH, CTP, COT, LWP, IWP, JCH	Comparison
HIRS OLR	Space-based observations of outgoing longwave radiation (See Section A.12 for details)	Level-3 TOA LW	Comparison
ERA-Interim	ECMWF Reanalysis (See Section A.13 for details)	Level-3 TOA, BOA SW and LW	Comparison

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx				
	Date:	03 February 2020				
	Issue:	6	Revision:	1	Page 16	

3. Validation of Cloud_cci products

This section lists the results of the validation effort inferred from comparison Cloud_cci cloud and radiation products to reference observations that are expected to be of higher accuracy. In more detail, this section contains pixel-based comparisons to A-Train data (CALIOP-CALIPSO, CPR-CloudSat and AMSR-E) in subsection 3.1, Level-3C (monthly mean) comparisons to ground-based SYNOP (Section 3.2) as well as Level-3C comparisons to BSRN and CERES (Section 3.3).

3.1 Level-2/Level-3U cloud products

Cloud_cci pixel-based (Level-2) cloud data were acquired from Level-3U data products and collocated to the footprints of A-Train sensors CALIOP for validating cloud mask (Section 3.1.1), cloud phase (Section 3.1.2) and cloud top height (Section 3.1.3). Collocation criteria were maximum spatial and temporal mismatch of 5 km and 3 minutes, respectively. For the ATSR2-AATSRv3 data, the temporal criterion was relaxed to 15 minutes to allow the collection of a sufficiently large data base.


3.1.1 Cloud mask

The cloud mask validation was separated into two scenarios: (a) counting all those CALIOP observations as clear-sky for which the reported CALIOP COT was equal to 0, and (b) additionally labelling also those CALIOP observations as clear-sky which have an CALIOP COT of greater than 0 but below 0.15. The latter scenario is included to account for and quantify the limited sensitivity of passive imagers to optically thin clouds under certain conditions. Table 3-1 reports the validation scores for both scenarios and all three datasets in addition to the number of underlying collocated data pairs.

Cloud mask (CMA) validation against CALIOP proves the good cloud detection performance of CC4CL with hit rates generally higher than 74%. For nearly all comparisons a slight underestimation of cloud occurrences is found in the Cloud_cci data compared to CALIOP, which is primarily due to a lack of sensitivity of passive imager data with respect to optically very thin clouds. Removing these optically thin clouds from the statistic, which is done by using CALIOP cloud optical thickness (COT), shows a clear improvement in the found detection scores (e.g. hit rates above 80% for all datasets).

Table 3-1: Summary of validation scores for cloud mask using CALIOP reference with and without applying an optical thickness filtering on CALIOP data. Collocation time window: ¹ 3 minutes, ² 15 minutes

	AVHRR-AMv3 ¹		AVHRR-PMv3 ¹		ATSR2-AATSRv3 ²	
	COT=0.0	COT=0.15	COT=0.0	COT=0.15	COT=0.0	COT=0.15
Matched FOVs	140 805	140 805	16 139 764	16 139 764	1 831 105	1 831 105
Bias	-18.3	-6.2	-17.4	-5.4	-16.5	-5.1
RMSE	0.51	0.45	0.46	0.41	0.49 %	0.44 %
POD cloudy	69.7 %	78.6 %	75.8 %	83.3 %	72.6 %	80.8 %
POD clear	86.3 %	82.7 %	92.0 %	82.5 %	85.6 %	81.5 %
FAR cloudy	6.9 %	12.5 %	2.8 %	9.5 %	6.5 %	12.1 %
FAR clear	48.3 %	28.5 %	49.5 %	29.0 %	47.6 %	28.3 %
Heidke Skill score	0.56	0.61	0.68	0.66	0.58	0.62
Hit Rate	74.2 %	80.2 %	79.23 %	83.0 %	76.0 %	81.1 %

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 17

3.1.2 Cloud phase

The cloud phase validation was done similarly to the cloud mask validation (previous section), thus separated into two scenarios, with, however, different definitions: (a) selecting the CALIOP cloud phase of the uppermost reported cloud layer as reference observations, and (b) selecting that CALIOP cloud phase as reference that belongs to the uppermost cloud layer after removing all top layers with an integrated COT of lower than 0.15. Also here, the latter scenario is included to account for and quantify the limited sensitivity of passive imagers to optically thin cloud layers. Table 3-2 reports the validation scores for both scenarios and all three datasets in addition to the number of underlying collocated data pairs.

Cloud phase (CPH) validation against CALIOP shows hitrates of 74% and higher when using the CALIOP phase of the uppermost detected cloud layer. Neglecting the very thin cloud layers (with COT lower than 0.15) improves the scores significantly: increasing hitrate scores to 77% and beyond (84% for AVHRR-PMv3). All datasets have a bias towards liquid clouds, when compared against to uppermost layer. The bias becomes negative for the scenario of neglecting the very thin cloud layer.

Table 3-2: Summary of validation scores for cloud phase using CALIOP reference with and without applying an optical thickness filtering on CALIOP profiles. Collocation time window: ¹ 3 minutes, ² 15 minutes


	AVHRR-AMv3 ¹ (NOAA-17,METOP-A)		AVHRR-PMv3 ¹ (NOAA-18,NOAA-19)		ATSR2-AATSRv3 ² (ERS-2, Envisat)	
	COT=0.0	COT=0.15	COT=0.0	COT=0.15	COT=0.0	COT=0.15
Matched FOVs	68 047	65 640	8 788 655	8 435631	940 837	909 370
Bias	4.1 %	-8.5 %	9.4 %	-2.7 %	6.1 %	-5.2 %
POD liquid	72.6 %	70.6 %	86.3 %	82.3 %	79.6 %	76.8 %
POD ice	75.3 %	85.4 %	75.4 %	87.1 %	78.1 %	87.8 %
Heidke Skill score	0.48	0.56	0.62	0.69	0.58	0.65
Hit Rate	74.2 %	77.7 %	79.7 %	84.7 %	78.7 %	82.4 %

3.1.3 Cloud top height

The cloud top height validation was done by separating the validation into three scenarios: (a) selecting the CALIOP cloud top height of the uppermost reported cloud layer as reference observations, (b) selecting that CALIOP cloud height as reference that belongs to the uppermost cloud layer after removing all top layers with an integrated COT of lower than 0.15 and (c) selecting that CALIOP cloud height as reference that belongs to the uppermost cloud layer after removing all top layers with an integrated COT of lower than 1.0. The latter scenario are done to account for and quantify the limited sensitivity of passive imagers to optically thin cloud layers. The validation is separated into liquid and ice clouds, for which both data sources had to agree on the phase for conducting the phase stratification.

Table 3-3 reports the validation scores for liquid clouds for all three scenarios and all three datasets in addition to the number of underlying collocated data pairs. For liquid clouds only very small biases (-0.1 to 0.12 km) and bc-RMSD (0.8 to 0.98 km) are found, which are only slightly sensitive to the chosen COT threshold when removing the uppermost clouds layers.

Table 3-4 reports the corresponding results for ice clouds. Here, the strong underestimation of cloud top height is evident and a common feature for all three v3 datasets, and mainly caused by high-level, optically

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 18

thin clouds. Biases are around -3.5 km and bc-RMSE around 2.3 km. Removing the optically very thin cloud layers at the top of the CALIOP profiles, improves the agreement between Cloud_cci and CALIOP significantly. When removing all cloud layers with COTs (integrated from layer bottom to top of all cloud layers) lower than 1.0, biases are reduced to approximately -1.5 km while bc-RMSE remain nearly unchanged.

Table 3-3: Validation scores for cloud top height (liquid clouds) using CALIOP reference with and without applying an optical thickness filtering on CALIOP profiles. Collocation time window: ¹ 3 minutes, ² 15 minutes

Liquid clouds	AVHRR-AMv3 ¹ (NOAA-17,METOP-A)			AVHRR-PMv3 ¹ (NOAA-18,NOAA-19)			ATSR2-AATSRv3 ² (ERS-2, Envisat)		
	COT=0.0	COT=0.15	COT=1.0	COT=0.0	COT=0.15	COT=1.0	COT=0.0	COT=0.15	COT=1.0
FOVs	19 485	24 239	24 545	2 603 163	3 016 985	2 982 690	247 224	306 512	311 296
Bias [km]	-0.05	-0.01	0.12	-0.10	-0.01	-0.05	0.12	0.13	0.24
Std [km]	0.98	1.00	0.94	0.86	0.91	0.80	0.97	1.00	0.93


Table 3-4: Validation scores for cloud top height (ice clouds) using CALIOP reference with and without applying an optical thickness filtering on CALIOP profiles. Collocation time window: ¹ 3 minutes, ² 15 minutes

Liquid clouds	AVHRR-AMv3 ¹ (NOAA-17,METOP-A)			AVHRR-PMv3 ¹ (NOAA-18,NOAA-19)			ATSR2-AATSRv3 ² (ERS-2, Envisat)		
	COT=0.0	COT=0.15	COT=1.0	COT=0.0	COT=0.15	COT=1.0	COT=0.0	COT=0.15	COT=1.0
FOVs	30 898	26 671	15 296	3 691 179	3 376 337	2 077 074	455 061	400 939	232 444
Bias [km]	-3.25	-2.60	-1.25	-3.54	-2.95	-1.62	-3.57	-3.01	-1.76
Std [km]	2.29	2.12	2.25	2.36	2.14	1.95	2.30	2.20	2.33

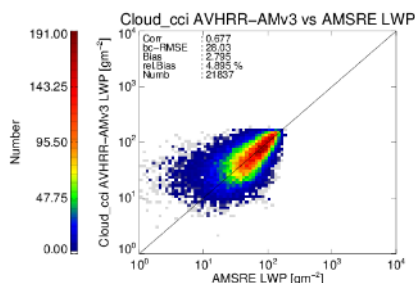
3.1.4 Liquid water path

Reference for the LWP validation is the AMSR-E products. Due to the large AMSR-E footprint, all Cloud_cci LWP observations that fall into that footprint had to be averaged including clear-sky pixels counting with a LWP of 0. AMSR-E footprints for which Cloud_cci reported an ice fraction of greater-equal 20% were excluded from the validation. This is done to excluded cases (ice clouds) for which the microwave signal is impacted by significant scattering at ice particles, which are unfavourable circumstances for the retrieval of liquid water path from AMSR-E. As AMSR-E observations are reliable over ocean only, all land regions were also excluded from this LWP validation.

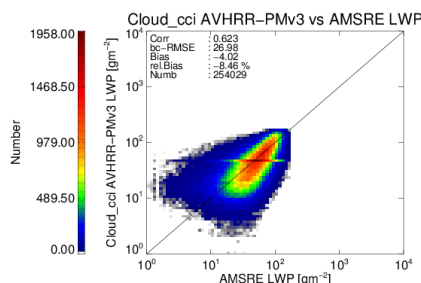
Figure 3-1 reports the LWP validation scores in addition to the number of underlying collocated data pairs. This validating gives very convincing results for all three Cloud_cci v3 datasets, with bc-RMSE values of 25 to 28 g/m², only small biases and high correlations (above 0.62).

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 19

AVHRR-AMv3¹



AVHRR-PMv3¹



ATSR2-AATSRv3²

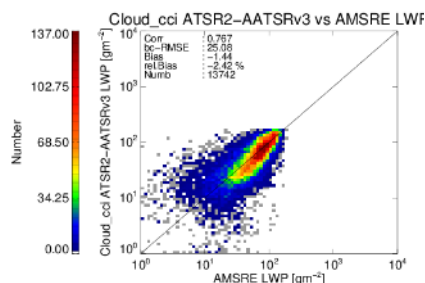


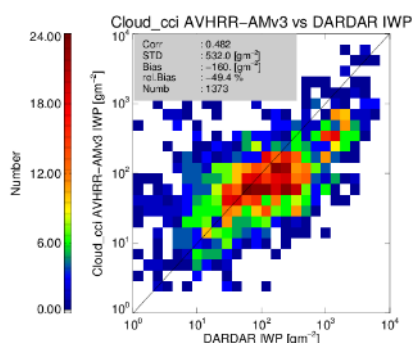
Figure 3-1 2d-frequency distribution and validation scores for cloud liquid water path using AMSR-E as reference. Collocation time window: ¹ 3 minutes, ² 15 minutes. All Cloud_cci data within an AMSR-E footprint was averaged.

3.1.5 Ice water path

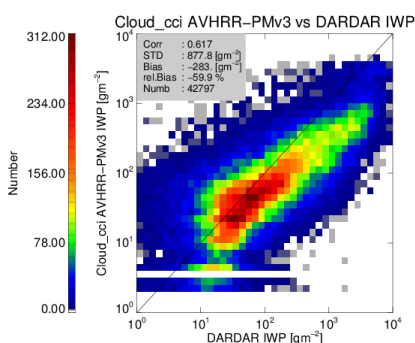
Reference for the IWP validation is the DARDAR IWP product, which make synergistic use of CALIOP-CALIPSO and CPR-CloudSat data. Pixels for which either Cloud_cci or the reference did not report ice clouds were excluded from the validation.

Figure 3-2 reports the IWP validation scores in addition to the number of underlying collocated data pairs. The results shows good agreement with correlations between 0.45 and 0.62. However, there is a general underestimation of IWP by Cloud_cci which in terms of relative bias partly exceeds 50%.

AVHRR-AMv3¹



AVHRR-PMv3¹



ATSR2-AATSRv3²

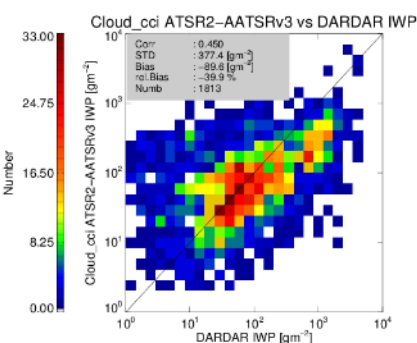



Figure 3-2 2d-frequency distribution and validation scores for cloud ice water path using DARDAR as reference. Collocation time window: ¹ 3 minutes, ² 15 minutes.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 20

3.2 Level-3C cloud fraction (CFC) against SYNOP

In this section we investigate the monthly mean cloud fraction of all datasets wrt. their agreement to ground based SYNOP observation (see Section A.4 for details on SYNOP data). For each SYNOP station the climatological monthly mean of the cloud fraction has been calculated from all available cloudiness observations (given in octa) during the day. For each SYNOP site, the Cloud_cci CFC of nearest Level-3C grid cell has been selected. For each dataset (Figure 3-3 to Figure 3-5) we present time series of CFC for Cloud_cci and SYNOP averaged over all SYNOP site locations as well as a time series of the corresponding monthly Bias and monthly standard deviation. In addition, we show the mean bias for each time period for each station as a map, separated into the four seasons.

It is important to note that the cloud mask validation against CALIOP, given in Section 3.1.1, are based on spatiotemporally collocated observations and are thus not affected by any sampling errors. For the validation against SYNOP shown in this section Level-3 data was utilized without accounting for the different temporal sampling of satellite sensors and SYNOP. Thus, the results shown in this section give the total error of the Cloud_cci monthly cloud fraction, including the sampling error.


Generally speaking, all datasets show a good agreement to SYNOP cloudiness. However, a few features are found that need to be mentioned:

(1) AVHRR-AM (entire period), AVHRR-PM (before 2006) and ATSR2-AATSR (before 2005) have a slight seasonal cycle in their agreement to SYNOP with a nearly zero bias in Northern Hemispheric summer while showing slight overestimations of 5% during winter. The latter seems to be caused by mid-latitude land regions in Europe and Asia.

(2) AVHRR-AM shows strong overestimation of cloudiness for the period before 1999 with highest amplitude in Northern Hemispheric winter. This is apparently a problem with NOAA-12 which is not fully understood yet. NOAA-12 flies in a twilight orbit, which are characterized by very difficult illumination conditions. However, so is NOAA-15, for which the same feature cannot be found.

(3) AVHRR-AM (1999 and onwards), AVHRR-PM (entire period) and ATSR2-AATSR (1997 onwards): small biases and standard deviations, temporally relative stable.

(4) Two distinct regions are visible in nearly all comparisons: a) the USA for which we have an underestimation of CFC against SYNOP nearly all the time and in all datasets, b) the Sahel zone at the southern end of the SAHARA desert, for which an underestimation of cloudiness is visible, again, nearly all the time for all datasets. The reasons for these two features are not known yet.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
		Page 21		

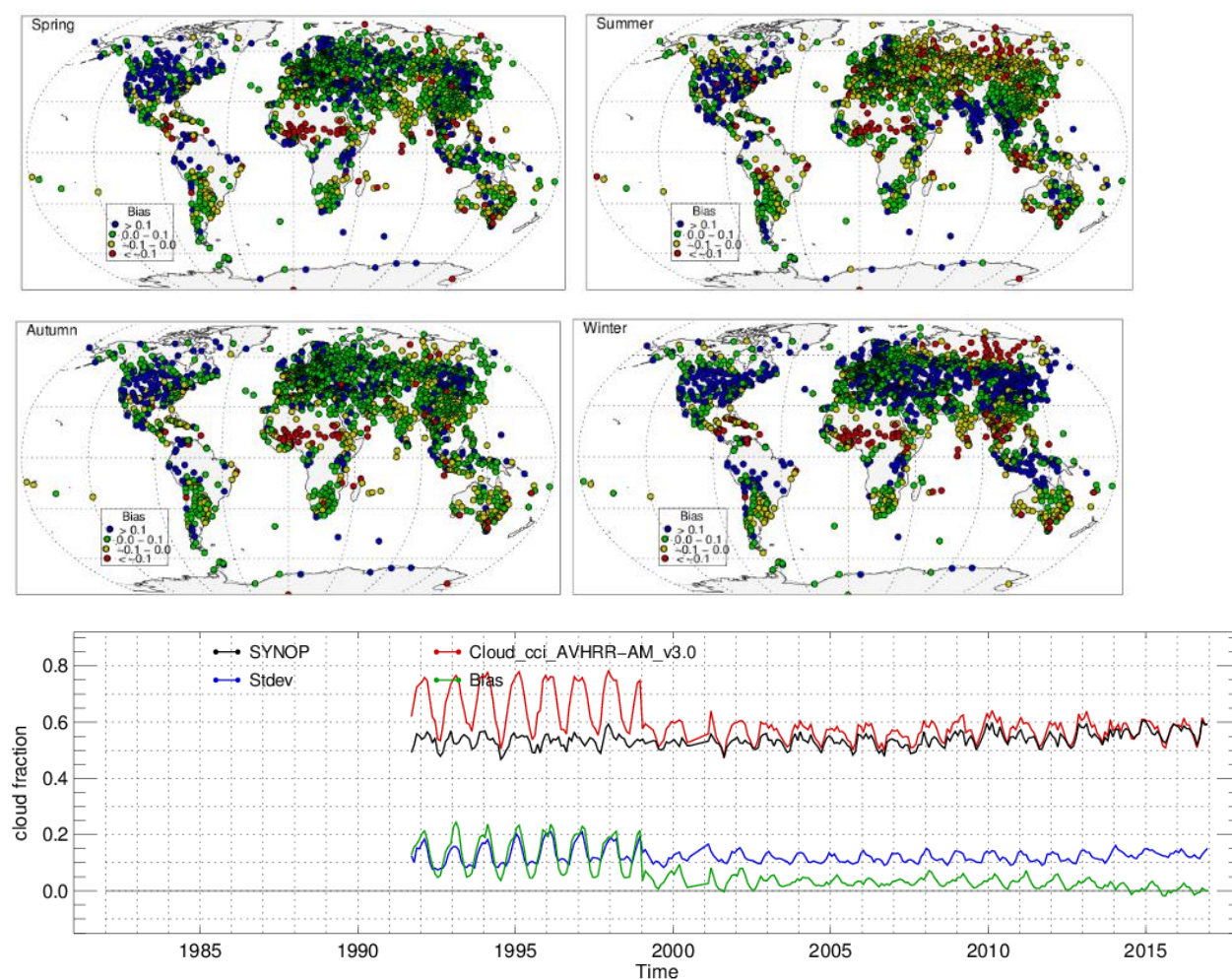



Figure 3-3 Top panels: Seasonal map of CFC Bias (Cloud_cci AVHRR-AMv3 minus SYNOP) at each SYNOP site, averaged over the time period of the dataset (1991-2016). Bottom panel: Time series of Cloud_cci AVHRR-AMv3 and SYNOP CFC averaged over all SYNOP sites, as well as corresponding monthly Bias and standard deviation.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 22				

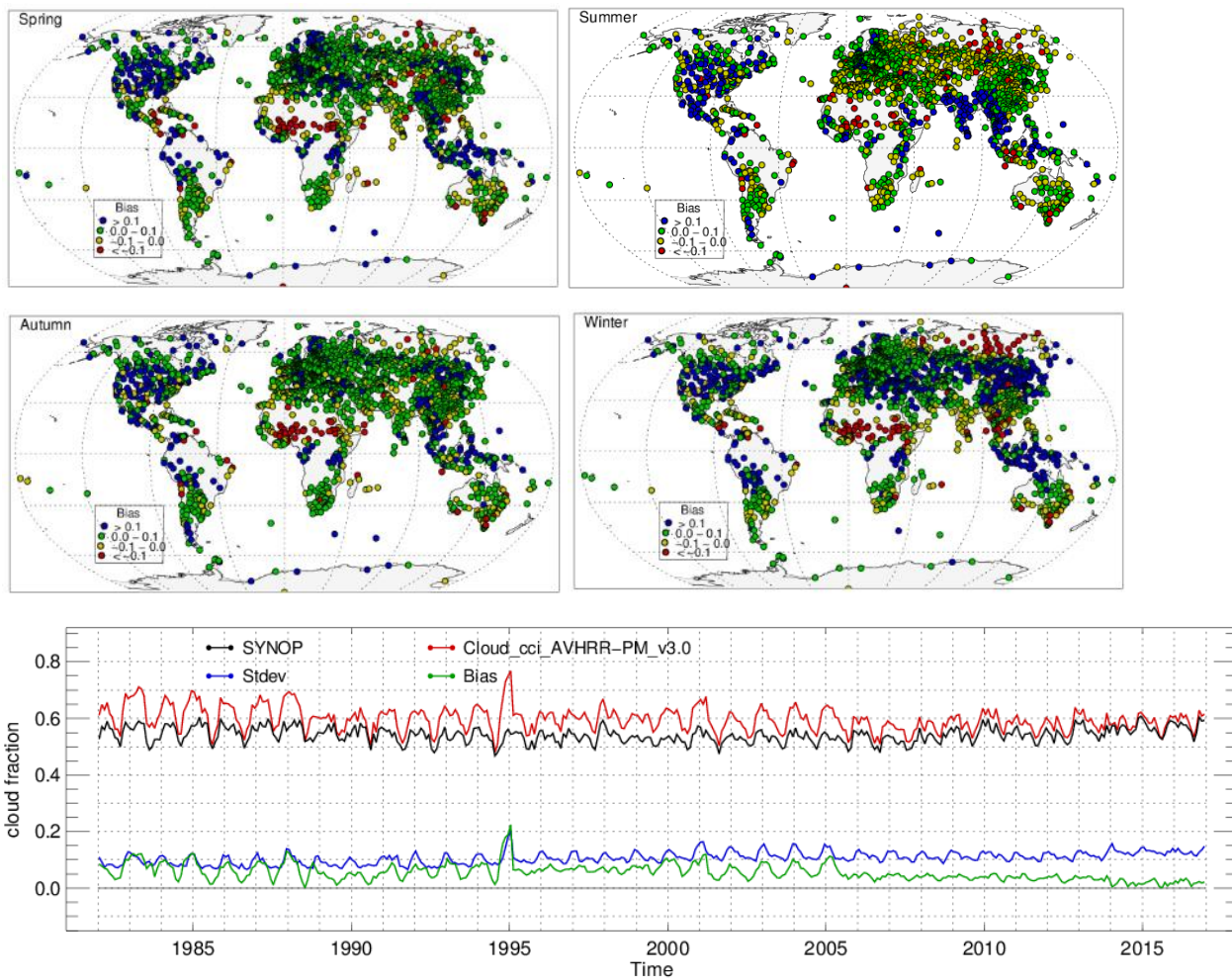


Figure 3-4 As Figure 3-3 but for Cloud_cci AVHRR-PMv3 (1982-2016).

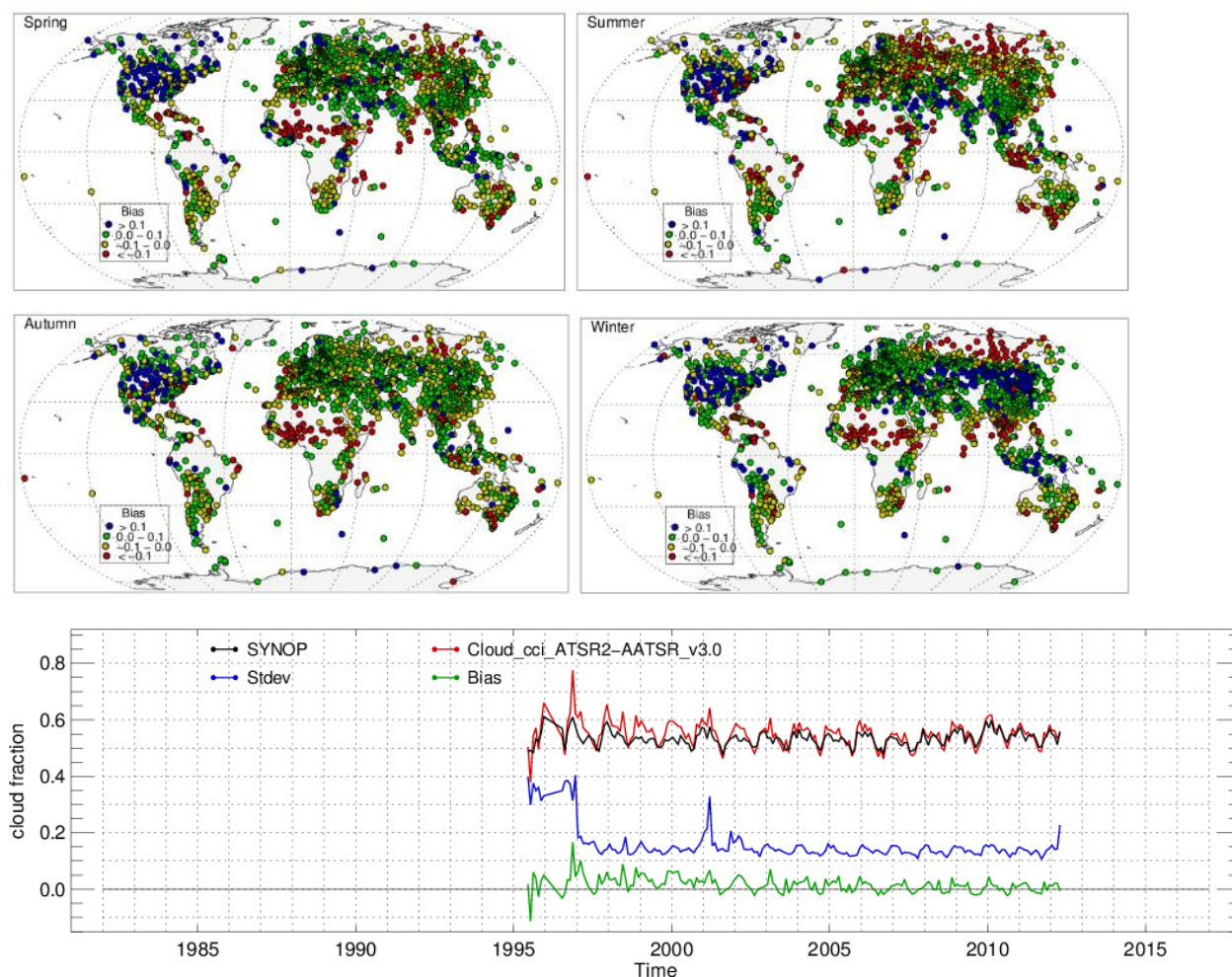



Figure 3-5 As Figure 3-3 but for Cloud_cci ATSR2-AATSRv3 (1995-2012).

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 24

3.3 Level-3C radiation products

Space-based CERES observations of top of atmosphere radiation as well as ground measurements of downwelling SW and LW radiation taken at sites of the international Baseline Surface Radiation Network (BSRN) were used to validate Cloud_cci TOA and BOA radiative fluxes.

3.3.1 TOA fluxes

Figure 3-6 reports two-dimensional frequency distributions and validation scores for monthly mean TOA upwelling shortwave fluxes for all three Cloud_cci datasets. The histograms reveal an excellent agreement of all Cloud_cci dataset with CERES with standard deviations ranging from 11.6 W/m² (AVHRR-PMv3) to 15.6 W/m² (ATSR2-AATSRv3), with the higher values for ATSR2-AATSRv3 probably due to less temporal sampling. Biases are all slightly positive ranging from 3.6 W/m² to 7 W/m². In Table 3-5 the validation scores are additionally stratified by USGS surface type. Biases are generally very low for ocean regions (with low surface albedo) and high in regions with only little vegetation (usually high surface albedo).

Figure 3-7 reports two-dimensional frequency distributions and validation scores for TOA upwelling longwave fluxes for all three Cloud_cci datasets. The agreement of all Cloud_cci datasets with CERES is even better than for shortwave: biases between -2 W/m² and 4 W/m² and standard deviations between 3.4 W/m² and 12.4 W/m². Correlations are extremely high with values above 0.93. In Table 3-6 the validation scores are stratified by USGS surface type. The validation scores are similar for all surface types with no significant outliers.

AVHRR-AMv3

AVHRR-PMv3

ATSR2-AATSRv3

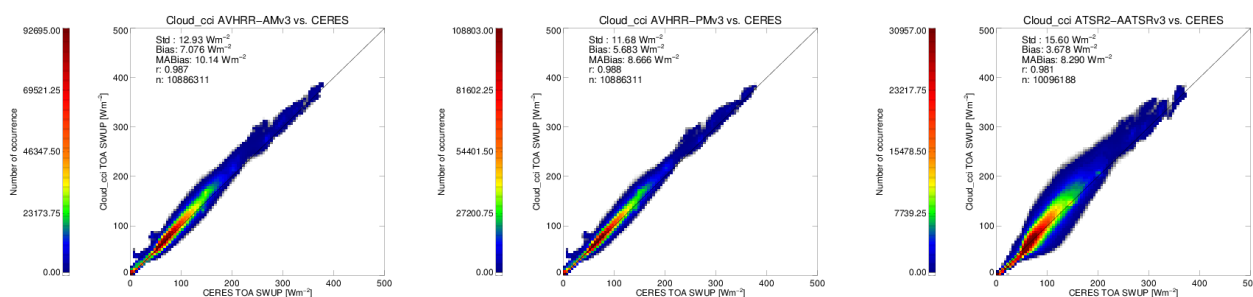


Figure 3-6 Cloud_cci top of atmosphere (TOA) monthly mean upwelling shortwave flux validation against to CERES EBAF-TOA Ed4.0. Data basis contains monthly mean, global data in the period 2003-2016.


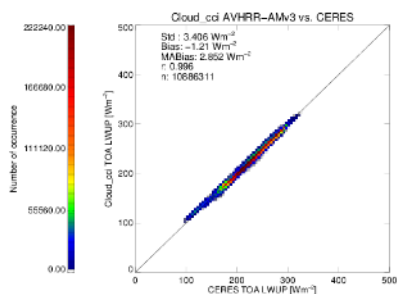
	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 25

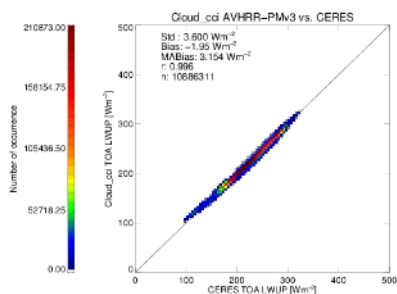
Table 3-5 *Cloud_cci top of atmosphere (TOA) monthly mean upwelling shortwave flux validation against to CERES EBAF-TOA Ed4.0 stratified by USGS surface type. Data basis contains monthly mean, global data in the period 2003-2016.*

	AVHRR-AMv3		AVHRR-PMv3		ATSR2-AATSRv3	
USGS class	Bias [Wm ⁻²]	Std [Wm ⁻²]	Bias [Wm ⁻²]	Std [Wm ⁻²]	Bias [Wm ⁻²]	Std [Wm ⁻²]
Water	6.05834	12.5741	3.06475	9.84631	2.23609	14.0283
Evergreen Needle leaf Forest	7.96467	14.6814	11.0898	14.2897	14.7714	22.6394
Evergreen Broadleaf Forest	9.96015	13.9383	11.2791	13.9853	6.24851	24.3378
Deciduous Needle leaf Forest	11.1215	15.6064	13.4994	15.1972	18.1622	23.8647
Deciduous Broadleaf Forest	6.43556	9.75003	9.20488	9.91884	8.72030	21.9931
Mixed Forests	8.58068	14.2590	11.8481	13.7353	14.8984	23.4949
Closed Shrublands	4.93758	13.3571	9.69724	11.5818	9.44300	18.8384
Open Shrublands	8.51614	11.2245	11.1522	11.6475	11.3251	16.6845
Woody Savannas	7.38031	12.6986	9.96802	12.8305	8.47494	20.8258
Savannas	7.28460	11.2011	11.2470	11.7535	6.98021	19.8989
Grasslands	6.01761	13.5122	9.57824	13.2994	13.0631	21.8042
Permanent Wetland	6.66528	11.3435	8.25270	11.7109	8.75107	19.4323
Croplands	6.09091	9.01186	8.09128	9.00241	10.3915	18.2135
Urban and Built-Up	4.31478	10.1817	7.17569	9.17166	8.27763	20.2296
Cropland/Natural Vegetation Mosaic	7.53897	11.4251	9.87849	11.7634	10.1917	20.8116
Snow and Ice	9.27160	13.6960	9.71686	13.6946	3.90981	14.9866
Barren or Sparsely Vegetated	19.1091	14.5303	20.3978	12.8995	27.8694	18.9095

AVHRR-AMv3



AVHRR-PMv3



ATSR2-AATSRv3

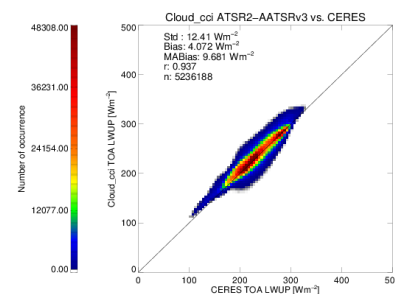


Figure 3-7 *Cloud_cci top of atmosphere (TOA) monthly mean upwelling longwave flux validation against to CERES EBAF-TOA Ed4.0. Data basis contains monthly mean, global data in the period 2003-2016.*


	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 26

Table 3-6 *Cloud_cci top of atmosphere (TOA) monthly mean upwelling longwave flux validation against to CERES EBAF-TOA Ed4.0 stratified by USGS surface type. Data basis contains monthly mean, global data in the period 2003-2016.*

	AVHRR-AMv3		AVHRR-PMv3		ATSR2-AATSRv3	
USGS class	Bias [Wm ⁻²]	Std [Wm ⁻²]	Bias [Wm ⁻²]	Std [Wm ⁻²]	Bias [Wm ⁻²]	Std [Wm ⁻²]
Water	-1.43841	2.93552	-1.90430	2.96760	1.92388	11.9424
Evergreen Needle leaf Forest	-2.07370	2.49140	-2.75351	2.57222	6.20787	9.53845
Evergreen Broadleaf Forest	-1.43247	4.58355	-5.14017	5.60383	11.5764	17.1009
Deciduous Needle leaf Forest	-1.70313	2.25475	-2.22613	2.20823	5.74411	6.73632
Deciduous Broadleaf Forest	-2.21790	3.54086	-4.20096	4.44610	5.41626	13.5648
Mixed Forests	-2.34746	2.82871	-3.40389	3.09671	5.21104	10.1855
Closed Shrublands	-0.504582	4.21184	-3.56148	4.51125	10.2281	13.9226
Open Shrublands	-2.31509	3.52065	-2.38175	3.84209	8.93722	11.0713
Woody Savannas	-1.47239	3.65843	-2.98450	4.42628	10.3412	13.0816
Savannas	-0.972261	4.18875	-4.10281	5.33563	11.9062	14.8795
Grasslands	-1.53631	4.03848	-2.96349	4.37169	9.46873	12.3968
Permanent Wetland	-2.01237	2.89896	-2.73342	3.44681	7.11700	10.9527
Croplands	-1.37773	3.42826	-2.86797	3.95718	8.29827	12.8002
Urban and Built-Up	-1.05430	3.90562	-2.67787	3.88509	6.94820	13.1080
Cropland/Natural Vegetation Mosaic	-1.61452	3.86298	-3.69254	4.68043	9.31071	14.3124
Snow and Ice	0.862723	4.33736	-0.412011	4.46359	6.02921	8.03100
Barren or Sparsely Vegetated	-0.491311	4.39070	-0.729419	4.99863	13.5958	11.7947


3.3.2 BOA fluxes

BSRN stations measure direct, diffuse and global downwelling shortwave and longwave fluxes in 1 min temporal resolution. The manned stations are located at positions, which are representative of a relatively large surrounding area for the use in satellite data validation among others. The 1-min data was aggregated to monthly averages which were used as validation data.

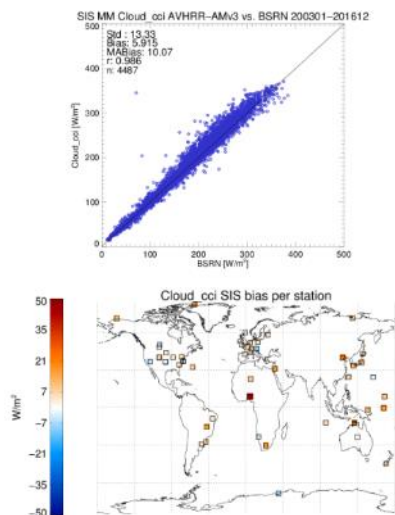
Figure 3-8 shows scatter plots and validation scores for monthly mean BOA downwelling shortwave radiation for all three datasets. Confirming the good TOA validation results reported in Section 3.3.1, also the BOA validation of Cloud_cci v3 datasets against BSRN highlight the good quality. Standard deviations lie between 13 W/m² and 24 W/m² with lower values for the AVHRR datasets, and biases range from 1.9 W/m² to 8.2 W/m². Correlations are again very high and above 0.96 for all three datasets. The shown bias maps reveal a rather uniform performance over all stations, with the exceptions of three stations (in South-west Africa, North-Australia, Eastern China) with relative high biases, consistently for all three Cloud_cci v3 datasets.

Figure 3-9 shows the validation results for BOA downwelling longwave radiation for all three datasets. Similar to the TOA validation, the longwave fluxes of Cloud_cci are of even better agreement with the references in terms of standard deviation (11.4 W/m² to 14.6 W/m²) and correlations (0.97 to 0.99) and only the biases are slightly higher (7.3 W/m² to 11.9 W/m²). The bias maps reflect the found positive bias at nearly all BSRN stations.

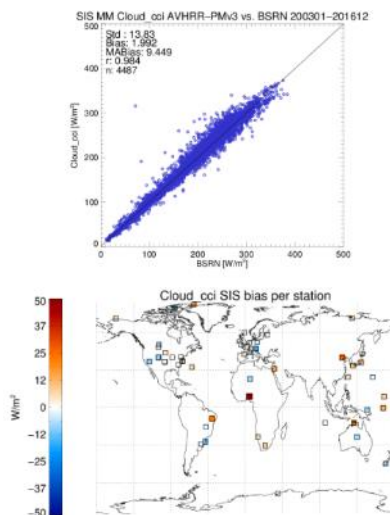
Figure 3-10 shows the validation results for monthly mean BOA PAR for AVHRR-AM and AVHRR-PM datasets (ATSR2-AATSRv does not include PAR). The data basis is relative sparse, but the results nevertheless indicate a reasonable agreement with high correlations (above 0.8) small positive biases (4.7 W/m² to 5.8 W/m²) and standard deviations around 20 W/m².

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 27

AVHRR-AMv3



AVHRR-PMv3



ATSR2-AATSRv3

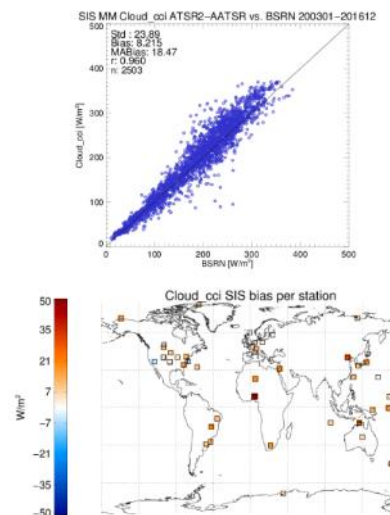
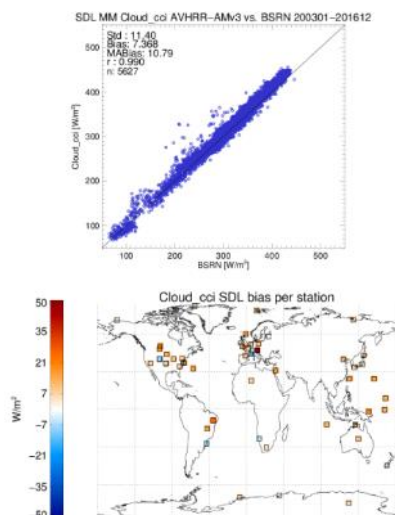
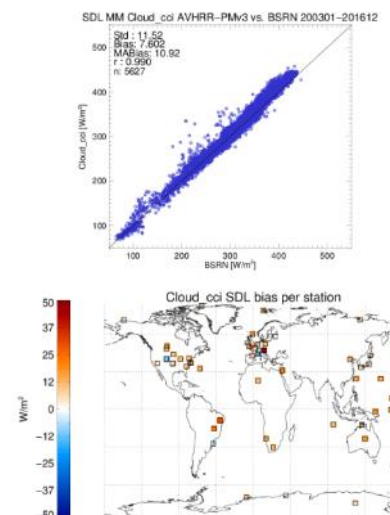


Figure 3-8 Cloud_cci bottom of atmosphere (BOA) monthly mean downwelling shortwave flux validation against to BSRN. Data basis contains monthly mean, global data in the period 2003-2016.

AVHRR-AMv3



AVHRR-PMv3



ATSR2-AATSRv3

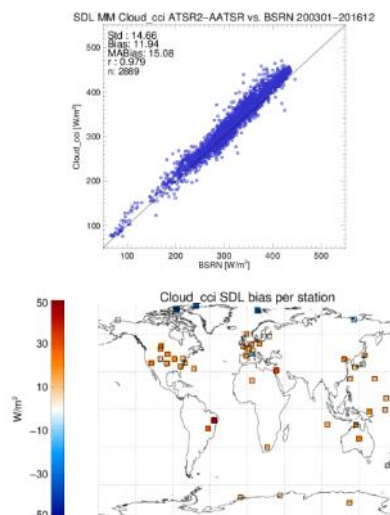

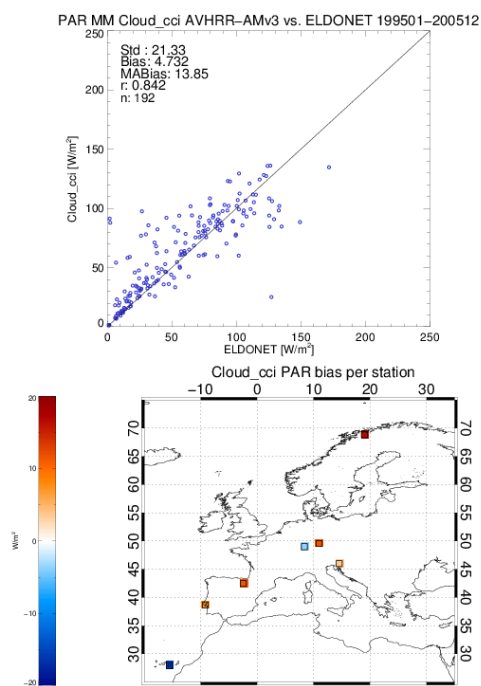


Figure 3-9 Cloud_cci top of atmosphere (TOA) monthly mean downwelling longwave flux validation against to BSRN. Data basis contains monthly mean, global data in the period 2003-2016.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 28

AVHRR-AMv3



AVHRR-PMv3

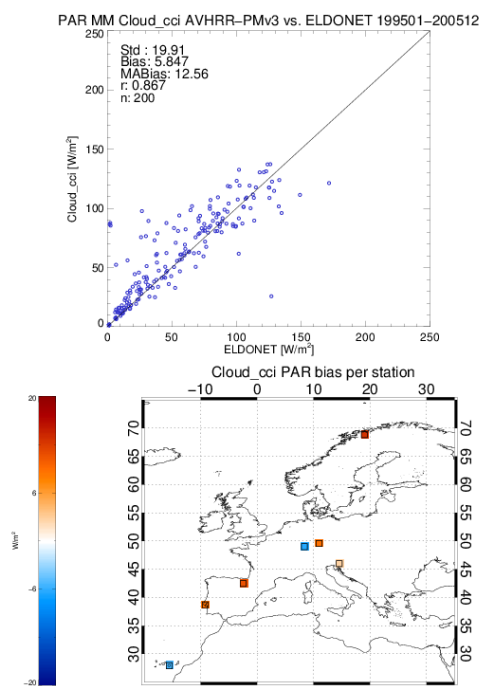



Figure 3-10 *Cloud_cci bottom of atmosphere (BOA) monthly mean Photosynthetically Active Radiation (PAR) validation against to ELDONET. Data basis contains monthly mean, European-area data in the period 1997-2009 in which the ELDONET data density is very diverse.*

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 29

4. Intercomparison of Cloud_cci cloud products with other satellite-based datasets

In this section the Cloud_cci datasets are compared to other, well-established, satellite-based datasets. The bases for all comparisons are monthly mean values (Level-3C in ESA CCI notation). The comparisons are separated into two classes for morning (Section 4.1) and afternoon (Section 4.2) datasets, according to the local (daytime) observation time of the individual, underlying satellites:

- Morning satellites: Terra, NOAA-12, NOAA-15, NOAA-17, Metop-A, ENVISAT, ERS2
- Afternoon satellites: Aqua, NOAA-7, NOAA-9, NOAA-11, NOAA-14, NOAA-16, NOAA-18, NOAA-19

The International Satellite Cloud Climatology Project (ISCCP), which is composed of geostationary and polar-orbiting data is utilized in the afternoon satellites discussion. All comparisons are stratified into the individual cloud properties. For each of these subsections, global maps of multi-annual means (of a common time period) and multi-annual (monthly) standard deviations are shown, together with zonal mean plots. In addition, time series plots of latitude-weighted global mean values (for a latitude band of 60S-60N) are given without restriction to the common time period. Using MODIS Collection 6.1 data as reference, the Level 3C scores bias and bias-corrected root mean square difference (bc-RMSD) are calculated for each Cloud_cci v3 dataset for the common period and 60S-60N. The stability of the Cloud_cci v3 datasets is investigated by calculating the linear trends and comparing these again to MODIS Collection 6.1 data.

4.1 Morning Satellites

In this subsection the morning satellite datasets are compared, these are:

- Cloud_cci AVHRR-AMv2 (see Section 2.2)
- Cloud_cci AVHRR-AMv3 (see Section 2.2)
- Cloud_cci ATSR2-AATSRv2 (see Section 2.2)
- Cloud_cci ATSR2-AATSRv3 (see Section 2.2)
- CLARA-A2 (only subset of morning satellites used; see Section A.8 for details)
- MODIS Collection 6.1 Terra (see Section A.10 for details)

For interpreting the following comparisons of cloud optical thickness, cloud effective radius, liquid water path and ice water path correctly it is necessary to revisit the channel settings of all sensors/datasets:

For Cloud_cci ATSR2-AATSRv2 the 3.7 μ m channel was used as near-infrared channel, while with Cloud_cci ATSR2-AATSRv3 the 1.6 μ m channel was used. For the first half of Cloud_cci AVHRR-AMv2 and v3 the 3.7 μ m channel was used (NOAA-12 and NOAA-15), while the second half of AVHRR-AM datasets (NOAA-17 and Metop-A) the 1.6 μ m channel was used as near-infrared channel. In CLARA-A2 the same channels as in AVHRR-AM were used. MODIS C6.1 Terra has three NIR channels and MODIS Collection 6.1 data usually has optical property products for each of them. In our comparisons, the maps always show the 2.1 μ m product, while in the time series and zonal mean plots 1.6 μ m and 3.7 μ m MODIS Collection 6.1 products are visualized if both exist.

Table 4-1 Near-infrared channels utilized for morning satellite datasets that are included in the comparisons in this section. This information is needed for correctly interpreting the comparison results shown for cloud optical thickness, cloud effective radius, liquid water path and cloud water path. Time periods given are approximated.

Channel used	Cloud_cci AVHRR-AMv2	Cloud_cci AVHRR-AMv3	Cloud_cci ATSR2-AATSRv2	Cloud_cci ATSR2-AATSRv3	CLARA-A2	MODIS Collection 6.1 Terra
1.6 μ m	2002-2014	2002-2016	-	1995-2012	2002-2015	(2000-2017) ^T
2.1 μ m	-	-	-	-	-	(2000-2017) ^M
3.7 μ m	1991-2002	1991-2002	1995-2012	-	1991-2002	(2000-2017) ^T

4.1.1 Cloud Fraction

In this subsection the cloud fraction of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS Collection 6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-1), zonal mean (Figure 4-2) and standard deviation (Figure 4-3), all for a common time period, and time series plots (Figure 4-4).

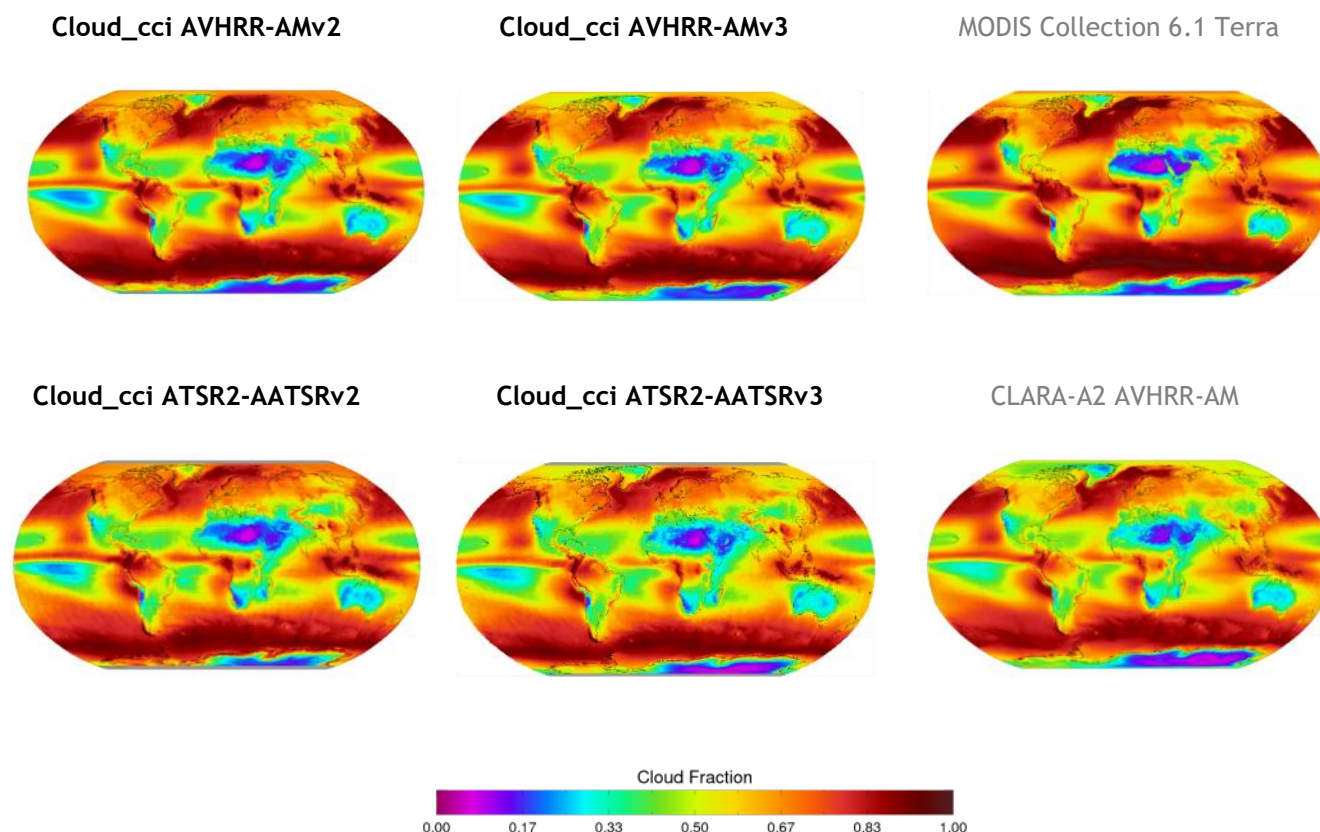


Figure 4-1 Globally gridded means of cloud fraction for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

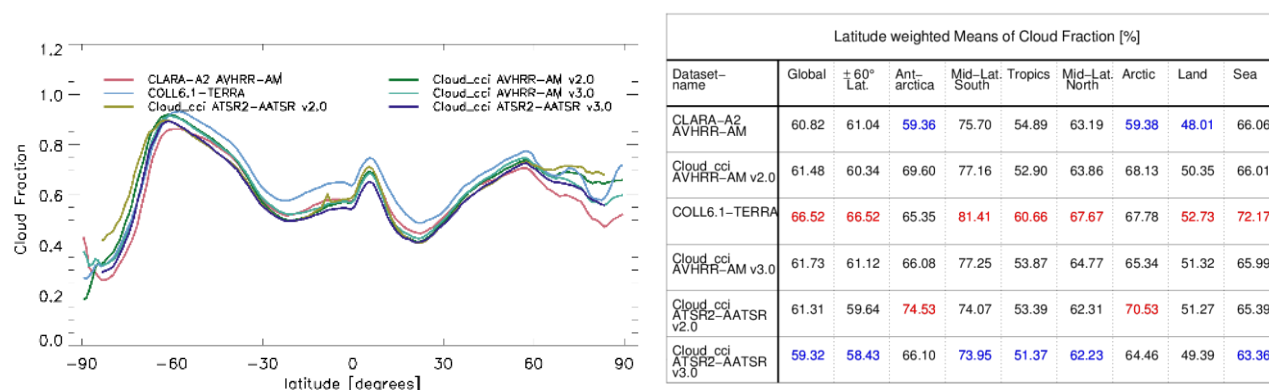



Figure 4-2 Cloud Fraction from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 31

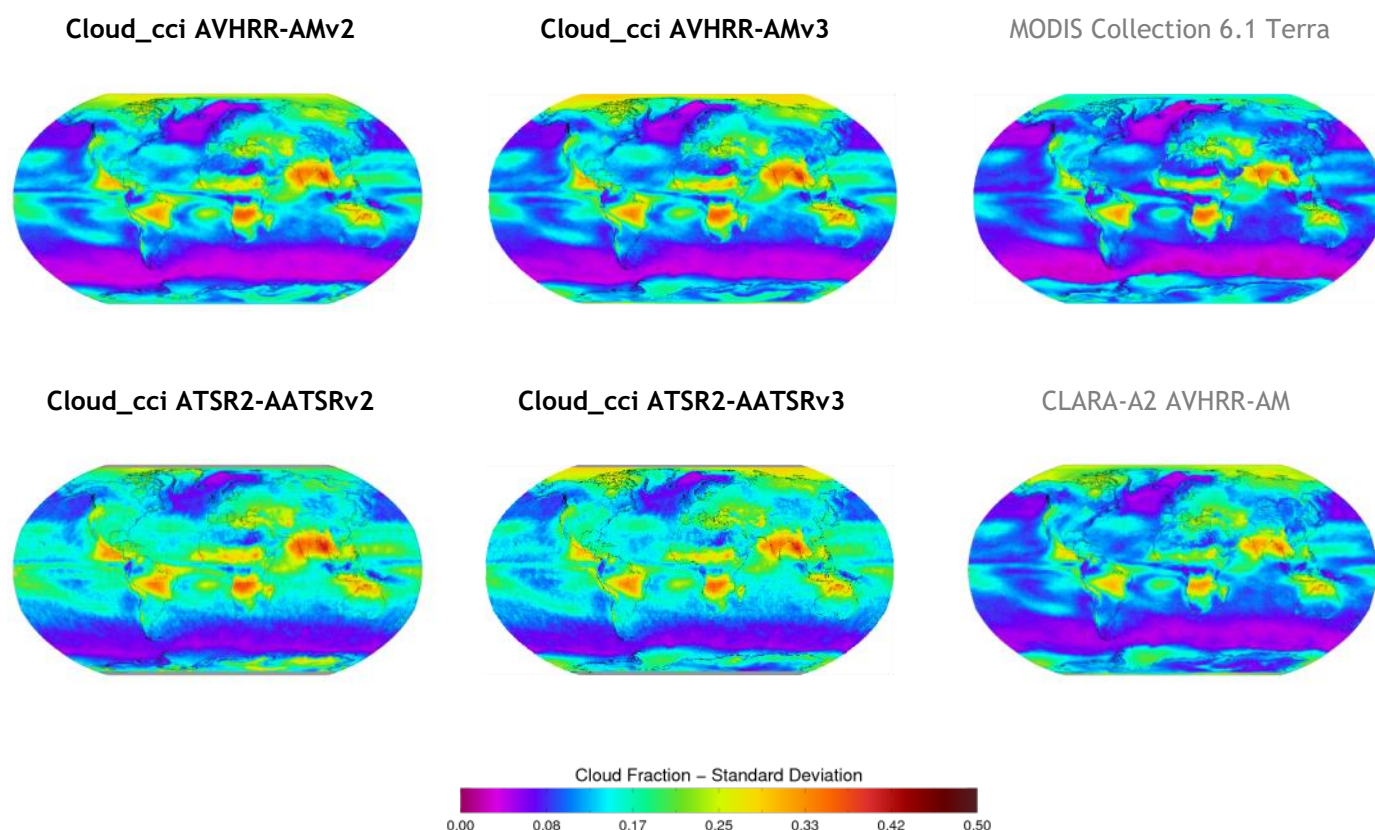


Figure 4-3 Globally gridded standard deviation of cloud fraction for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- Multi-annual averages of all datasets generally compare well with each other; regionally higher disagreement is visible.
- Best agreement in Mid-Latitudes and Tropics with all datasets being with approx. 10% cloud fraction. Exception is MODIS Collection 6.1 Terra dataset which has higher cloud fraction in the Tropics.
- Largest spread among datasets is found in polar regions, mainly arising from polar night conditions.
- Temporal variability is highest for Cloud_cci ATSR2-AATSRv3 and lowest for MODIS C6.1 Terra. Other datasets lie in between and seem relative consistent.
- Time series plots reveal jumps in the time series for (1) Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-AMv2 and CLARA-A2 at the transition from NOAA-12 to NOAA-15 around 1999 and (2) for ATSR2-AATSRv2 at the transition from ATSR2 (onboard ERS2) to AATSR (onboard ENVISAT) around 2003. In contrast, the time series of Cloud_cci ATSR2-AATSRv3 and MODIS Collection 6.1 Terra reveal a high stability over time.

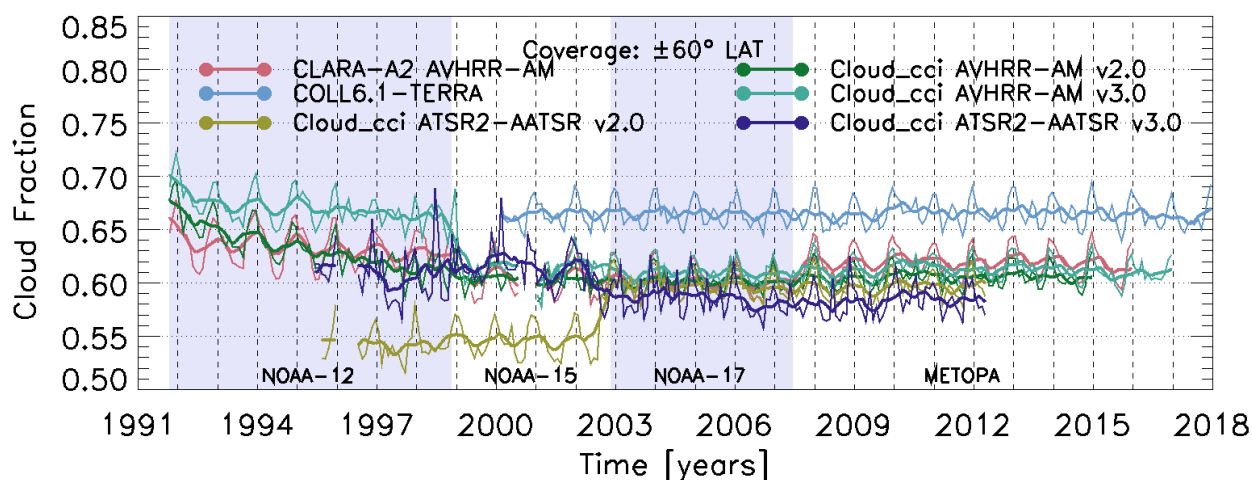


Figure 4-4 Time series of monthly cloud fraction for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

All Cloud_cci datasets are comparable to the reference datasets for nearly all parts of the globe. In general, MODIS C6.1 Terra gives highest CFC for large parts of the globe, which might be due to a combination of relative fine spatial resolution and using more spectral bands than Cloud_cci. Globally gridded maps and zonal mean plots exhibit a very high agreement among all datasets. The large spread in the polar regions might be due to polar night conditions in which no visible information is available and in which the surface temperature can be very low, both leading to only little or no contrast between signals from surfaces and clouds. This is a well-known problem for cloud detection based on passive imaging sensors.


In the time series of AVHRR-AM datasets, problems of the early AVHRRs (in particular NOAA-12) become evident. The NOAA-12 orbit is steadily drifting leading to a drift in local observation time, moving from late evening/twilight to night observations in case of the descending orbit node. As the cloud mask used in Cloud_cci AVHRR-AM has different branches for different illumination condition with significantly varying accuracies (twilight cloud detection is worst - large overestimation of cloudiness), the mean cloud fraction reduces with time as the observation time moves away from twilight condition. Important to note that this feature has worsened for AVHRR-AMv3 compared to v2.

Another point to mention is that an overestimation of cloudiness for CC4CL is found in condition of high aerosol loadings in the atmosphere. As discussed later for the afternoon satellites, the Pinatubo eruption in 1991 creates a significant positive anomaly in the cloud detection. This effect decreases as the aerosol loading is decreasing after some time after the eruption. Both, the reduction in twilight condition and the decreasing aerosol loadings after the Pinatubo outbreak result into negative trends in cloud fraction in the early years of Cloud_cci AVHRR-AMv2 and for v3 as well. CLARA-A2 shows similar behaviour with slightly decreased trend as CLARA-A2 cloud detection seems less impacted by the Pinatubo eruption. It should be noted that the gap in the Cloud_cci AVHRR-AMv2/v3 and CLARA-A2 datasets around the year 2000 is due to blacklisting of a large number of AVHRR-NOAA-15 orbits due to AVHRR scan motor errors.

After the year 2000 no trends are seen in all Cloud_cci datasets and Cloud_cci AVHRR-AMv3 agrees well with CLARA-A2. CLARA-A2 shows a positive jump of about 2 percentage points when switching from NOAA-17 to METOP-A. This is also seen in the Cloud_cci AVHRR-AM datasets but not that pronounced. Possible reasons for that can be a different equator crossing time of half an hour, slightly different viewing angle geometry, and potentially differences in recalibration accuracy.

The Cloud_cci ATSR2-AATSRv2 and especially v3 datasets are a bit lower than the AVHRRs. The reason for the differences between MODIS, AVHRR-AM and ATSR2-AATSR in Cloud_cci could be due to:

- 1) Passive sensors with higher viewing angles see more clouds: At higher viewing angles radiation has a longer path through the atmosphere, thus the impact of optically thin clouds on radiation is larger.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 33

- 2) The major differences between all sensors are their spectral response functions of the AVHRR heritage channels. In particular the differences of the visible channels are very high between AVHRR and ATSR2-AATSR. Within the CC4CL cloud detection all channels of ATSR2-AATSR are corrected to mimic AVHRR by using slope and offset found by comparing collocated measurements. For ATSR2-AATSR the offset of the reflectance channel correction is not applied to avoid a too strong increase for very low reflectance, which would lead to an increase in falsely detected clouds.

Another point worth mentioning is the transition from ATSR2 to AATSR, for which a significant jump of 5% for v2 is reduced for v3 (with a different sign). The ATSR2 part of that time series seems a little bit more unstable for v3 compared to v2.

Table 4-2 presents evaluation scores inferred from comparisons of Cloud_cci datasets with MODIS C6.1 Terra. Both Cloud_cci v3 datasets show negative CFC biases compared to MODIS C6.1 up to -8%. The bc-RMSD ranges from 6 to 10% with lower values for Cloud_cci AVHRR-AMv3 and higher values for Cloud_cci ATSR2-AATSRv3. The trend in 2003-2011 for is relative small for the Cloud_cci datasets and in relative good agreement to MODIS C6.1 Terra.

Table 4-2: Evaluation scores for Cloud_cci v3 Level-3C cloud fraction (morning satellites) based on comparison to C6.1 Terra from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM	ATSR2-AATSR
Bias [%]	-5.41	-8.09
bc-RMSD [%]	+6.81	+9.80
Cloud_cci v3.0 trend [% / decade]	+0.82	-0.35
C6.1-Terra trend [% / decade]	+0.17	+0.17
Δ trend [% / decade]	+0.65	-0.52

4.1.2 Cloud Top Pressure

In this subsection the cloud top pressure of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS Collection 6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-5), zonal mean (Figure 4-6) and standard deviation (Figure 4-7), all for a common time period, and time series plots (Figure 4-8).

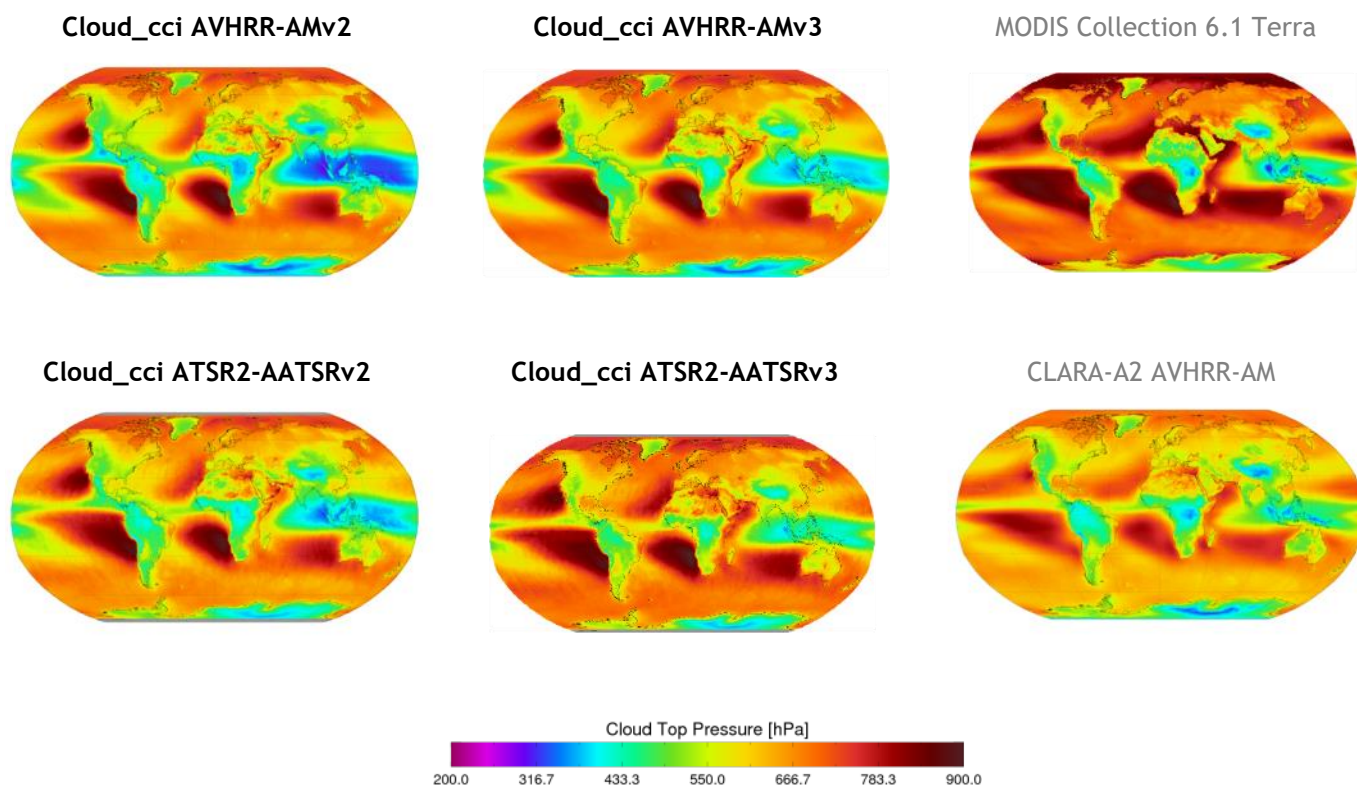


Figure 4-5 Globally gridded means of cloud top pressure for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

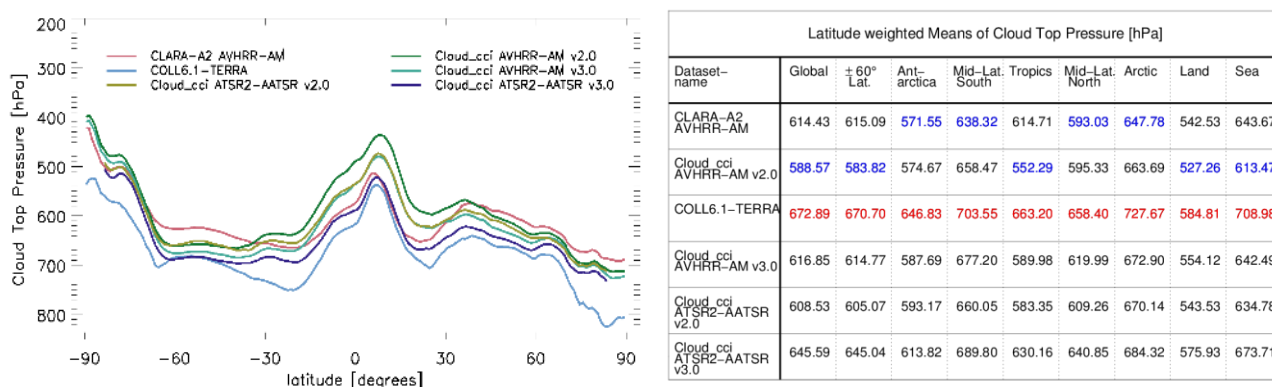



Figure 4-6 Cloud Top Pressure from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. **Highest** and **lowest** values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 35

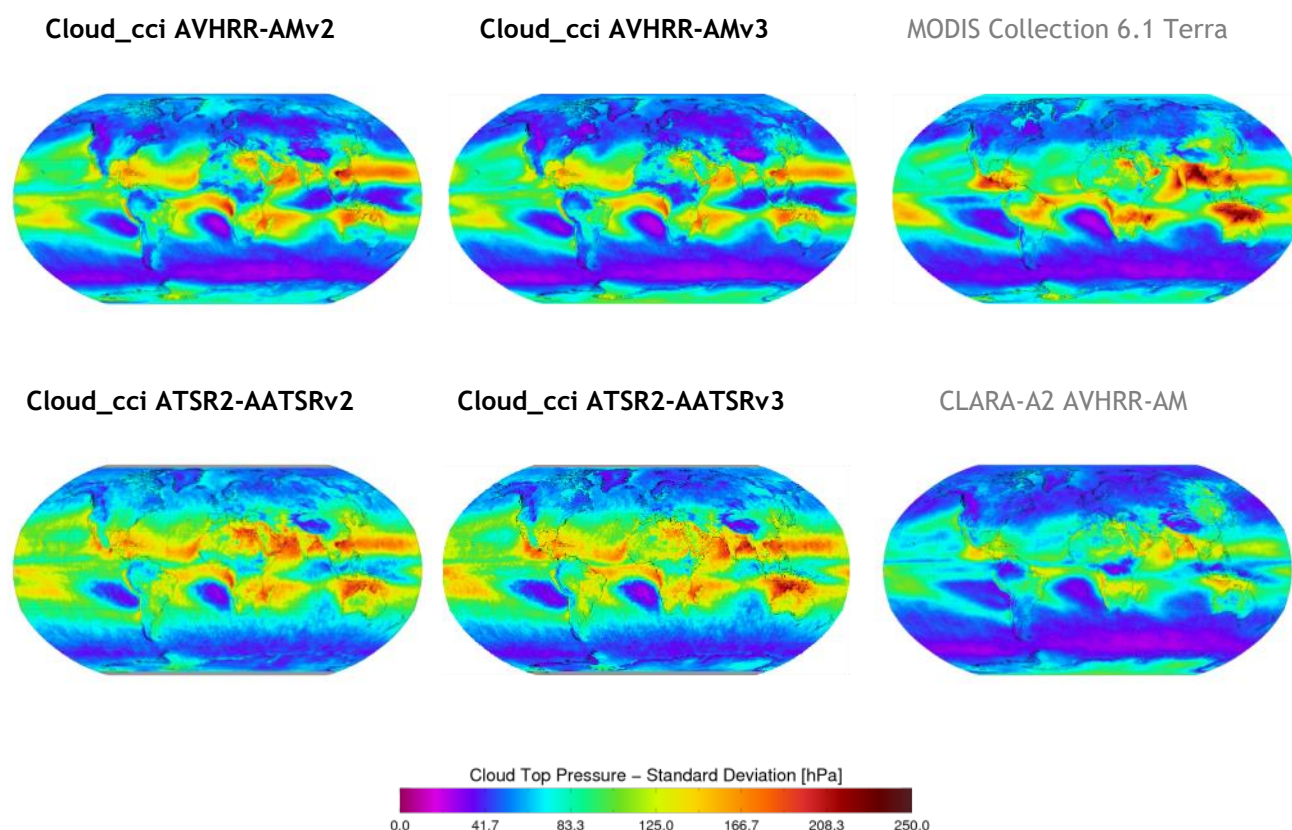


Figure 4-7 Globally gridded standard deviation of cloud top pressure for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- All datasets reflect the main characteristic of global cloud top distribution. Lowest mean CTP are found in the Tropic, highest mean CTPs in the stratocumulus regions. Mean mid-level CTPs are found for the storm track region in the Mid-Latitudes of both hemispheres.
- Lowest mean CTP is found for CLARA-A2, highest for MODIS C6.1 Terra especially in the tropics over sea. Despite the agreement on mean distributions, the spread among the datasets is partly large in nearly all regions of the globe.
- Largest temporal variability is found for the sub-tropic regions, which is reflected in all datasets (to a smaller or larger extent) and most likely connected to the seasonal oscillation of the ITCZ.
- The time series plots confirm the large spread in mean CTP among all datasets, while most datasets for themselves remain rather stable throughout the time period covered. An exception here is a jump found for CLARA-A2 at the transition from NOAA-12 to NOAA-15, which is also seen in AVHRR-PMv3 although with a different direction. This jump is not so pronounced in Cloud_cci AVHRR-AMv2. Cloud_cci ATSR2-AATSR v2 and v3 show some instability at the transition from ATSR2 to AATSR.

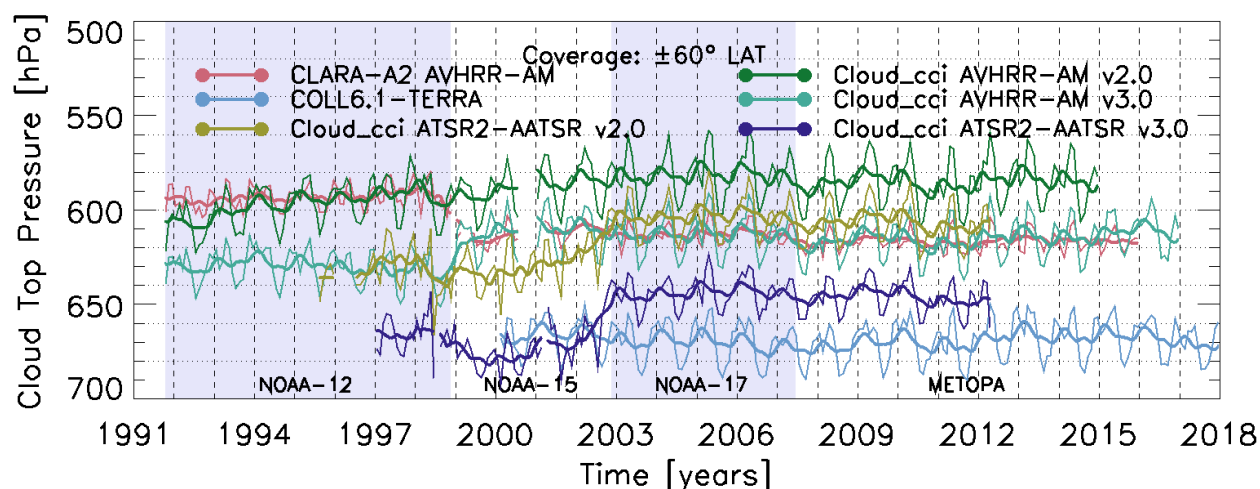


Figure 4-8 Time series of monthly cloud top pressure for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.


Discussion

Cloud_cci L3C CTP data show characteristic global patterns as seen in the other datasets. However, partly systematic deviations to the reference datasets are found. The most significant of the latter are the higher values for MODIS C6.1 Terra and the lower values CTP, which might be explained by more clouds being detected for this dataset (see higher cloud fraction for the MODIS C6.1 Terra dataset in Section 4.1.1) assuming these additional clouds are low, small scale clouds. In addition, CLARA-A2 AVHRR-AM shows a rather homogeneous spreading of CTP values, because the well-known stratocumulus fields are less pronounced than in the other datasets.

In Table 4-3 monthly 1-dimensional CTP histograms (also included in Cloud_cci and all other considered datasets) have been used to calculate the relative fraction of low, mid-level and high clouds on the total cloud amount using CTP thresholds of 440 hPa and 680 hPa according to the ISCCP definition. For most datasets we find a relative good agreement within 10 to 15%, but there are some outliers: MODIS C6.1 Terra has 58.06% low clouds while CLARA-A2 has only 35.30%; Very few mid-level clouds in MODIS C6.1 Terra of 9.62%, in contrast all others have values around 20 to 25%. For high clouds Cloud_cci ATSR2-AATSRv3 has the lowest amounts of 23.52% while CLARA-A2 has the most with 39.95%.

Table 4-3 Relative fraction of low ($CTP > 680$ hPa), mid-level (680 hPa $> CTP > 440$ hPa) and high ($CTP < 440$ hPa) clouds calculated from 1-dimensional histograms for a common time period of 2003 to 2011 (excluding the polar regions higher/lower than $\pm 60^\circ$ latitude). Highest values are in red, while lowest values are shown blue.

Dataset	Low clouds	Mid-level clouds	High clouds
CLARA-A2 AVHRR-AM	35.30%	24.75%	39.95%
Cloud_cci AVHRR-AMv2	45.51%	20.45%	34.04%
MODIS Collection 6.1 TERRA	58.06%	9.62%	32.32%
Cloud_cci AVHRR-AMv3	47.13%	24.07%	28.80%
Cloud_cci ATSR2-AATSRv2	46.71%	22.55%	30.74%
Cloud_cci ATSR2-AATSRv3	51.84%	24.64%	23.52%

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 37

A general conclusion is that MODIS C6.1 Terra includes many more small-scale clouds over ocean with high pressure values and the utilization of more than just the AVHRR-heritage channel may also lead to a better vertical placement of thin, high clouds in multi-layer cloud situation, while for AVHRR-heritage channel (without a absorption channel) clouds are often placed in the middle of two layers leading to more mid-level clouds.

In terms of stability, all Cloud_cci datasets seem relative stable, except small trends existing for Cloud_cci AVHRR-AMv2 and CLARA-A2 before 2001 due to the drift of NOAA-12. Cloud_cci AVHRR-AMv3 as well as CLARA-A2 contain significant jumps between NOAA-12 and NOAA-15, for CLARA-A2 towards higher pressure values, for Cloud_cci AVHRR-AMv3 towards lower values. Cloud_cci ATSR2-AATSRv2 and v3 also contain small trends before 2003 due to some remaining problems with ATSR2.

Table 4-4 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. Biases are below 60 hPa for both Cloud_cci v3 datasets and bc-RMSD is between 80 and 90 hpa. There is a slight positive trend in both Cloud_cci datasets. MODIS C6.1 Terra also shows a positive trend about 3.6 hPa per decade and is thus very close to the ATSR2-AATSR v3 trend.

Table 4-4: Evaluation scores for Cloud_cci v3 Level-3C Cloud Top Pressure (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM	ATSR2-AATSR
Bias [hPa]	-55.95	-25.52
bc-RMSD [hPa]	+80.81	+86.09
Cloud_cci v3.0 trend [hPa / decade]	+7.08	+3.99
C6.1-Terra trend [hPa / decade]	+3.59	+3.54
Δ trend [hPa / decade]	+3.49	+0.45

4.1.3 Cloud Optical Thickness - Liquid clouds

In this subsection the cloud optical thickness of liquid clouds (COT_{liq}) of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS C6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-9), zonal mean (Figure 4-10) and standard deviation (Figure 4-11), all for a common time period, and time series plots (Figure 4-12). The COTs of the products refer to slightly different wavelengths¹

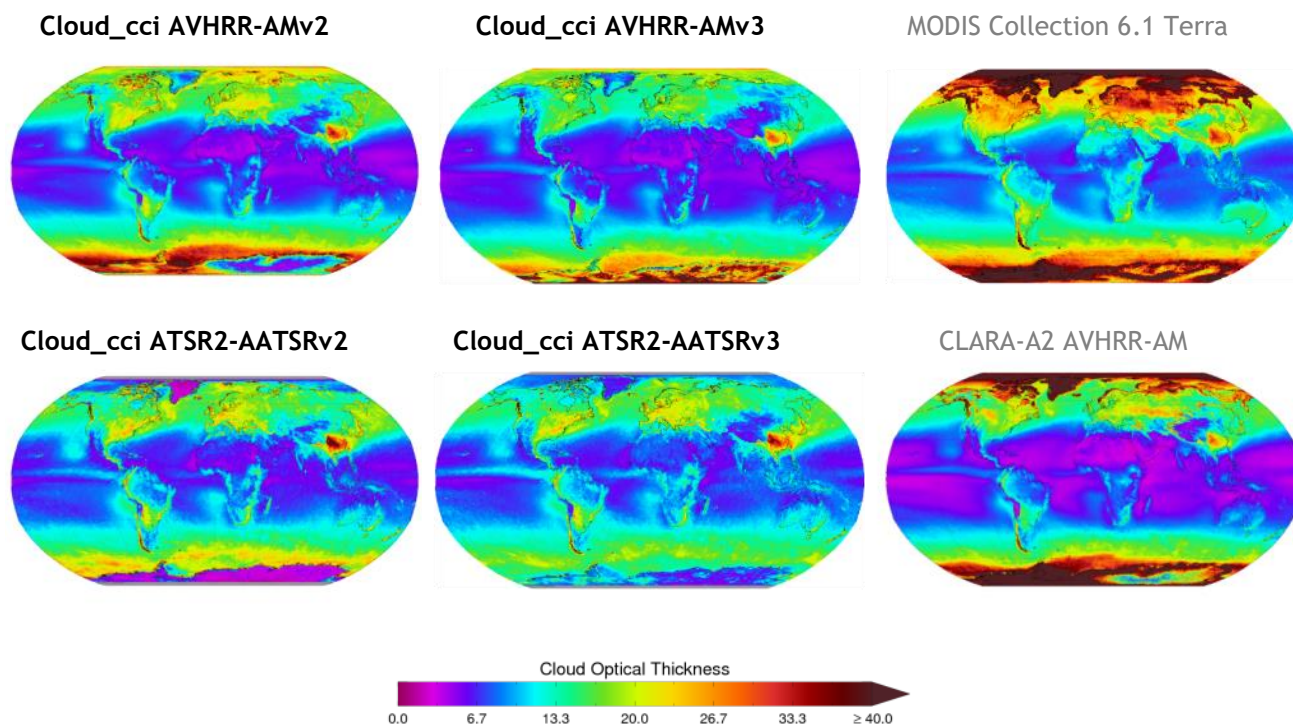


Figure 4-9 Globally gridded means liquid cloud optical thickness for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

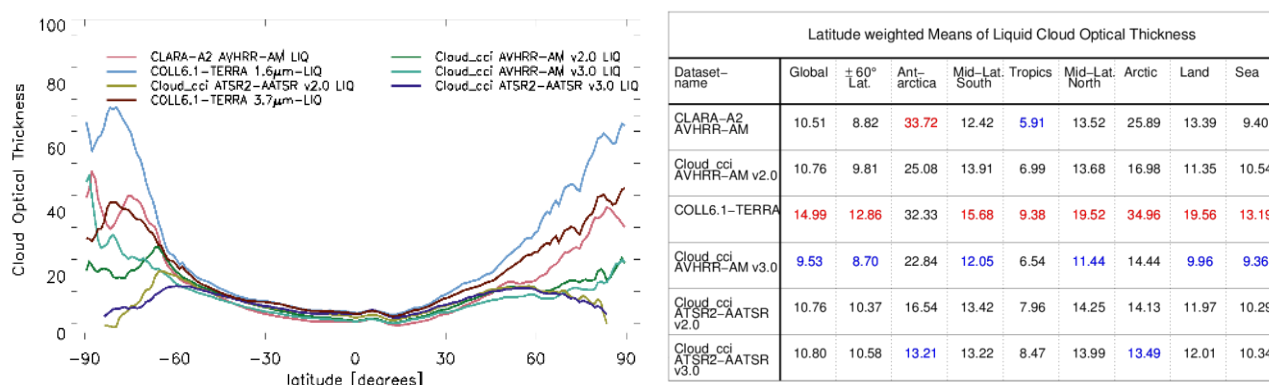



Figure 4-10 Liquid Cloud Optical Thickness from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6 Terra the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 39

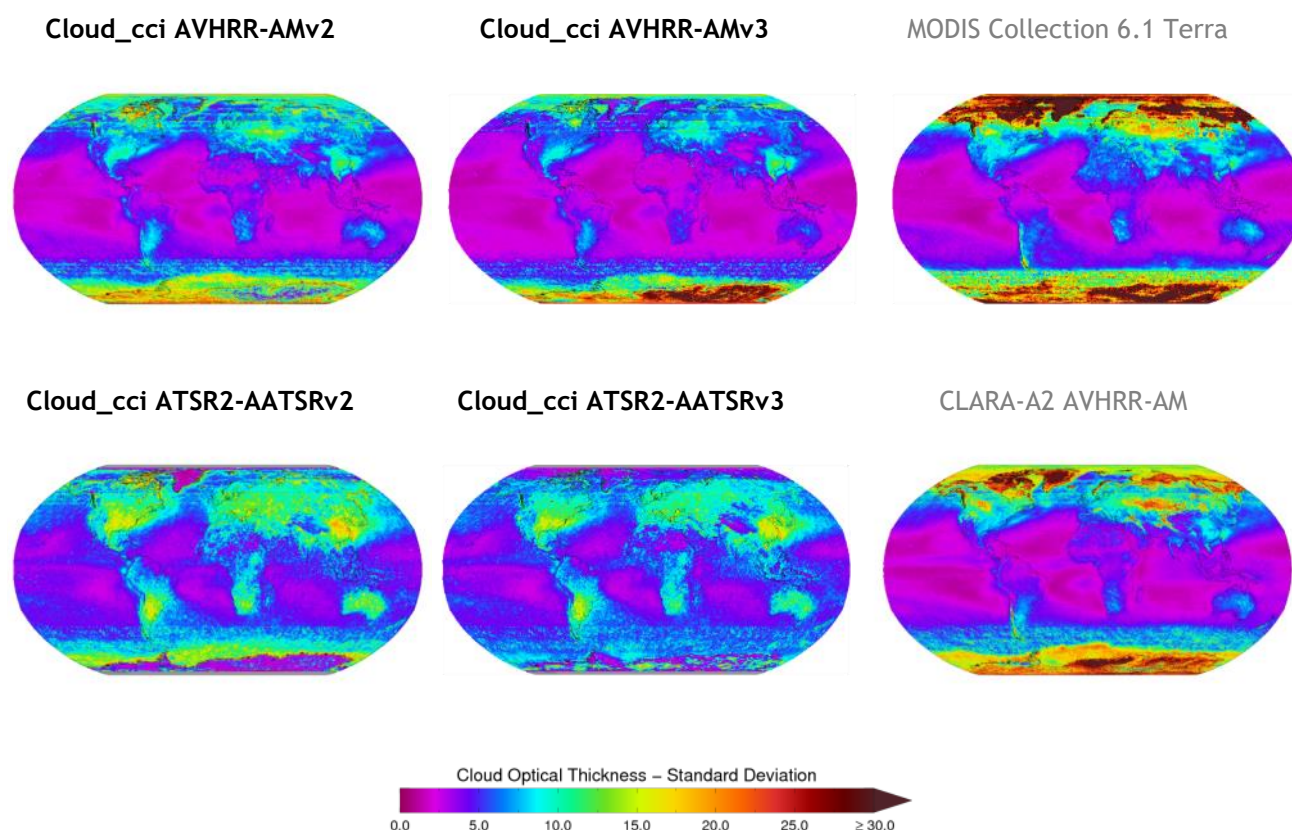


Figure 4-11 Globally gridded standard deviation of liquid cloud optical thickness for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

General findings

- For large parts of the globe the mean COT_{liq} looks similar among the datasets with highest values for MODIS C6.1 Terra.
- By far largest deviations are found for polar regions where MODIS C6.1 and CLARA-A2 have extremely high values. To a smaller extend this can also be seen for AVHRR-AM v2 and v3 over the Antarctic sea ice. The spread at the poles is nearly covering one order of magnitude.
- Lowest mean COT_{liq} are generally found in the subtropical regions with lowest values for CLARA-A2.
- In the polar regions, also the highest temporal variability is found, again most pronounced in MODIS C6.1 Terra and CLARA-A2. Outside the polar regions, largest temporal deviation is found over land regions in South-America (i.e. in ATSR2-AATSR datasets) and high Northern Latitudes (i.e. in MODIS C6.1 and CLARA-A2).
- Global mean values (between 60S and 60N) are generally between 5 and 15. Seasonal variations of these values are similar with higher values in boreal summer and lower in boreal winter.
- Time series plots show all datasets being relative stable in time. Jumps occur for Cloud_cci AVHRR-AMv2 (improved in AVHRR-AMv3) in 1999 (transition from NOAA12 to NOAA15) and for CLARA-A2 at the transitions from NOAA-15 to NOAA-17. Cloud_cci v3 datasets show high stability.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

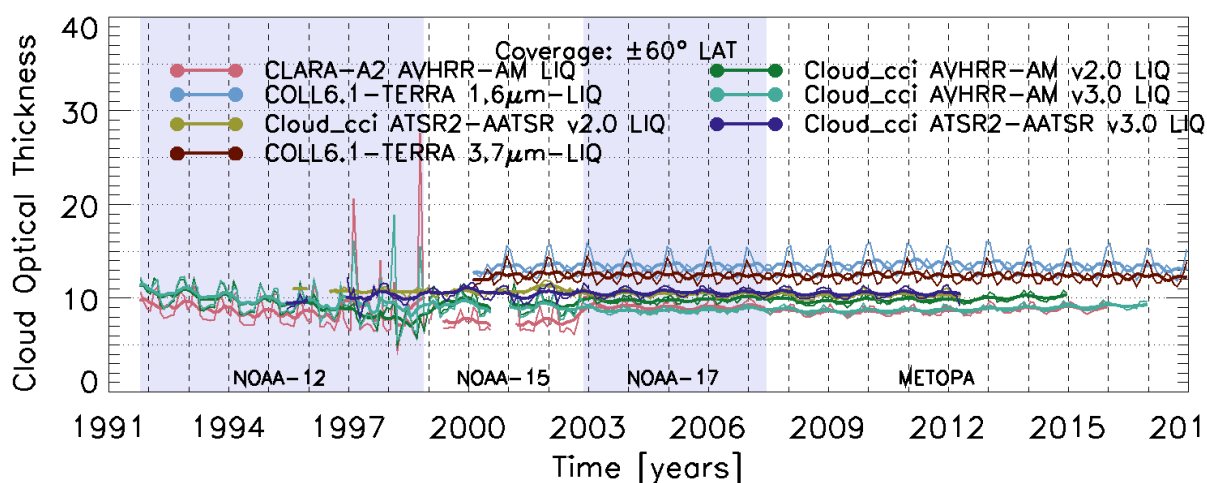


Figure 4-12 Time series of monthly liquid cloud optical thickness for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.

Discussion

For large parts of the globe the Cloud_cci COT_{liq} is in good agreement with the reference datasets. Mean COT_{liq} values around 10 in the tropics, slowly increasing with latitude to about 20 in the mid-latitudes, with only little spreads among the datasets and only small zonal variability. In contrast to that, in the high latitudes the spread becomes enormous, with Cloud_cci ATSR2-AATSRv3 and v2 being at the lower end and MODIS C6.1 Terra at the upper end and one order of magnitude between them. The very high COT_{liq} values in the polar regions seem not realistic since there is no reason to believe that clouds on the polar regions should be specifically optically thick. A retrieval artefact due to improper handling of snow and ice covered surfaces which usually causes a high reflectance in the visible seems more likely.

In case of the Cloud_cci v2 datasets for which CC4CL was used a related bug has been identified (wrong treatment of the sun zenith angle in the calculation of the BRDF over snow and ice surfaces) after processing. The larger impact of this for Cloud_cci AVHRR-AMv2 compared to Cloud_cci ATSR2-AATSRv2 is likely due to the observation time differences and the narrower swath of ATSR2-AATSR, which could lead to a smaller number of conditions with high solar zenith angles. Also, wrongly detected clouds over snow and ice surfaces might play a role here. For the Cloud_cci v3 datasets the results have improved and show more realistic values. However, all Cloud_cci datasets seem relative stable in time, except a small decreasing trend for Cloud_cci AVHRR-AMv3 and v2 for NOAA-12, thus before 1999, leading also to a small jump at the transition to NOAA-15. Reasons for that are likely to be the trend in cloud fraction seen before and the fact that drifting away from daytime leads to fewer daytime pixels for which COT_{liq} could be retrieved.

Table 4-5 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. The biases with respect to MODIS C6.1 Terra are low (between -5 and -2 optical thickness, with slightly lower mean values for Cloud_cci ATSR2-AATSRv3). The bc-RMSD is very similar for the Cloud_cci v3 datasets (around 7 optical thickness). The stability of Cloud_cci datasets is very high when compared against MODIS C6.1.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 41

Table 4-5: Evaluation scores for Cloud_cci v3 Level-3C Optical Thickness - Liquid (morning satellites) based on comparisons to MODIS C6.1 Terra from 2003-2011. All datasets use the 1.6 μ m channel. The scores were calculated separately for each Cloud_cci v3 dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias	-4.43	-2.40
bc-RMSD	+7.04	+7.81
Cloud_cci v3.0 trend [/ decade]	+0.07	-0.01
C6.1-Terra trend [/ decade]	+0.22	+0.38
Δ trend [/ decade]	-0.14	-0.40

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

4.1.4 Cloud Optical Thickness - Ice clouds

In this subsection the cloud optical thickness of ice clouds (COT_{ice}) of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS C6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-13), zonal mean (Figure 4-14) and standard deviation (Figure 4-15), all for a common time period, and time series plots (Figure 4-16). The COTs of the products refer to slightly different wavelengths¹.

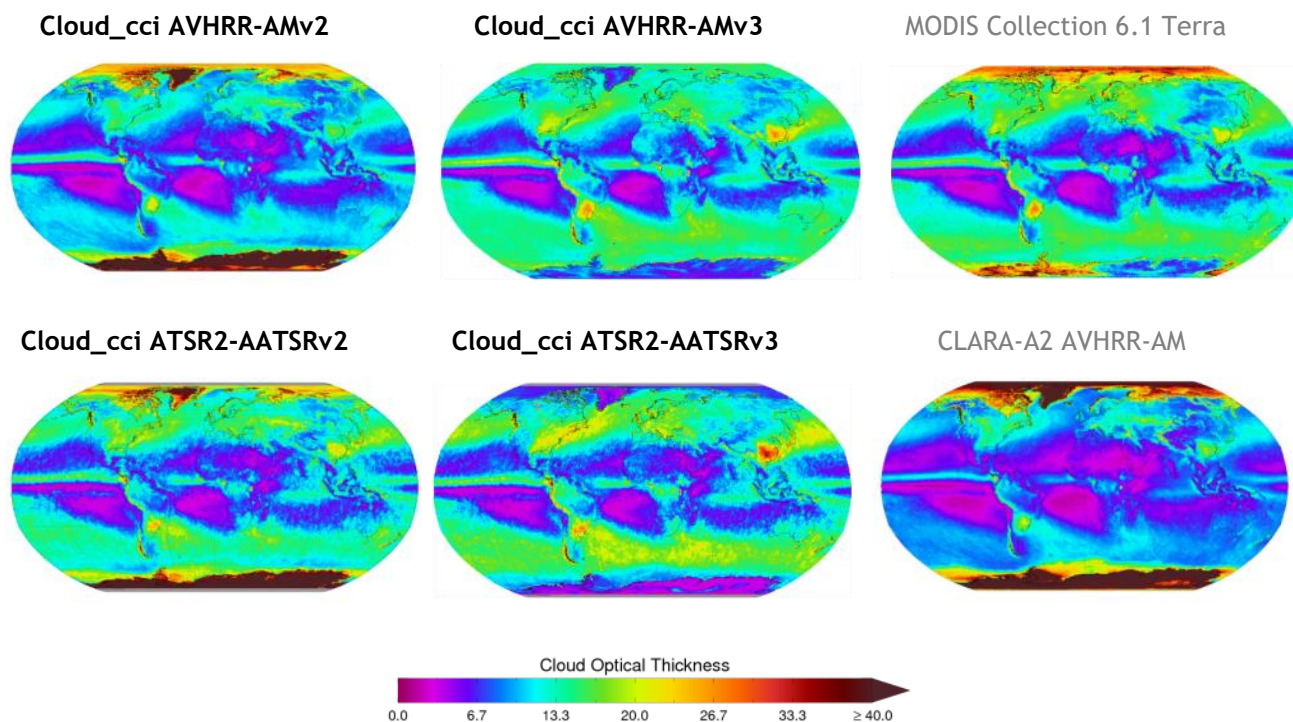


Figure 4-13 Globally gridded means of ice cloud optical thickness for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

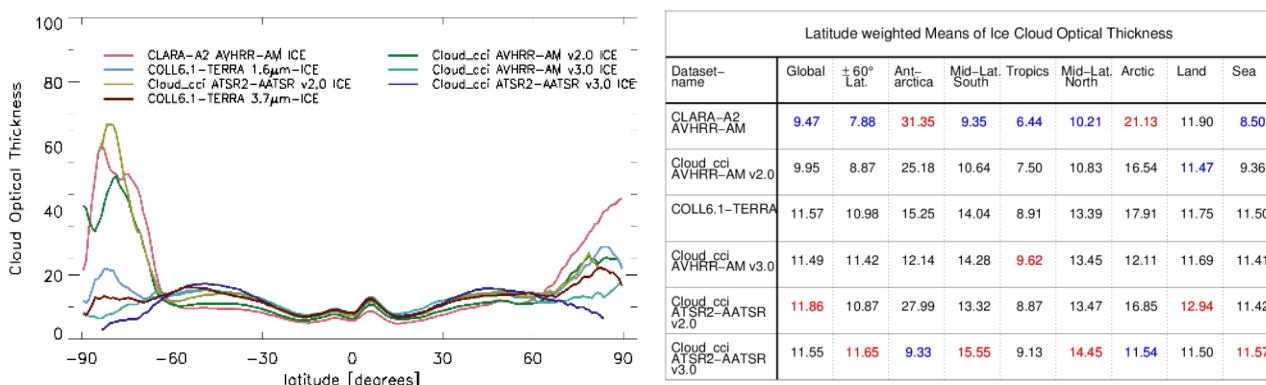



Figure 4-14 Ice Cloud Optical Thickness from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Terra the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 43

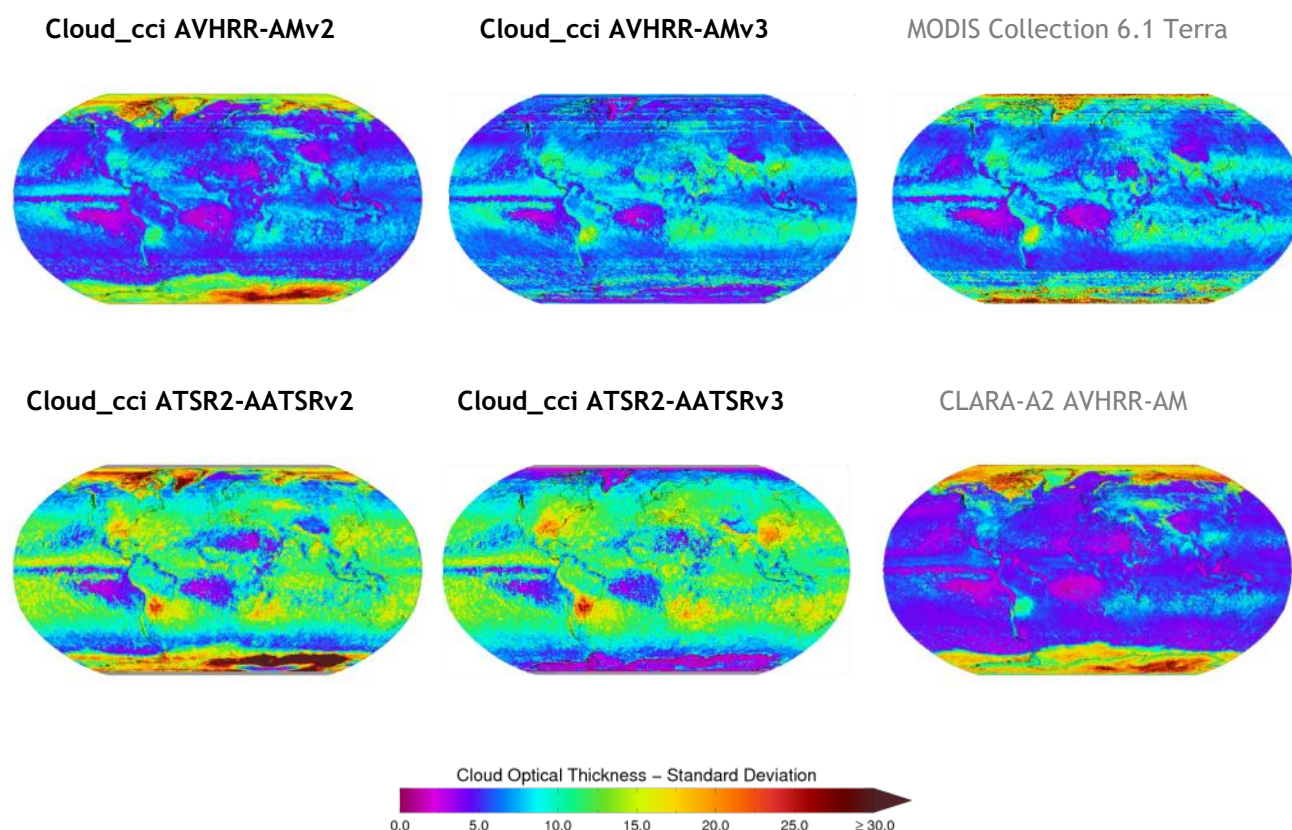


Figure 4-15 Globally gridded standard deviation of ice cloud optical thickness for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

General findings

- Comparable to COT_{liq} , the mean COT_{ice} is highest in polar regions for Cloud_cci v2 datasets, CLARA-A2 and MODIS Collection 6.1 Terra in particular in the Antarctic (above 30 optical thicknesses) with an exception for both Cloud_cci v3 datasets. Lowest values are found in the subtropical regions, i.e. the stratocumulus regions, with mean values below 5. The general global distribution patterns are similar for all datasets.
- The temporal variability is highest for ATSR2-AATSRv2 and v3, which seems to be two times as high as for the other datasets.
- The time series plots show a relatively stable series for all datasets beyond 2003. Before 2003 the Cloud_cci AVHRR-AMv3, AVHRR-AMv2 and CLARA-A2 series suffer from difficulties in correctly treating the early morning satellites NOAA-12 and NOAA-15.
- The mid-morning satellites of AVHRR-AM v2 and v3 that are using the 1.6 μ m channel instead of the 3.7 μ m (NOAA-17 and Metop-A) show a very stable time series with low variability.
- Cloud_cci AVHRR-AMv3 show a higher stability than the Cloud_cci v2 datasets, especially for the transition from NOAA-15 to NOAA-17, respectively. For ATSR2-AATSRv3 the jump between ATSR2 and AATSR is of different sign compared to v2 with about the same amplitude.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

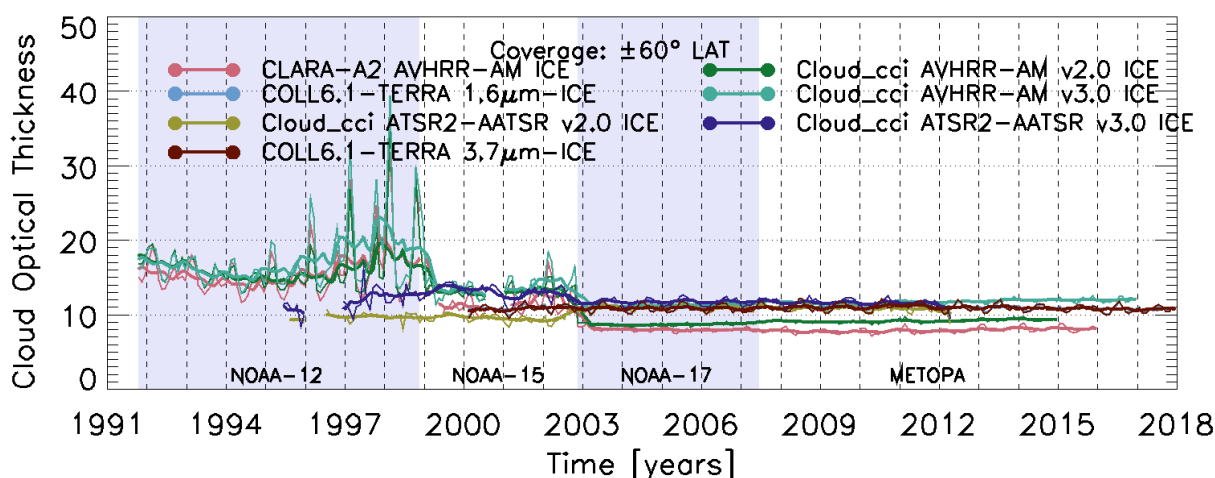


Figure 4-16 Time series of monthly ice cloud optical thickness for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6µm and 3.7µm products are shown.

Discussion

As seen for COT_{liq} , in the 60S-60N part of the globe the COT_{ice} of the Cloud_cci datasets is very similar among themselves (slightly higher in v3 compared to v2) and in comparisons to the reference datasets. All of them represent the local COT_{ice} maximum in the ITCZ, the minimum in the subtropics and show slightly increasing mean COT_{ice} values with increasing latitude, with Cloud_cci v2 datasets and CLARA-A2 showing highest values towards the poles.

The spread among the datasets in the polar regions is most likely due to the presence of snow/ice surfaces and the correspondingly difficult cloud detection, which could lead to pixel that contain snow/ice covered surfaces being erroneously classified as cloudy. In these cases, or in cases for which an optically thin cloud overlays a snow/ice surface, the retrieval schemes might attribute much of the measured reflection to the cloud optical thickness. In case of Cloud_cci v2 datasets the data was affected by a bug in the calculation of the BRDF components over snow and ice surfaces and high solar zenith angle condition, which led to a significant overestimation of COT_{ice} at these conditions. This bug is fixed in Cloud_cci AVHRR-AMv3 and Cloud_cci ATSR2-AATSRv3, clearly showing more realistic results.

In general, the early-morning orbit is very close to twilight condition, thus solar zenith angles are very high, which makes the retrieval of optical cloud properties, e.g. optical thickness, extremely difficult. Differences to the liquid phase can be found in the behaviour of NOAA-12 which is not showing the strong negative trend as seen in the liquid phase.

Table 4-6 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. The biases of COT_{ice} are relatively low and similar with value of +0.53 (Cloud_cci AVHRR-AMv3) and +0.58 (Cloud_cci ATSR2-AATSRv3). The bc-RMSD for Cloud_cci AVHRR-AMv3 is with approx. 5 optical thicknesses relative low; while the bc-RMSD for Cloud_cci ATSR2-AATSRv3 is twice as large. Decadal trends remain below 1 optical thickness for all datasets - with highest values for Cloud_cci AVHRR-AMv3 (0.63 / decade) which is close to the MODIS C6.1 figures, while ATSR2-AATSRv3 shows nearly no trend.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55µm, CLARA-A2: 0.6µm, MODIS C6.1: 0.65µm (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7µm on COT is assumed to be very minor.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 45

Table 4-6: Evaluation scores for Cloud_cci v3 Level-3C Optical Thickness - Ice (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. All datasets use the 1.6 μ m channel. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias	+0.53	+0.58
bc-RMSD	+4.84	+9.60
Cloud_cci v3.0 trend [/ decade]	+0.63	-0.02
C6.1-Terra trend [/ decade]	+0.41	+0.45
Δ trend [/ decade]	+0.21	-0.47

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1: 0.65 μ m (assumed to be the same as for C5), ISCCP (unknown). The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

4.1.5 Cloud Effective Radius - Liquid clouds

In this subsection the cloud effective radius of liquid clouds (CER_{liq}) of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS Collection 6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-17), zonal mean (Figure 4-18) and standard deviation (Figure 4-19), all for a common time period, and time series plots (Figure 4-20).

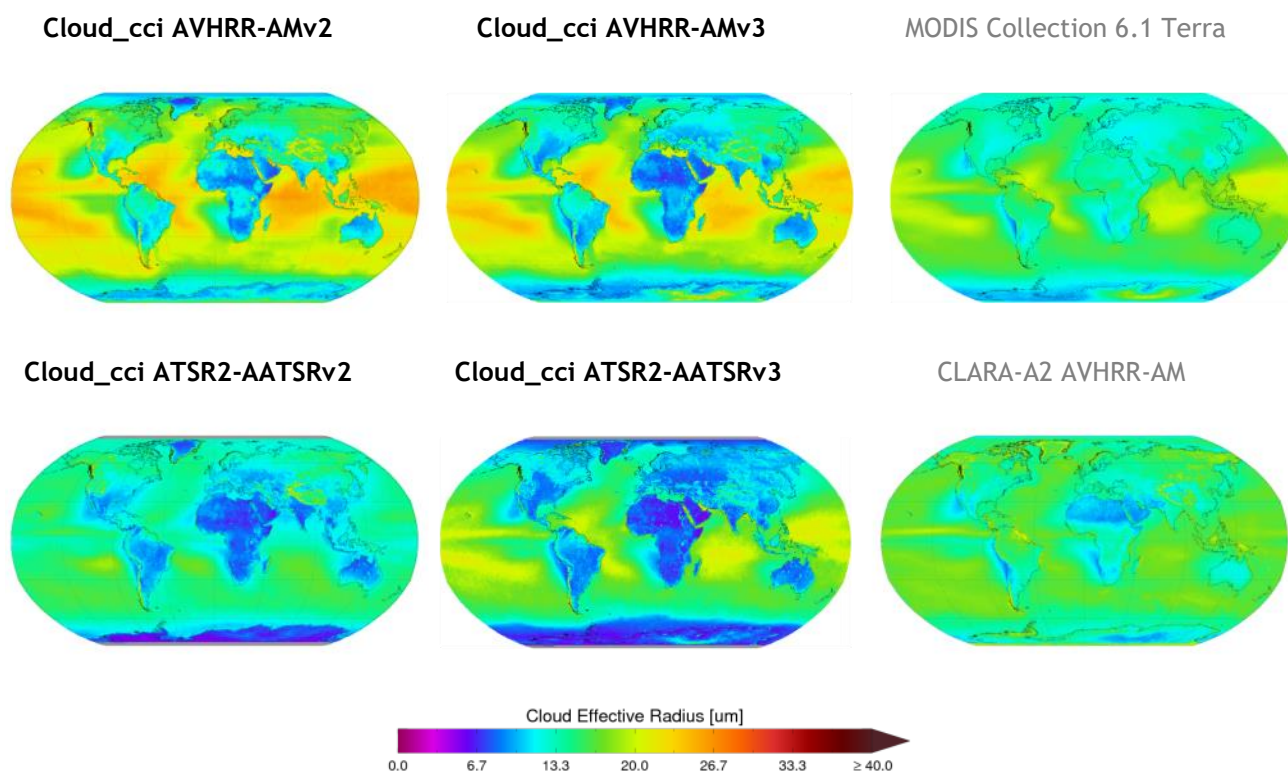


Figure 4-17 Globally gridded means of liquid cloud effective radius for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1μm product is plotted.

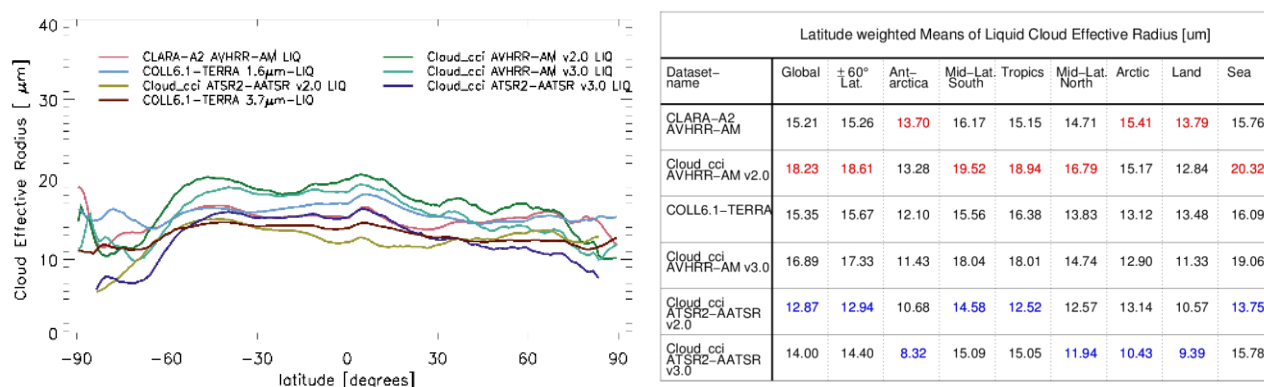



Figure 4-18 Liquid Cloud Effective Radius from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the 1.6μm and 3.7μm products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Terra the 2.1μm product is used. **Highest** and **lowest** values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 47

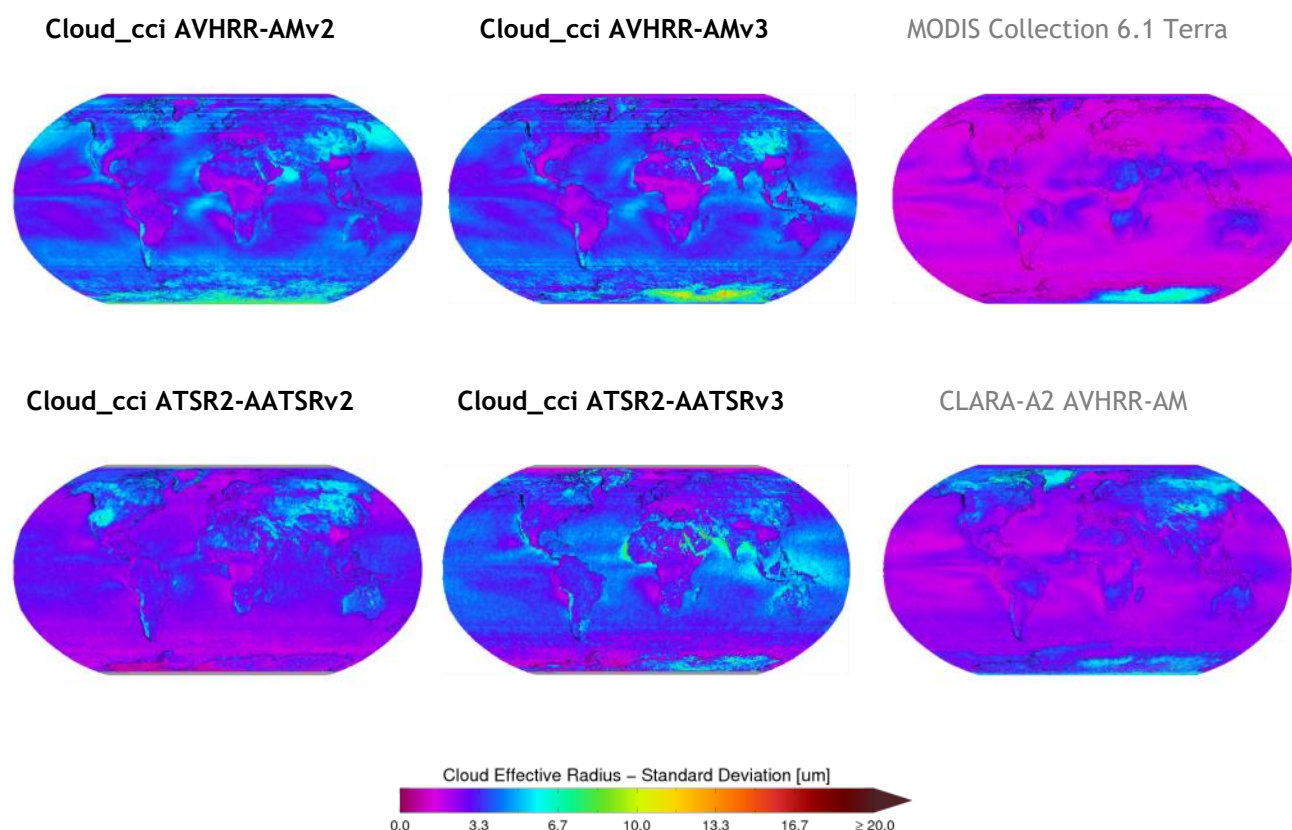


Figure 4-19 Globally gridded standard deviation of liquid cloud effective radius for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the $2.1\mu\text{m}$ product is plotted.

General findings

- Mean CER_{liq} are lowest over land and highest over tropical and subtropical oceans in all datasets; however some deviations are found in the actual mean values, which are highest in Cloud_cci AVHRR-AM datasets and lowest in Cloud_cci ATSR2-AATSR datasets, especially in Cloud_cci ATSR2-AATSRv2.
- In zonal mean plots it is seen that the mean CER_{liq} is between approx. 10 and $20\mu\text{m}$ for all latitude bands.
- The temporal variability is very low, with slightly higher values in Cloud_cci AVHRR-AMv3, AVHRR-AMv2 and Cloud_cci ATSR2-AATSRv3 compared to the other datasets.
- All datasets are stable beyond 2003. As for Cloud_cci AVHRR-AMv2, AVHRR-AMv3 reveals jumps in the time series of the global mean (60S-60N) values of CER_{liq} between the satellites NOAA-12 and NOAA-15. The jump from NOAA-15 to NOAA-17 is overlaid by a switch of the NIR channel available at daytime for the specific satellites. Which is $3.7\mu\text{m}$ for NOAA-12 and NOAA-15, and $1.6\mu\text{m}$ for NOAA-17 and Metop-A.
- A small jump is also visible at the transition of the two sensors in ATSR2-AATSRv2 dataset, which can similarly be found for ATSR2-AATSRv3.

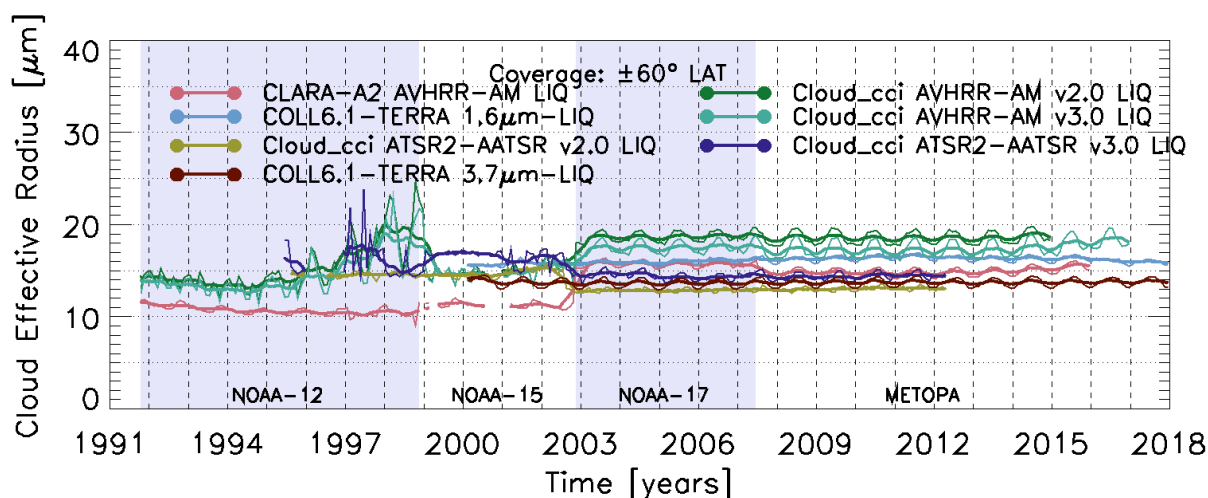


Figure 4-20 Time series of monthly liquid cloud effective radius for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μm and 3.7 μm products are shown.

Discussion

The Cloud_cci ATSR2-AATSRv3 and v2 datasets agree to the reference data for CER_{liq} for most parts of the globe. However, over land both ATSR2-AATSR datasets show values which are clearly lower than in the reference datasets CLARA-A2 and MODIS C6.1 Terra, which is also seen in AVHRR-PMv3 to some extent. Both datasets Cloud_cci AVHRR-AMv3 and v2 show values up to 5 μm higher nearly everywhere over the ocean. A common feature among all datasets is that smaller mean CER_{liq} exist for land than over ocean. One could speculate if this potentially speaks for a higher number of cloud droplets over land due to a potentially higher aerosol concentration. Additionally, all datasets agree on smaller mean CER_{liq} for stratocumulus regions, compared to for example the trade cumulus regions.

There is only little year-to-year variability of the 60S-60N mean values in all datasets beyond 2003. All datasets seem to have a similar seasonal cycle with a minimum in the boreal winter and a maximum in the boreal summer, the amplitudes being a bit diverse. The transitions from ATSR2 to AATSR and from NOAA-15 to NOAA-17 can be identified. For the AVHRR-AM based datasets this is mainly due to the switch of the available daytime NIR channel, which is 3.7 μm for NOAA-12 and NOAA-15 while 1.6 μm for NOAA-17 and Metop-A. As the signal at 1.6 μm can be assumed to come from lower parts of the cloud compared to measurements at 3.7 μm , we can infer information about the vertical profiles of CER_{liq} in liquid clouds. Cloud_cci AVHRR-AMv3 as well as Cloud_cci AVHRR-AMv2 (specifically looking at NOAA-15 and NOAA-17) agree well with CLARA-A2 and MODIS C6.1 Terra (1.6 and 3.7 μm products) with respect to showing the 1.6 μm CER_{liq} being about 5 μm larger than the 3.7 μm CER_{liq} on average for 60S-60N; in other words smaller cloud droplets at the cloud top compared to below.

Noticeable again is the problem with NOAA-12 increasingly shifting during its lifetime. However, as this is only seen for Cloud_cci AVHRR-AMv3 and v2 and not for CLARA-A2, we assume some shortcomings in the CC4CL retrieval either for CER, cloud detection or cloud phase determination.

Table 4-7 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. The biases of 1.17 μm for AVHRR-AMv3 and -1.8 μm for ATSR2-AATSRv3 are very small. The bc-RMSD is higher for Cloud_cci AVHRR-AMv3 with values of 4.36 μm than the bc-RMSD for Cloud_cci ATSR2-AATSRv3 dataset with 3.88 μm . Decadal trends of both Cloud_cci v3 datasets are very small negative while the MODIS C6.1 Terra trends for the 1.6 CER_{liq} are small positive with approximately 0.9 μm per decade.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 49

Table 4-7: Evaluation scores for Cloud_cci v3 Level-3C Effective Radius - Liquid (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. All datasets used the 1.6 μ m channel. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias [μ m]	+1.17	-1.80
bc-RMSD [μ m]	+4.36	+3.88
Cloud_cci v3.0 trend [μ m / decade]	-0.46	-0.09
C6.1-Terra trend [μ m / decade]	+0.87	+0.88
Δ trend [μ m / decade]	-1.33	-0.96

4.1.6 Cloud Effective Radius - Ice clouds

In this subsection the cloud effective radius of ice clouds (CER_{ice}) of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, MODIS Collection 6.1 Terra, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3 and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-21), zonal mean (Figure 4-22) and standard deviation (Figure 4-23), all for a common time period, and time series plots (Figure 4-24).

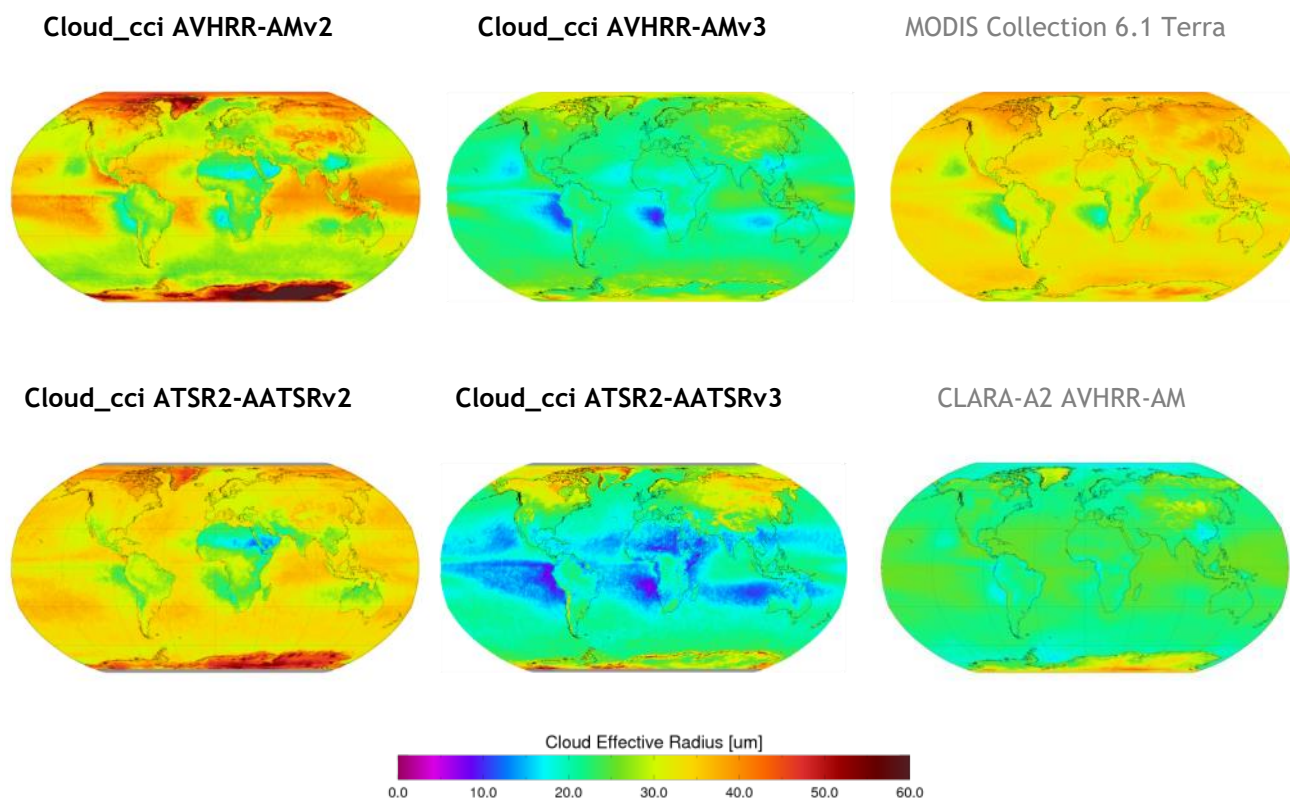


Figure 4-21 Globally gridded means of ice cloud effective radius for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the $2.1\mu m$ product is plotted.

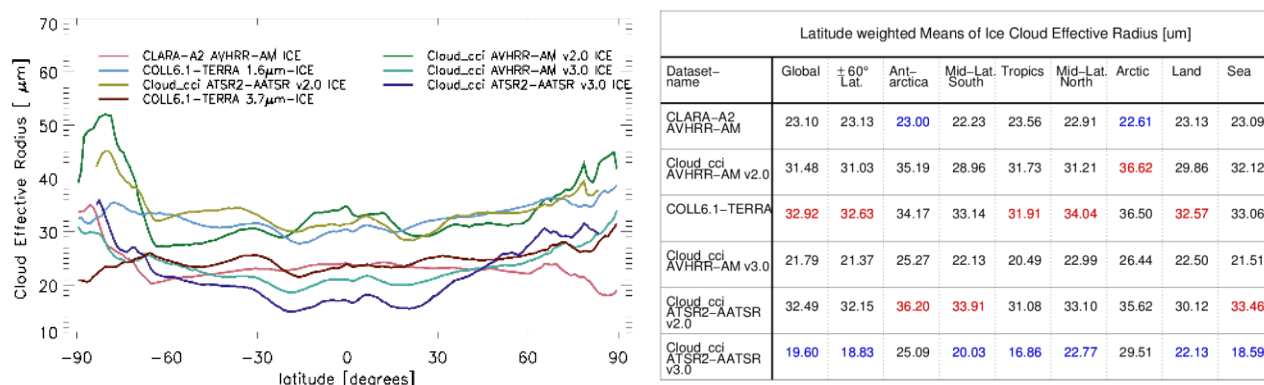



Figure 4-22 Ice Cloud Effective Radius from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the $1.6\mu m$ and $3.7\mu m$ products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Terra the $2.1\mu m$ product is used. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 51

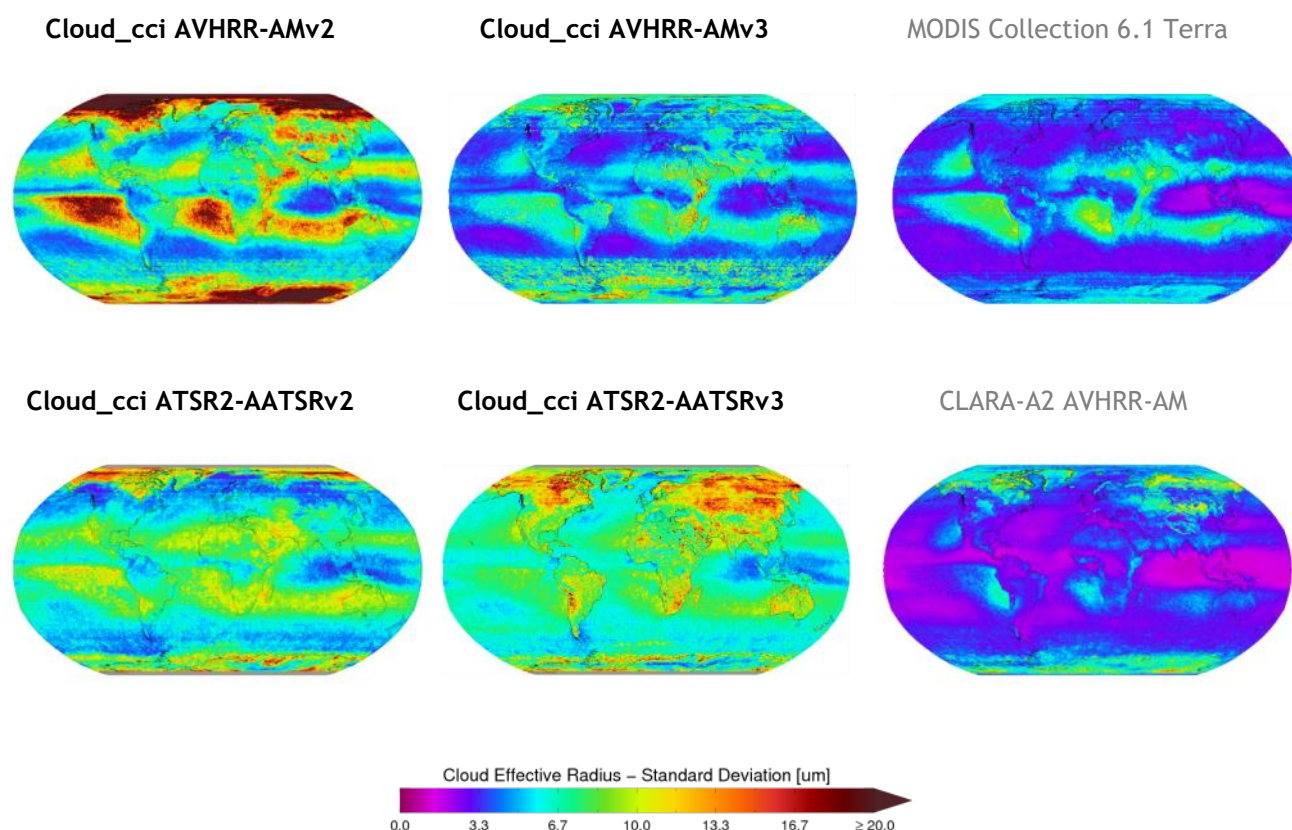


Figure 4-23 Globally gridded standard deviation of ice cloud effective radius for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

General findings

- Global maps of mean CER_{ice} show quite some systematic deviation between the Cloud_cci v2 datasets together with MODIS C6.1 Terra compared to Cloud_cci v3 datasets and CLARA-A2. The deviation is nearly a factor of two. However, major global pattern remains similar in all datasets with lower values in the stratocumulus regions and higher values in the trade-cumulus regions and Tropics. Clearly higher values are found over polar land regions in all Cloud_cci and reference datasets.
- In terms of temporal variability, all datasets show highest variability in subtropical and polar regions. However, the variability spread among the datasets is, similar to the mean values, partly deviating by a factor of two, with highest variability values for Cloud_cci AVHRR-AMv2 and Cloud_cci ATSR2-AATSRv3.
- Wrt. the long-term stability, one can again see the switch of available (and used) near-infrared channel from NOAA-15 to NOAA-17 (from 3.7 to 1.6 μ m) in Cloud_cci AVHRR-AMv2 and CLARA-A2, however, with different implications. Cloud_cci v3 datasets show an increased stability, also for satellite transitions and channel switches. ATSR2-AATSRv3 shows a jump from ATSR2 to AATSR as also visible for v2.

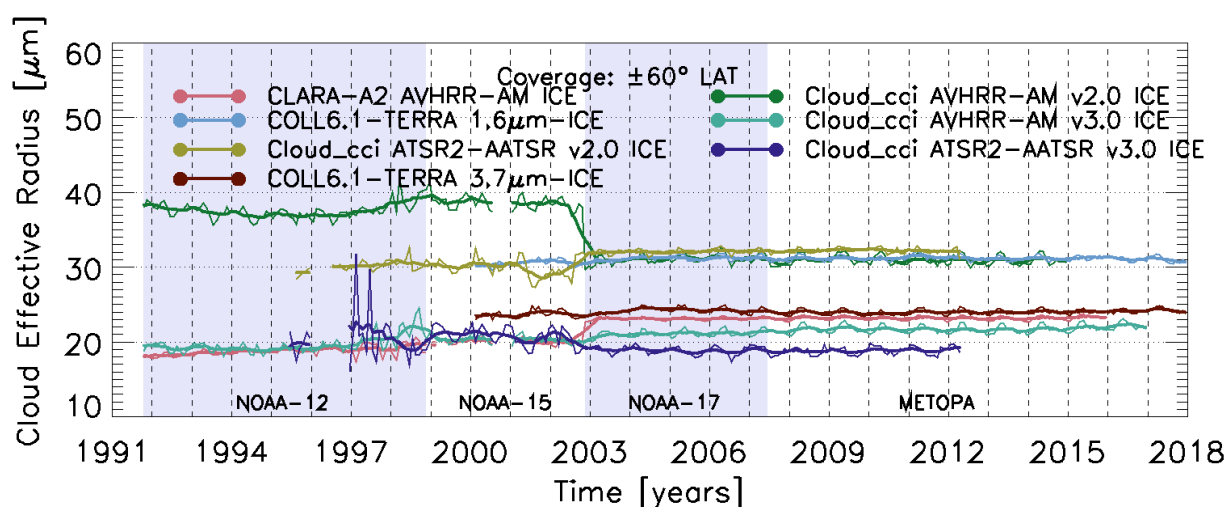


Figure 4-24 Time series of monthly ice cloud effective radius for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.


Discussion

The Cloud_cci v2 time series of CER_{ice} revealed a problem found in an older CC4CL 3.7 μ m ice retrieval. Especially the early AVHRR-AM CER_{ice} satellites were unreasonable high. After bug-fixing, Cloud_cci AVHRR-AMv3 and Cloud_cci ATSR2-AATSRv3 and the time series show significantly lower and stable values without any significant transition jumps.

V3 Cloud_cci time series are rather stable with nearly no year-to-year variability for the 60S-60N average nor long-term trend. A small season cycle is present, similar for all datasets, with a maximum in the boreal spring and a minimum at boreal autumn.

As for liquid clouds, the difference between 1.6 μ m and 3.7 μ m retrievals of CER_{ice} can give information about the vertical structure of particle sizes in ice clouds. Cloud_cci v3 AVHRR-AMv3 also reveals what is seen for CLARA-A2 and MODIS C6.1 Terra: (slightly) higher CER_{ice} for 1.6 μ m than for 3.7 μ m on average for 60S-60N; which in other words means smaller ice particles found at the cloud top than below. MODIS C6.1 lies somewhere in the middle MODIS C6.1 lies somewhere in the middle in this respect.

Table 4-8 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. Compared to liquid clouds, the agreement between Cloud_cci and MODIS C6.1 Terra datasets is less. For AVHRR-AMv3 as well as for ATSR2-AATSRv3 we find large biases against the corresponding MODIS C6.1 Terra product (smaller values in Cloud_cci). The bc-RMSD against MODIS C6.1 Terra is very similar for both Cloud_cci v3 datasets with values between 4.5 and 8 μ m. There is only little trend in the Cloud_cci ATSR2-

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 53

AATSRv3 dataset (below $-0.2\mu\text{m}$ / decade), while Cloud_cci AVHRR-AMv3 shows a positive trend about $1\mu\text{m}$ per decade. MODIS C6.1 lies somewhere in the middle in this respect.

Table 4-8: Evaluation scores for Cloud_cci v3 Level-3C Effective Radius - Ice (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. All datasets used the $1.6\mu\text{m}$ channel. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias [μm]	-9.79	-12.50
bc-RMSD [μm]	+4.89	+7.46
Cloud_cci v3.0 trend [μm / decade]	+1.06	-0.12
C6.1-Terra trend [μm / decade]	+0.25	+0.21
Δ trend [μm / decade]	+0.81	-0.33

4.1.7 Liquid Water Path

In this subsection the cloud liquid water path of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3, MODIS C6.1 Terra and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-25), zonal mean (Figure 4-26) and standard deviation (Figure 4-27), all for a common time period, and time series plots (Figure 4-28 and Figure 4-29).

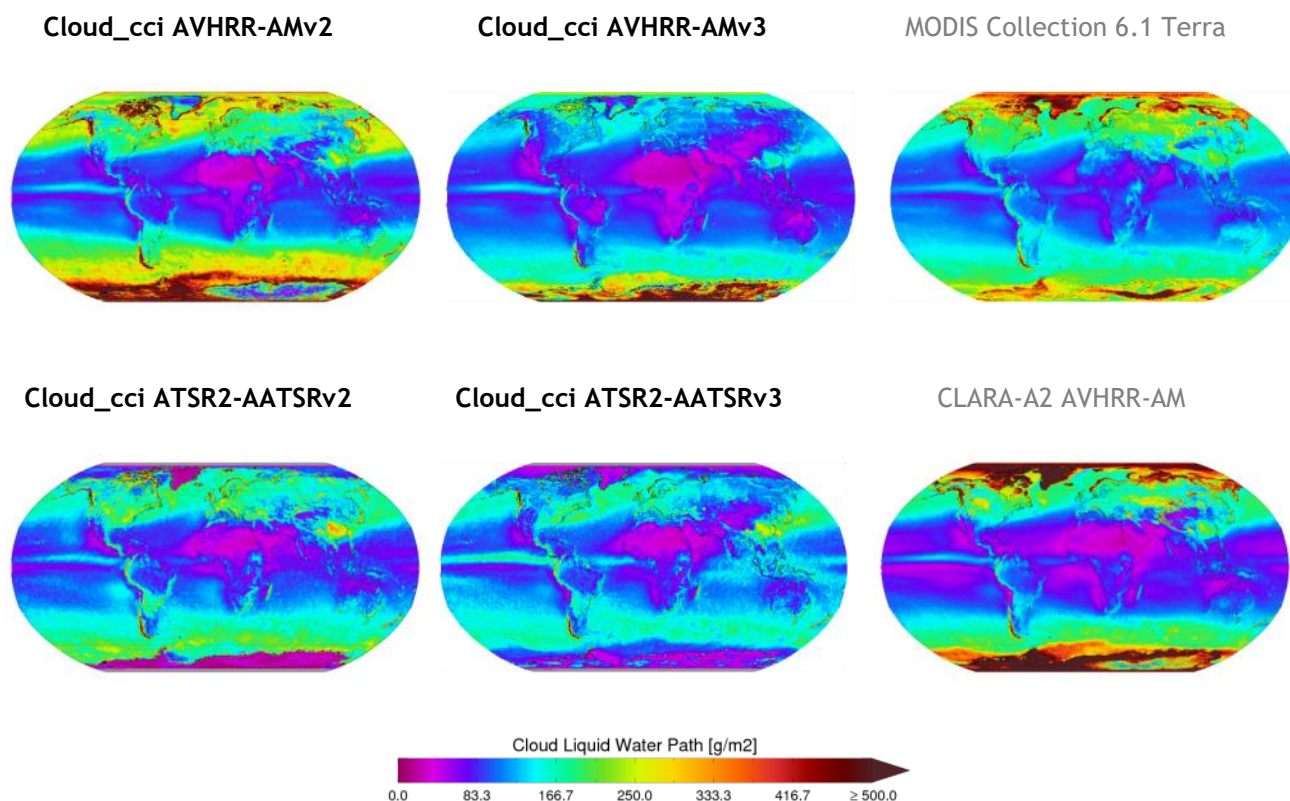


Figure 4-25 Globally gridded means of cloud liquid water path for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

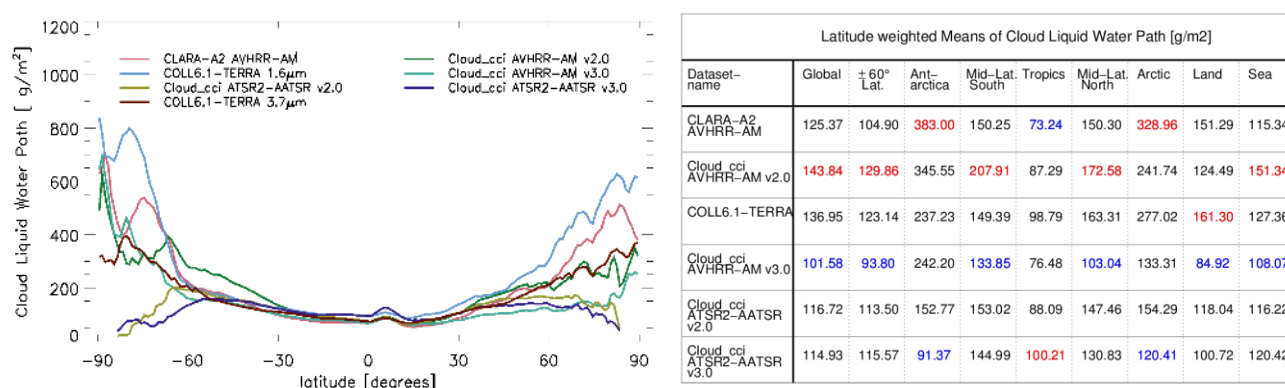



Figure 4-26 Cloud Liquid Water Path from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS C6.1 Terra the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 55

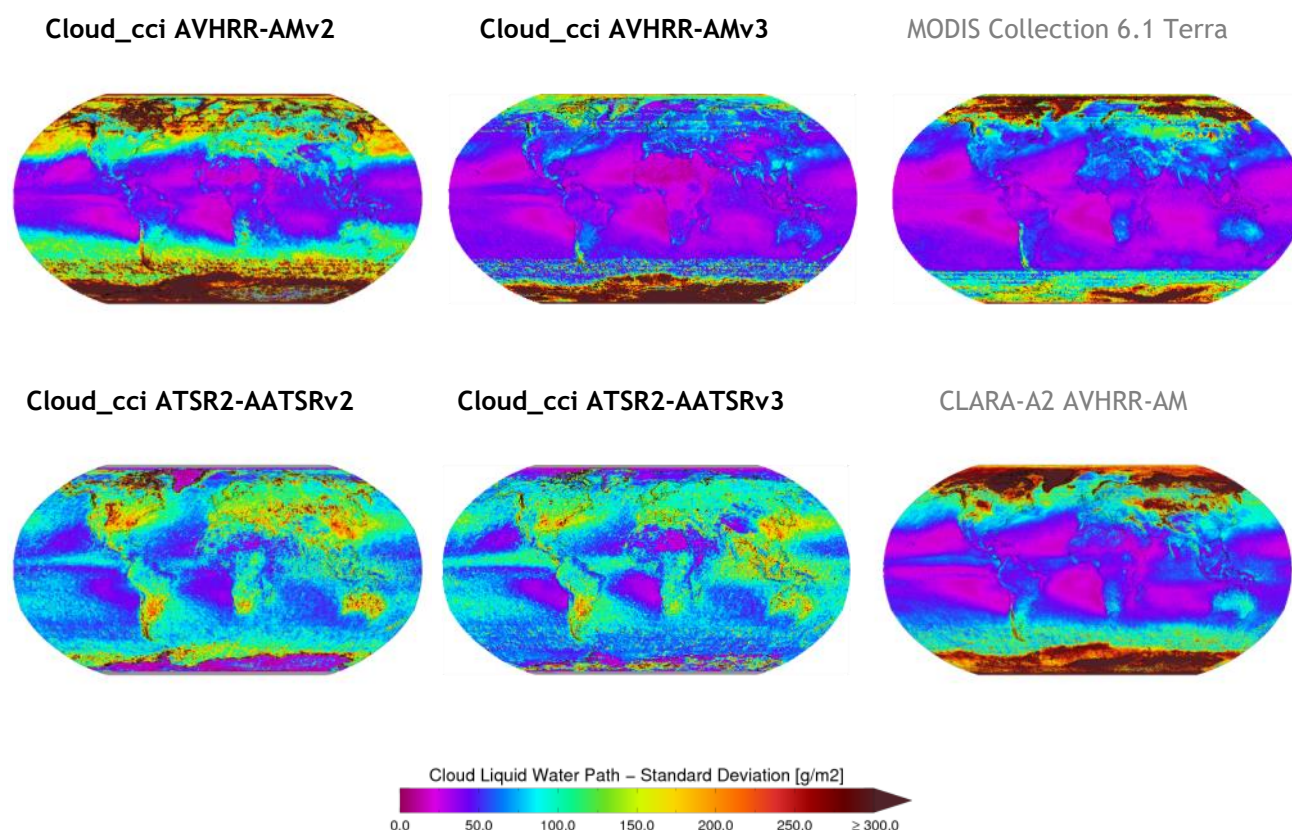



Figure 4-27 Globally gridded standard deviation of cloud liquid water path for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

General findings

- For large parts of the globe the spatial patterns of liquid water path looks comparable among the datasets, with a small maximum in the inner tropics, somewhat less LWP in the subtropics (in particular in the stratocumulus regions) and increased LWP again for the mid-latitudes. For MODIS C6.1, CLARA-A2 and AVHRR-PMv2 very high mean LWP values are visible for the polar regions, except in Cloud_cci ATSR2-AATSRv3 and Cloud_cci ATSR2-AATSRv2. The high mean LWP in these regions partly reaches 800 g/m², while the mean LWP is below 200 g/m² in all datasets for the rest of the globe.
- The temporal variability of LWP is strongly correlated to the mean LWP values, meaning the regions with low LWP exhibit low temporal variability and vice versa. Polar regions show highest variability at least in those datasets which also have high mean values there.
- Global (60S-60N) mean values are lowest for Cloud_cci AVHRR-AMv3 (approx. 93 g/m²) and highest for the previous version Cloud_cci AVHRR-AMv2 with approx. 130g/m², but all datasets are stable beyond 2003. Transitions from one sensor to another caused jumps in the time series for Cloud_cci AVHRR-AMv3 and v2, CLARA-A2 and ATSR2-AATSRv2. In addition, LWP of Cloud_cci AVHRR-AMv3 and v2 as well as CLARA-A2 is significantly impacted by the satellite drift of NOAA-12 between 1991 and 1999.
- Time series of the LWP highlight a stable behaviour for all datasets beyond 2003, although the mean values differ partly by 50%

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 56				

- Time series (in 60S-60N) of all-sky means suggests a good agreement among all datasets with the exception of AVHRR-AMv2 which is significantly higher.

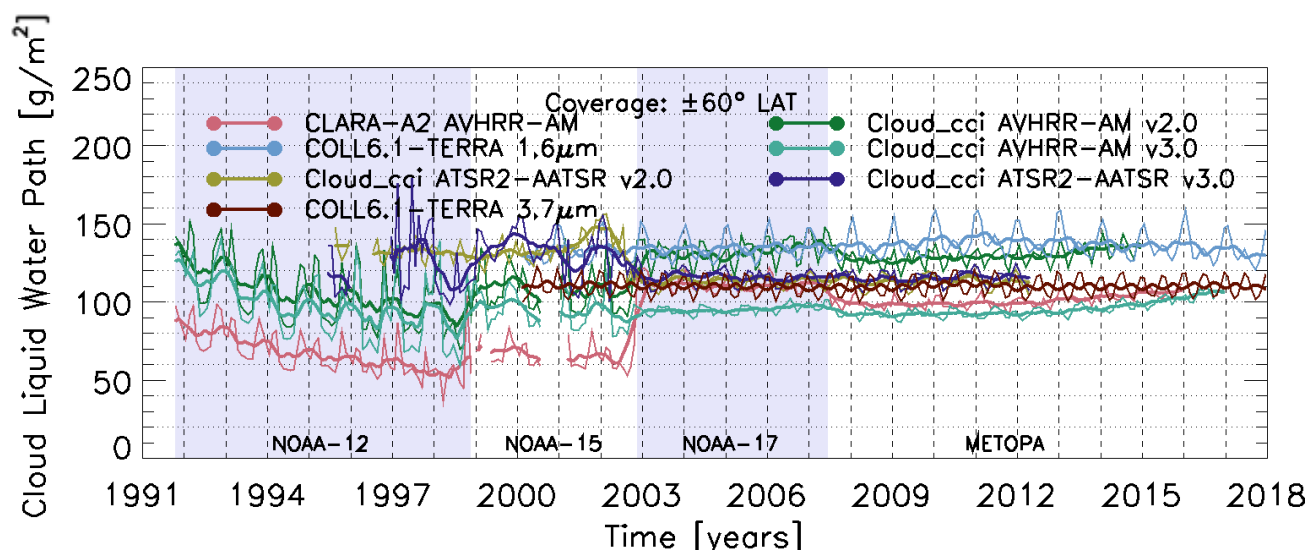


Figure 4-28 Time series of monthly “in-cloud” cloud liquid water path for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.

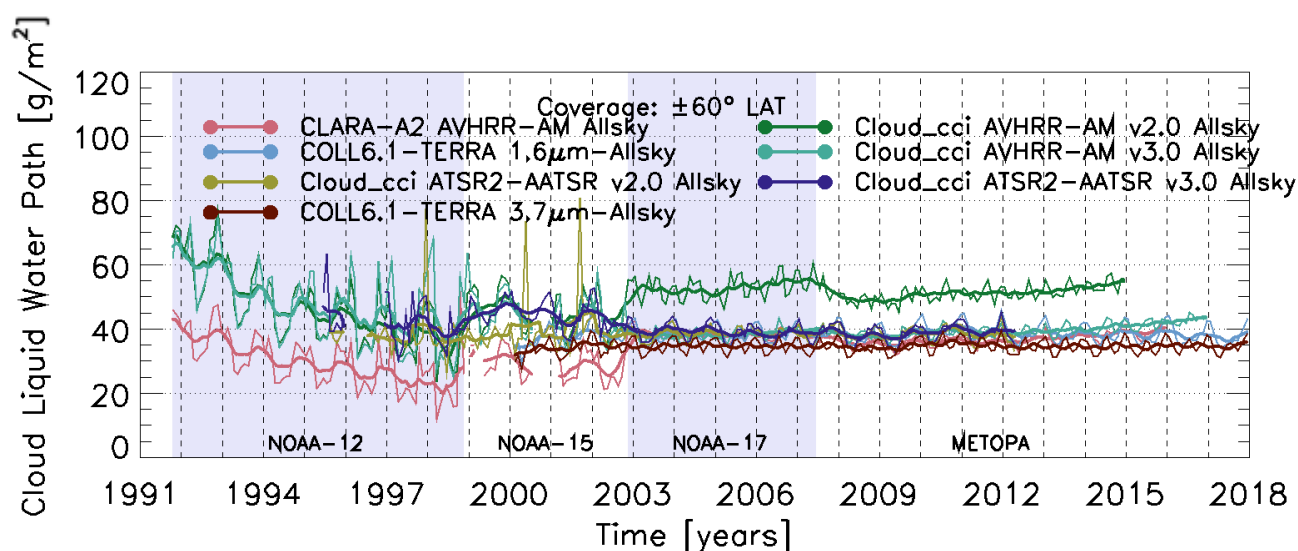



Figure 4-29 Time series of monthly “all-sky” cloud liquid water path for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 57

Discussion

All Cloud_cci datasets reflect the typical meridional distribution of LWP, with a small local maximum in the inner tropic, smaller values in the outer tropics and slight increasing mean LWP values towards the mid and higher latitudes. Cloud_cci AVHRR-AMv3 generally presents the lowest values while Cloud_cci AVHRR-AMv2 shows the maximum values for nearly all latitude bands. Zonal mean values and the spread among the datasets increase dramatically in higher latitudes. Evaluating the time series plots, from NOAA-17 onwards Cloud_cci AVHRR-AM v3 agrees very well with the respective MODIS C6.1 Terra retrievals for 1.6 μ m. Also Cloud_cci ATSR2-AATSRv3 (using 1.6 μ m) is very similar to MODIS Terra C6.1 for 1.6 μ m. A resulting jump can be seen at the transition from NOAA15 to NOAA17, which has been improved in Cloud_cci AVHRR-AMv3. The LWP of ATSR2 v2 (using 3.7 μ m) seemed very high compared to the other 3.7 μ m data, in particular visible when compared against AATSR v2. This feature has also been fixed with Cloud_cci ATSR2-AATSRv3. The found seasonal cycle of 60S-60N time series is similar for the Cloud_cci datasets compared to the reference data, although with deviating amplitude.

Looking at the time series of the all-sky means, the differences appear to become less. Cloud_cci datasets are very close together with the reference data after 2003, except Cloud_cci AVHRR-AMv2, which points to shortcomings of either cloud fraction, cloud phase determination or both.

Table 4-9 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. The biases between -40 and -15 g/m² as well as the bc-RMSD are relative large (values between 80 and 100 g/m²). Decadal trends are divers among the Cloud_cci v3 datasets, but none is as large as the trend of MODIS C6.1 Terra 1.6 μ m with about +7 g/m² /decade.

Table 4-9: Evaluation scores for Cloud_cci v3 Level-3C Liquid Water Path (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. All datasets used the 1.6 μ m channel. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias [g/m ²]	-39.77	-17.29
bc-RMSD [g/m ²]	+80.18	+97.38
Cloud_cci v3.0 trend [g/m ² / decade]	-3.22	+0.99
C6.1-Terra trend [g/m ² / decade]	+7.01	+8.06
Δ trend [g/m ² / decade]	-10.23	-7.06

4.1.8 Ice Water Path

In this subsection the cloud ice water path of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, MODIS C6.1 Terra, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3 and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-30), zonal mean (Figure 4-31) and standard deviation (Figure 4-32), all for a common time period, and time series plots (Figure 4-33 and Figure 4-34).

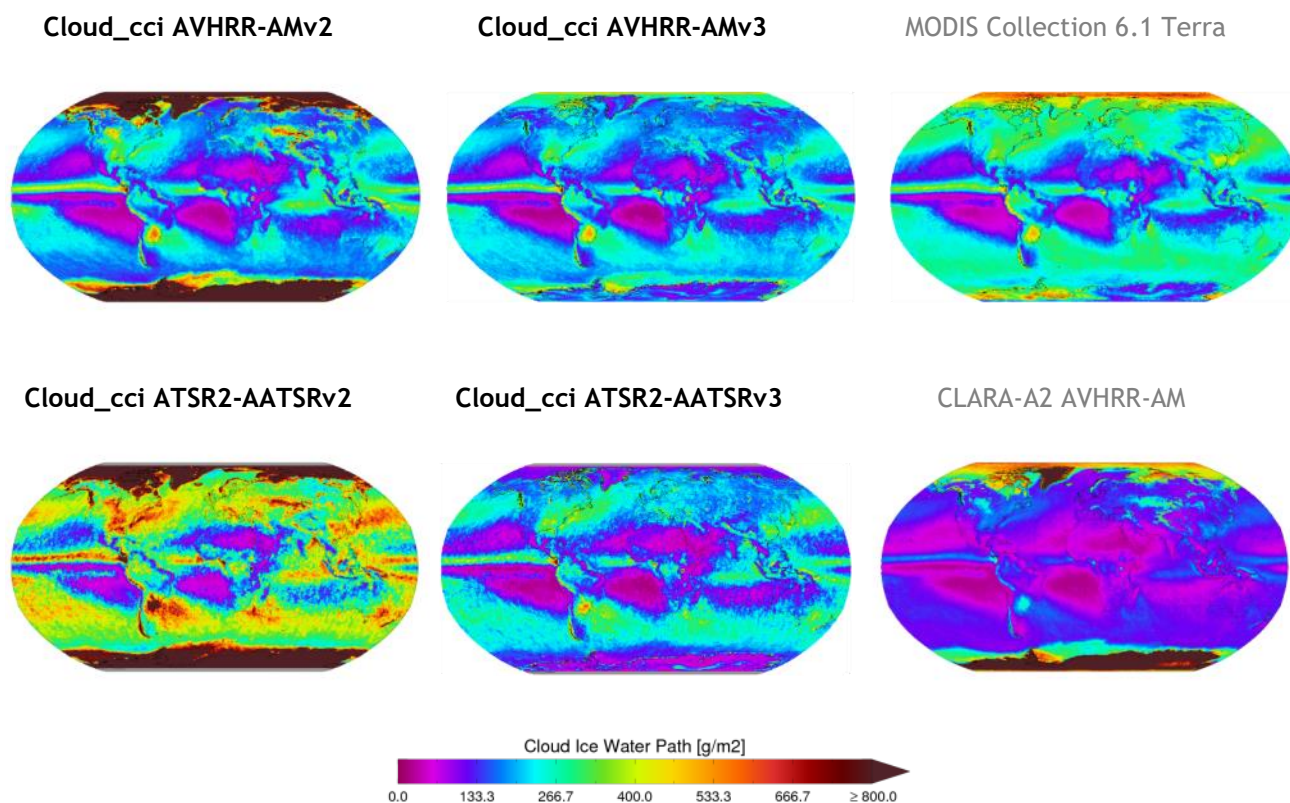


Figure 4-30 Globally gridded means of cloud ice water path for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the 2.1 μ m product is plotted.

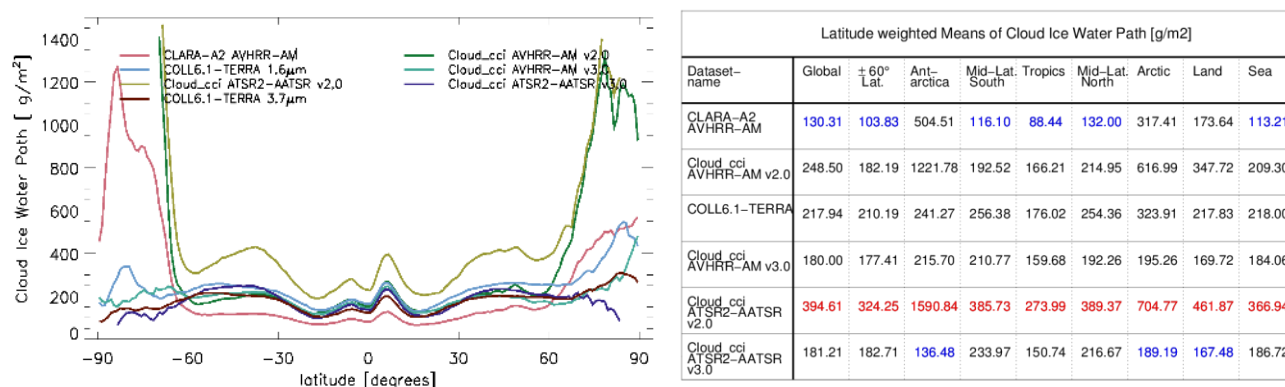



Figure 4-31 Cloud Ice Water Path from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Terra the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 59

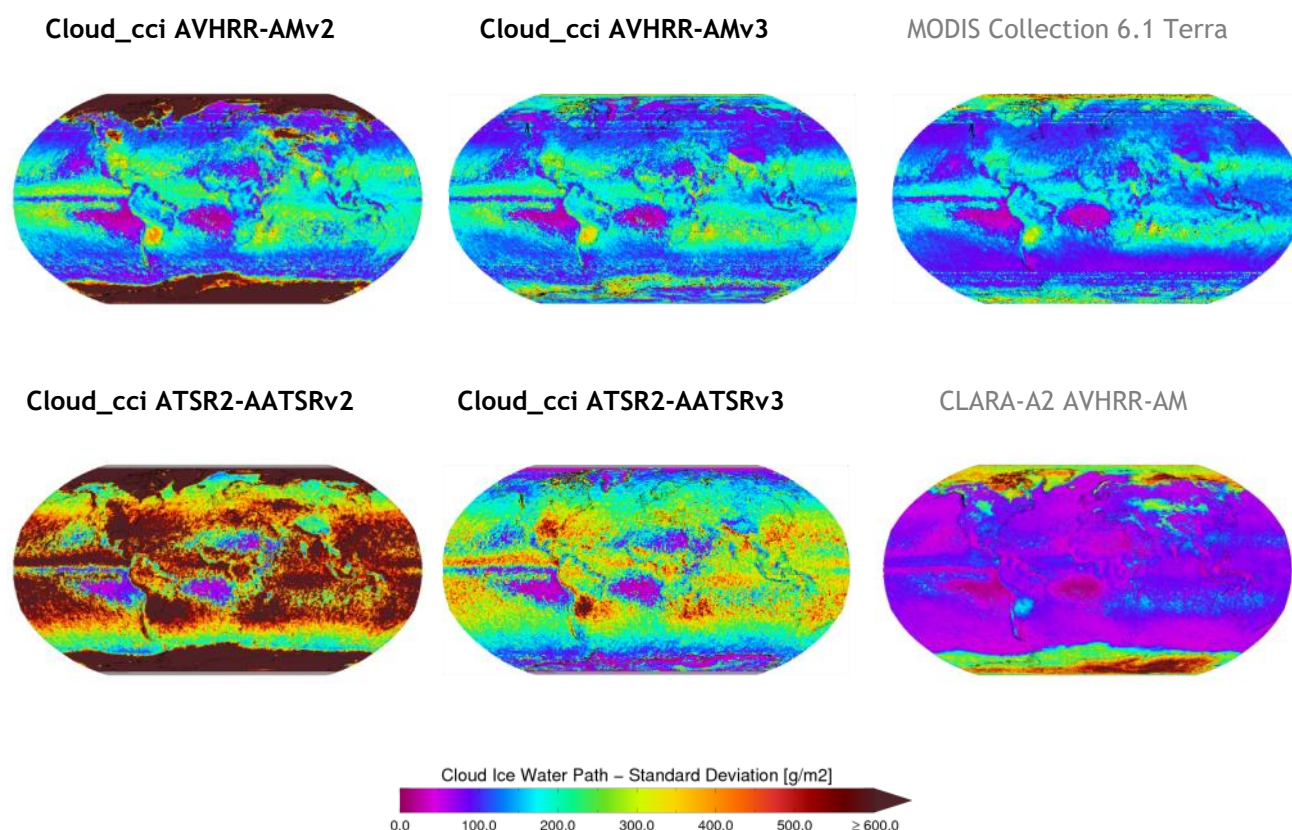


Figure 4-32 Globally gridded standard deviation of cloud ice water path for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Terra the $2.1\mu\text{m}$ product is plotted.

General findings

- Large deviations are found for the different datasets, although the global patterns are similar, which show low mean ice water path values for the subtropics, in particular in the stratocumulus regions, and higher values in the inner tropic and mid-latitudes. Extremely high values are found again for the polar regions in the Cloud_cci v2 datasets, MODIS C6.1 Terra and CLARA-A2. The high mean IWP values in the polar regions for these datasets are partly one order of magnitude higher than in the rest of the globe.
- Similar to LWP, the largest temporal variability for IWP occurs in regions with highest IWP and vice versa. General highest variability is seen in Cloud_cci ATSR2-AATSRv2, while lowest variability is seen for CLARA-A2.
- All datasets are stable in time beyond 2003 in terms of global (60S-60N) mean IWP. Highest values here are approx. 324 g/m^2 for Cloud_cci ATSR2-AATSRv2, lowest values are approx. 100 g/m^2 for CLARA-A2. Transition jumps are seen for Cloud_cci AVHRR-AMv3 and even more clearly for v2 for the change of NOAA-12 to NOAA-15 and of NOAA-15 to NOAA-17. Values for NOAA-12 and NOAA-15 are very high compared to NOAA-17 and Metop-A. A small jump in IWP is also found for the Cloud_cci ATSR2-AATSRv2 transition from ATSR2 to AATSR in the end of 2002. This is improved in ATSR2-AATSRv3.

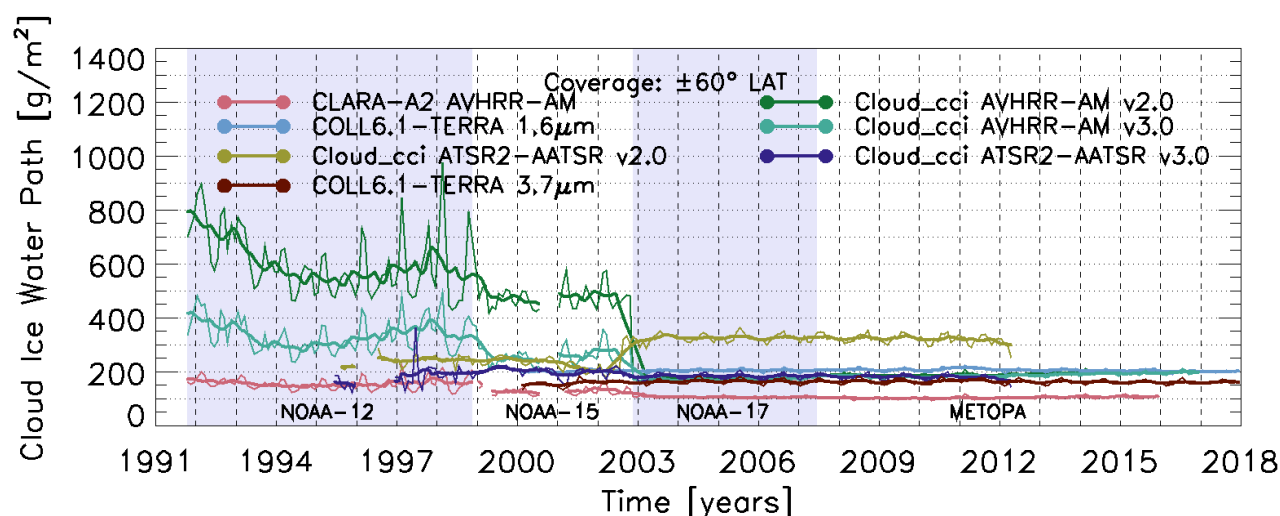


Figure 4-33 Time series of monthly “in-cloud” cloud ice water path for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.

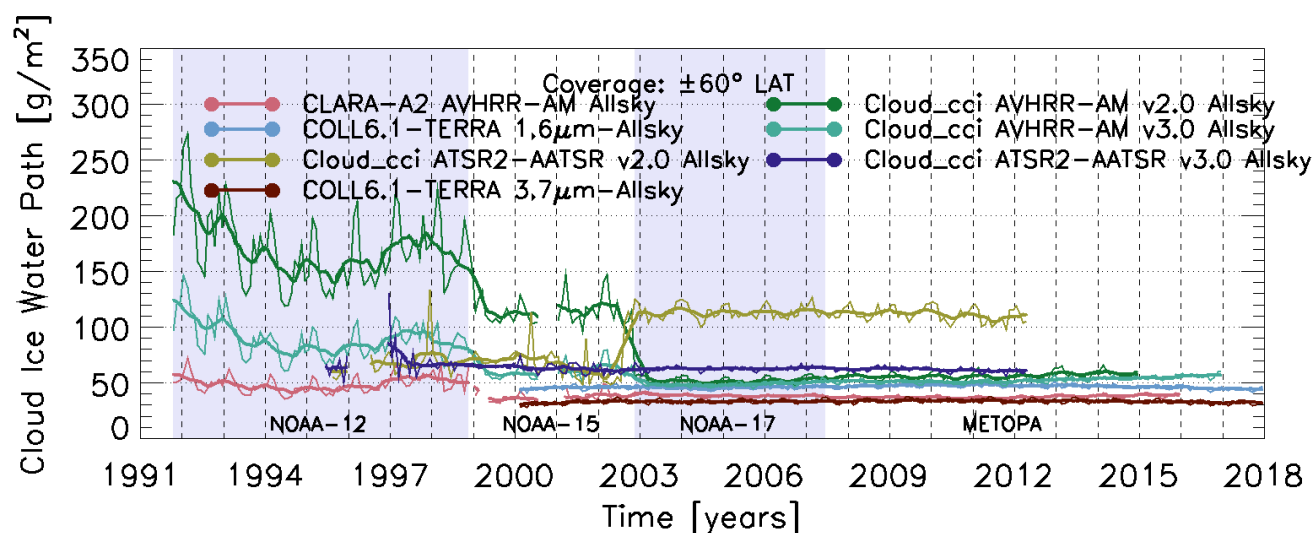



Figure 4-34 Time series of monthly “all-sky” cloud ice water path for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Terra the 1.6 μ m and 3.7 μ m products are shown.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 61

Discussion

As mentioned in the general findings, large deviations are found among all datasets including the Cloud_cci datasets, although the general meridional structure is present in all datasets. Cloud_cci ATSR2-AATSRv2 has much higher IWP compared to the references datasets. For Cloud_cci AVHRR-AMv2 some compensating errors are likely to be present. In addition, Cloud_cci v2 datasets exhibit extremely high, and certainly unrealistic, values for the polar regions, basically above/below 60N/60S. These errors, mainly caused by an ice LUT bug and the BRDF bug already mentioned in the discussion for CER_{ice} are fixed in Cloud_cci v3 datasets. Cloud_cci AVHRR-AMv3 as well as ATSR2-AATSRv3 show less deviations and very similar global patterns but without the extremely high values over the polar regions.

The time series of IWP and all-sky IWP reveal only little year-to-year variability for the 60S-60N mean IWP, this is common among all datasets. Cloud_cci ATSR2-AATSRv3 shows lesser problems at the satellite transition than the older version 2; the same is seen for Cloud_cci AVHRR-AMv3 compared to v2 for NOAA-12 to NOAA-15 and NOAA-15 to NOAA-17 transitions. For NOAA-15 and NOAA-17, Cloud_cci AVHRR-AMv3 values are still higher than MODIS Collection 6.1 Terra, but less unstable as it was in version 2.0 due to the bad illumination condition in their near-twilight orbit which is in addition drifting with time.

Table 4-10 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS Collection 6.1 Terra. The biases for Cloud_cci AVHRR-AMv3 and Cloud_cci ATSR2-AATSRv3 after 2003 are relatively high. For AVHRR-AMv3 and ATSR2-AATSRv3 (both using the 1.6 μ m channel and are compared against MODIS C6.1 Terra 1.6 μ m) have very large biases around -30 g/m². The bc-RMSD is between 100 and 215 g/m² with lower value for Cloud_cci AVHRR-AMv3 and larger value for Cloud_cci ATSR2-AATSRv3. The reason for the strong positive trend of AVHRR-AMv3 in the common period (2003-2011) needs to be further investigated. Otherwise, a small negative trend is found for Cloud_cci ATSR2-AATSRv3, while positive trends are seen in MODIS C6.1 Terra 1.6 μ m data.

Table 4-10: Evaluation scores for Cloud_cci v3 Level-3C Ice Water Path (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. All datasets use the 1.6 μ m channel. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM _{1.6}	ATSR2-AATSR _{1.6}
Bias [g/m ²]	-29.52	-28.77
bc-RMSD [g/m ²]	+102.02	+213.60
Cloud_cci v3.0 trend [g/m ² / decade]	+18.75	-2.27
C6.1-Terra trend [g/m ² / decade]	+8.45	+7.82
Δ trend [g/m ² / decade]	+10.30	-10.09

4.1.9 Liquid Cloud Fraction

In this subsection the liquid cloud fraction of Cloud_cci AVHRR-AMv2, Cloud_cci AVHRR-AMv3, MODIS Collection 6.1 Terra, Cloud_cci ATSR2-AATSRv2, Cloud_cci ATSR2-AATSRv3 and CLARA-A2 (morning satellites only) are compared by means of multi-annual mean (Figure 4-35), zonal mean (Figure 4-36) and standard deviation (Figure 4-37), all for a common time period, and time series plots (Figure 4-38). It should be noted that for MODIS C6.1 Terra the “Cloud_Phase_Infrared” phase product is chosen for the comparison in this section.

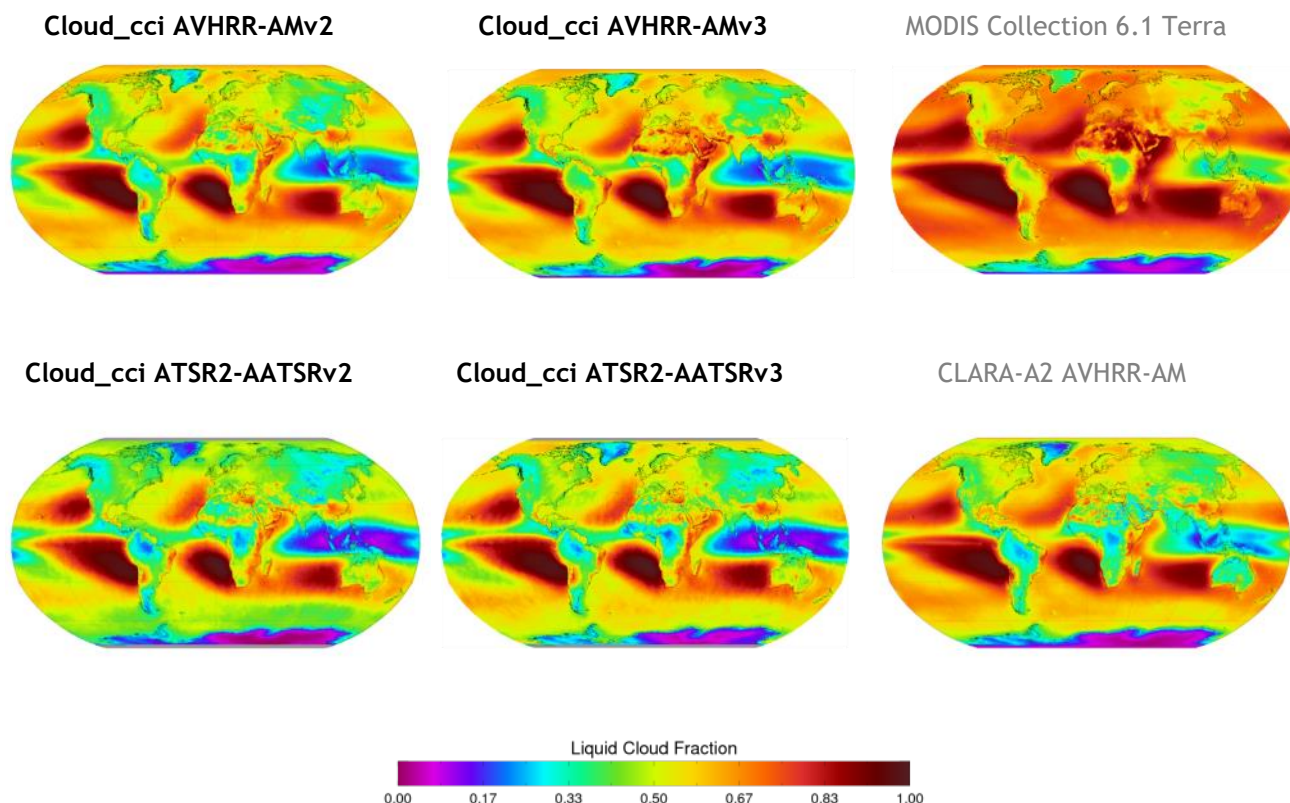


Figure 4-35 Globally gridded means of liquid cloud fraction for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

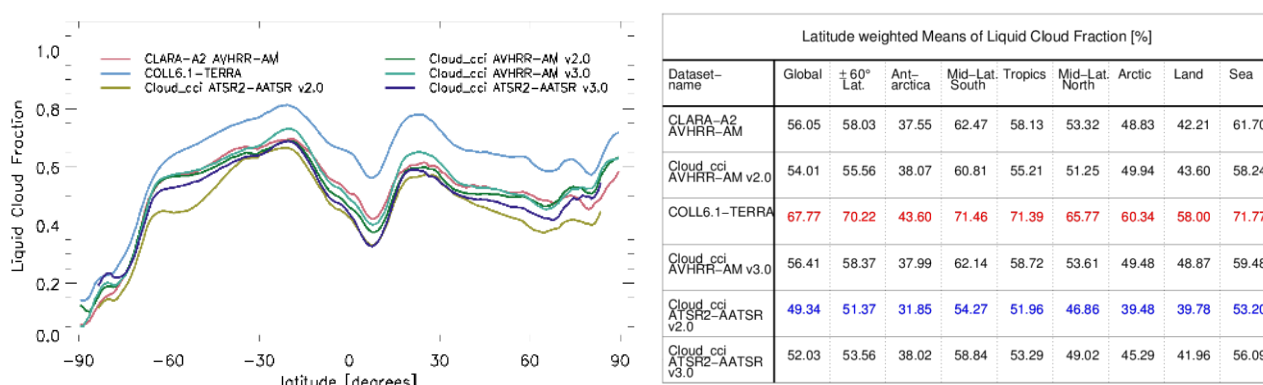



Figure 4-36 Liquid Cloud Fraction from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. **Highest** and **lowest** values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 63

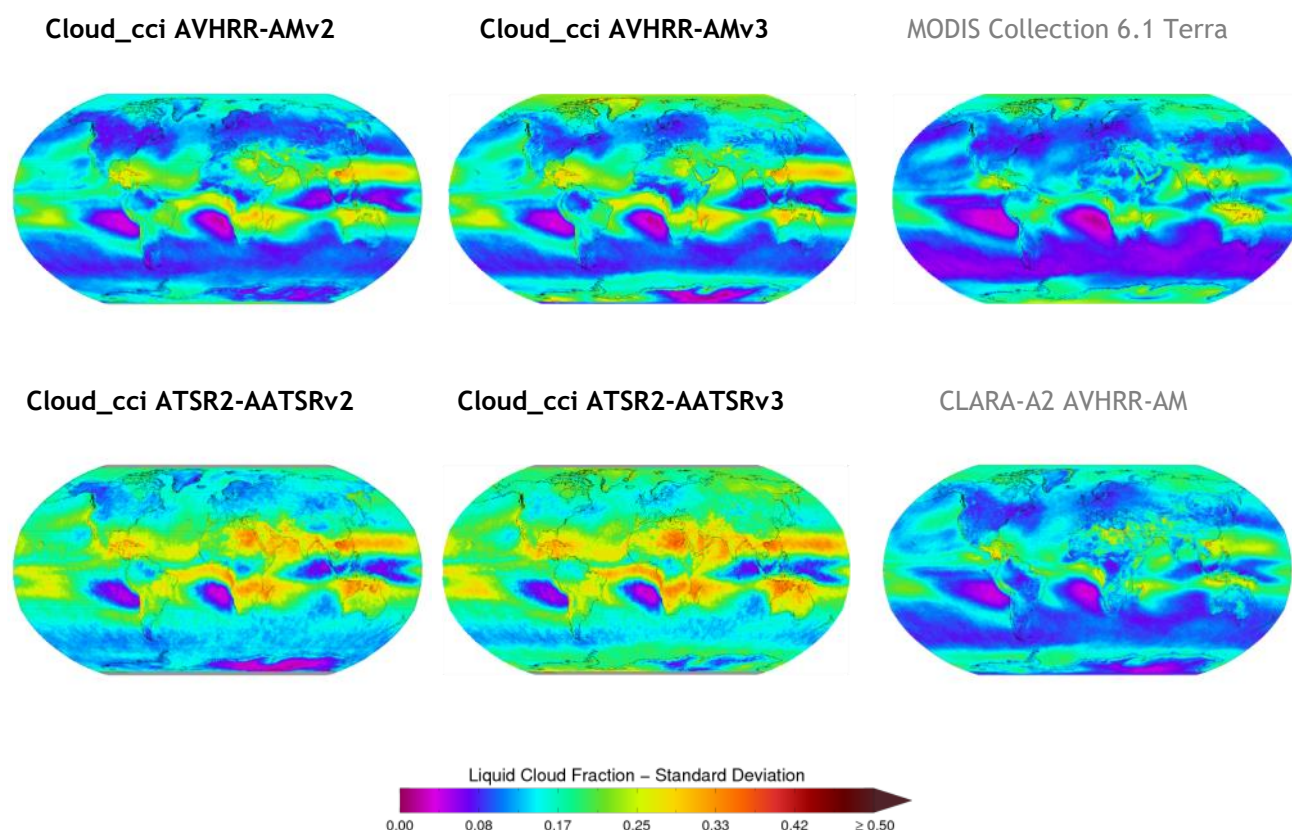



Figure 4-37 Globally gridded standard deviation of liquid cloud fraction for all morning satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- All datasets agree on the main global patterns of liquid cloud fraction (LCF): high LCF (close to 1) in the stratocumulus regions, low values in the inner tropics with a distinct local minimum in the pacific warm pool region. Another local minimum is found for the Antarctica region in all datasets with nearly no liquid cloud tops in some parts. MODIS C6.1 Terra has more liquid clouds then the other datasets.
- The spread among the datasets is up to 20% in some latitude band, but it does not increase towards the polar regions as it was seen for most other cloud variables in the subsections above.
- The temporal variability of LCF is relative similar for Cloud_cci AVHRR-AMv3 and v2, MODIS C6.1 Terra and CLARA-A2. It's a bit higher for Cloud_cci ATSR2-AATSRv2 and largest for Cloud_cci ATSR2-AATSRv3. However, for all datasets the largest temporal variability is found in the subtropical regions and lowest for southern hemispheric stratocumulus regions.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 64				

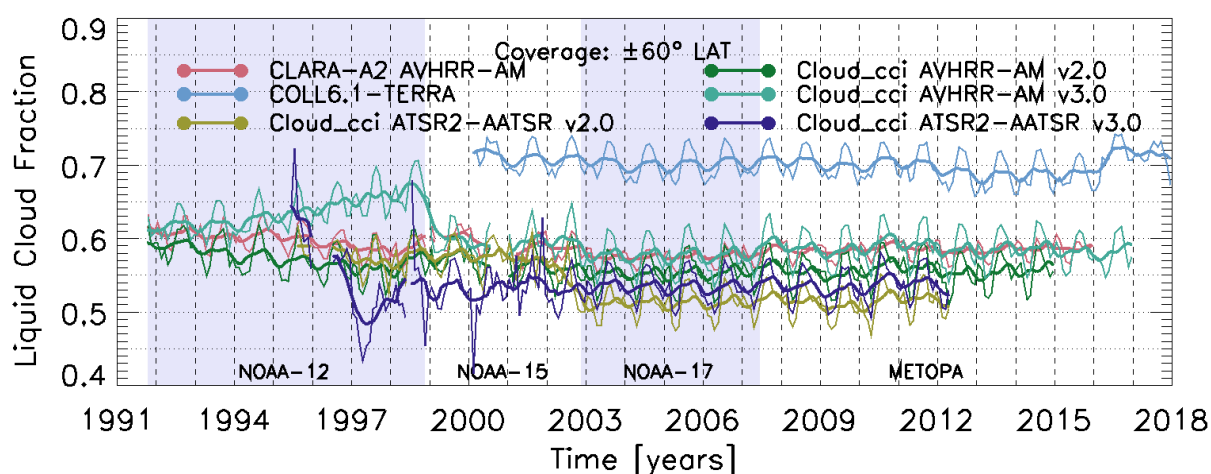


Figure 4-38 Time series of monthly liquid cloud fraction for all morning satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.


Discussion

The global patterns of LCF of the Cloud_cci datasets are very similar to the reference data. Regions with dominant liquid cloud occurrence are identified as well as regions with high occurrences of ice clouds, e.g. tropical regions over the eastern Indian Ocean and western Pacific Ocean. There is a LCF of about 50% of liquid clouds over the arctic regions while the liquid clouds fraction in Antarctica is very low. In particular the agreement in the latter is very convincing. Apart from the general agreement, the Cloud_cci datasets show systematic deviations among them and in comparisons to the reference datasets. MODIS Collection 6.1 Terra has relative high LCF, which could indicate that small, low cumulus clouds might be identified with higher spatial resolution of the MODIS sensor. However, ATSR2-AATSR based datasets do not show this high LCF even though being on the same spatial resolution.

All Cloud_cci LCF data are relative stable in time, except a decreasing trend seen for Cloud_cci AVHRR-AMv2 during NOAA-12's lifetime, an increasing trend seen for Cloud_cci AVHRR-AMv3 also during NOAA-12's lifetime and a jump found for Cloud_cci ATSR2-AATSRv2 at the sensor transition in 2002. The seasonal cycle for the 60S-60N mean values agrees well among the Cloud_cci datasets and also with respect to the reference data. It should be mentioned that for MODIS C6.1 Terra a noticeable jump in LCF is found in 2016 and possible reasons need to be further investigated.


Table 4-10 presents evaluation scores inferred from comparisons of Cloud_cci v3 datasets with MODIS C6.1 Terra. Since MODIS C6.1 Terra has exceptionally high LCF, both Cloud_cci datasets appear to have negative biases when compared against C6.1 Terra. The biases are somewhat larger for Cloud_cci ATSR2-AATSRv3 (-16.62%) while being a bit smaller for AVHRR-AMv3 (-11.85%). The bc-RMSD values range from 10 to 15 %. The trends of the Cloud_cci datasets are consistently of positive sign and MODIS C6.1 Terra only shows a very small positive trend of 0.3% / decade, which also shows the high stability of this time series.

Table 4-11: Evaluation scores for Cloud_cci v3 Level-3C Liquid Cloud Fraction (morning satellites) based on comparison to MODIS C6.1 Terra from 2003-2011. The scores were calculated separately for each Cloud_cci

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 65

dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset	
	AVHRR-AM	ATSR2-AATSR
Bias [%]	-11.85	-16.62
bc-RMSD [%]	+10.40	+15.17
Cloud_cci v3.0 trend [% / decade]	+2.44	+0.73
C6.1-Terra trend [% / decade]	+0.30	+0.30
Δ trend [% / decade]	+2.14	+0.43

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 66

4.2 Afternoon satellites

In this subsection the afternoon satellite datasets are compared, these are:

- Cloud_cci AVHRR-PMv2 (see Section 2.2 for details)
- Cloud_cci AVHRR-PMv3 (see Section 2.2 for details)
- CLARA-A2 (only subset of afternoon satellites used; see Section A.8 for details)
- MODIS Collection 6.1 Aqua (see Section A.10 for details)
- PATMOS-x (only subset of afternoon satellites used; see Section A.9 for details)
- ISCCP (0300AMPM data used, see Section A.11 for details)

For interpreting the following comparisons of cloud optical thickness, cloud effective radius, liquid water path and ice water path correctly it is necessary to revisit the channel settings of all sensors/datasets:

For Cloud_cci AVHRR-PMv3 and v2, CLARA-A2 and PATMOS-x the 3.7 μ m channel was used for the products compared, except for a 2-year period (approx. 2000-2002) for which the 1.6 μ m channel was switched on NOAA-16/AVHRR. MODIS Collection 6.1 Aqua has three NIR channels and MODIS Collection 6.1 data usually has optical property products for each of them. In our comparisons, the maps always show the 3.7 μ m product, while in the time series and zonal mean plots 1.6 μ m and 3.7 μ m MODIS Collection 6.1 Aqua products are visualized if both exist.

Table 4-12 Near-infrared channels utilized for afternoon satellite datasets compared in this section. This information is needed for correctly interpreting the comparison results shown for cloud optical thickness, cloud effective radius, liquid water path and cloud water path. ISCCP does not use any Near-infrared channel. Time periods given are approximated.

Channel used	Cloud_cci AVHRR-PMv2	Cloud_cci AVHRR-PMv3	CLARA-A2	PATMOS-x	MODIS Collection 6.1 Aqua	ISCCP
1.6 μ m	2000-2002	2000-2002	2000-2002	2000-2002	(2002-2017) ^T	-
2.1 μ m	-		-	-	(2002-2017) ^M	-
3.7 μ m	1982-2000 2003-2014	1982-2000 2003-2016	1982-2000 2003-2015	1982-2000 2003-2015	2002-2017 ^T	-

4.2.1 Cloud Fraction

In this subsection the cloud fraction of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS Collection 6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-39), zonal mean (Figure 4-40) and standard deviation (Figure 4-41), all for a common time period, and time series plots (Figure 4-42).

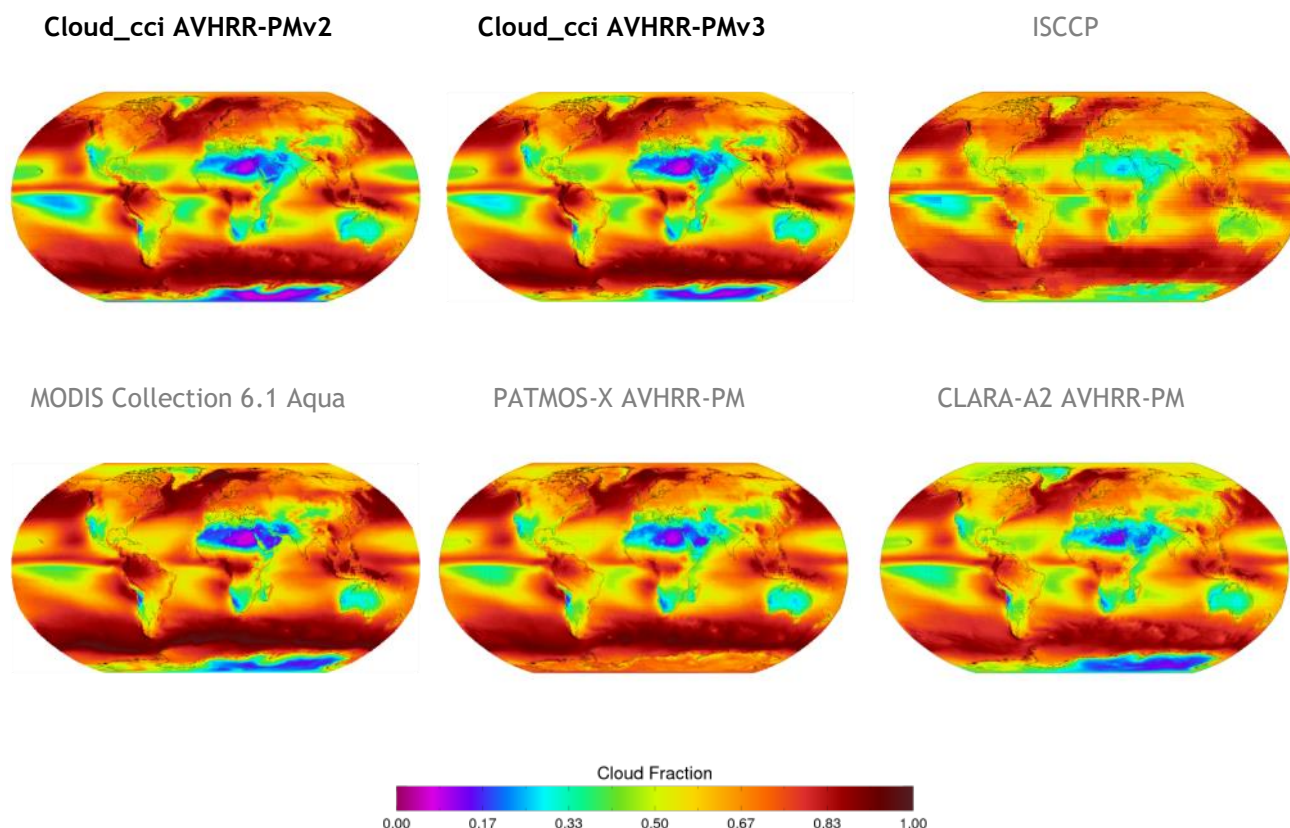


Figure 4-39 Globally gridded means of cloud fraction for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

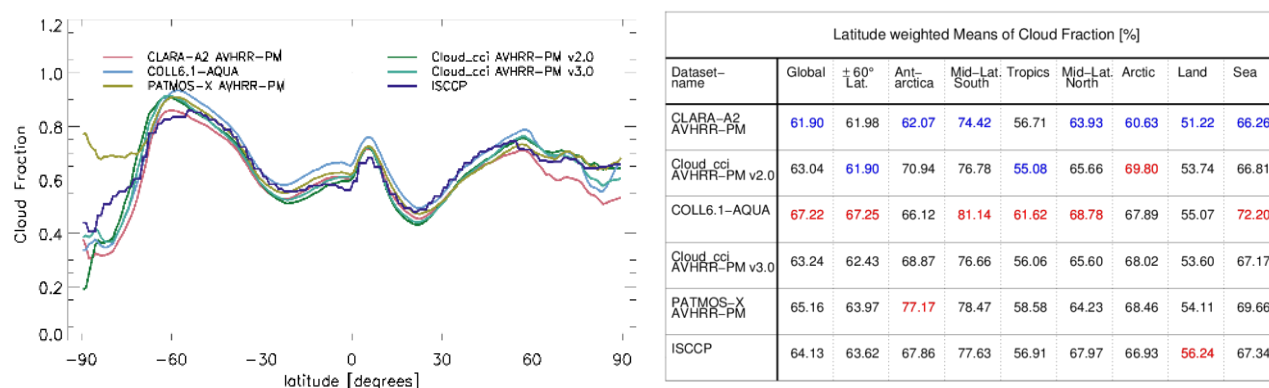



Figure 4-40 Cloud Fraction from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 68

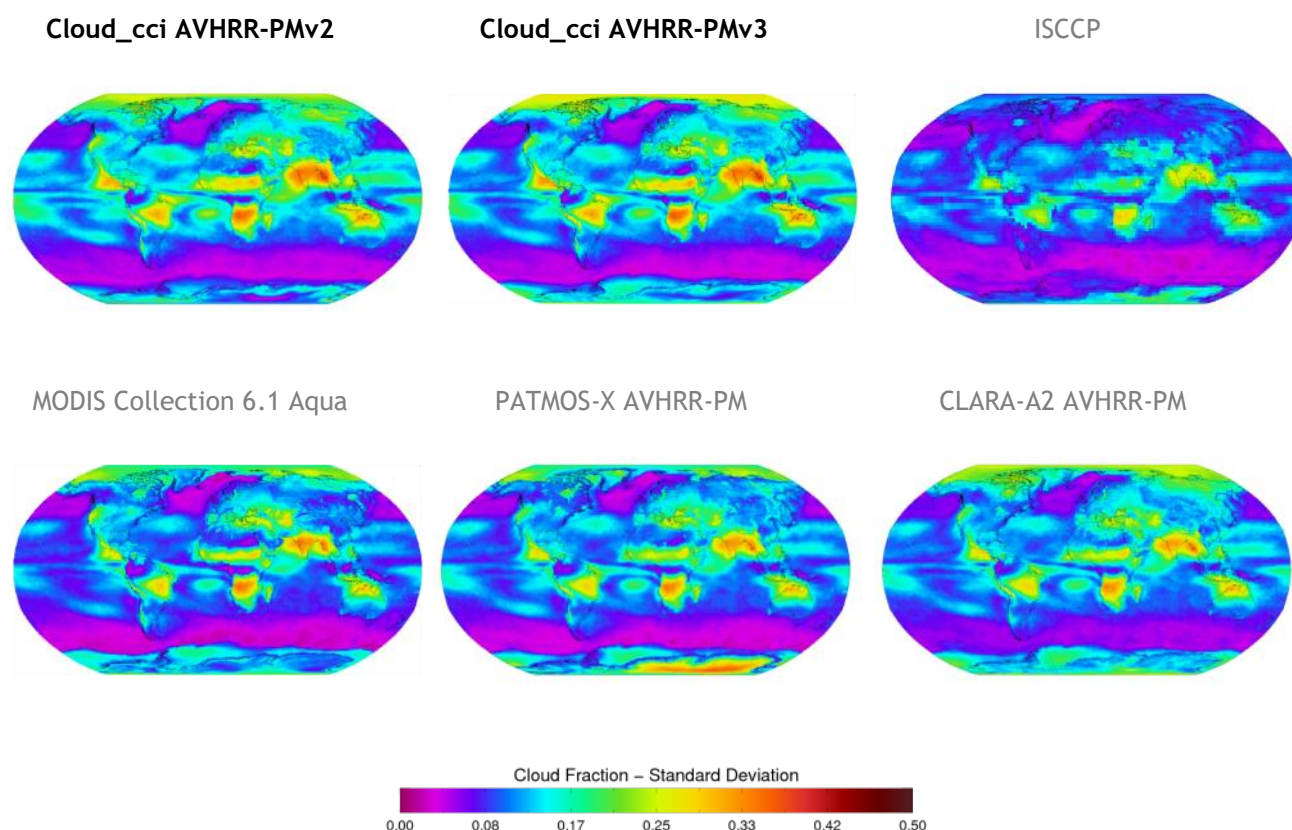


Figure 4-41 Globally gridded standard deviation of cloud fraction for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- The mean CFC is very similar for all datasets with very similar spatial patterns, e.g. (1) high CFC in the tropics, oceanic stratocumulus regions and mid-latitude storm track regions, (2) low CFC in subtropical subsidence regions and Antarctica. Largest spread among the datasets is found for polar regions up to 10% (Arctic) and 70% (Antarctic). On average, MODIS C6.1 Aqua has highest cloud fraction and CLARA-A2 PM lowest (Fig. 4-39).
- The temporal variability of CFC is also very similar between all datasets considered, with the exception of ISCCP which has a somewhat lower variability than the others (Fig. 4-41). Highest variability is found in the outer tropics, in particular over land. Lowest variability at the inner tropics and the oceanic storm track regions of the mid-latitudes.
- While after 2003 all datasets seem relative stable, relative strong variations are visible for PATMOS-x, CLARA-A2 and Cloud_cci AVHRR-PMv3 as well as for v2 at the satellite transitions and during the lifetimes of the early NOAA satellites NOAA-7 through NOAA-14. Similar issues are found for ISCCP which is likely to be caused by transitions between GEO satellites used.

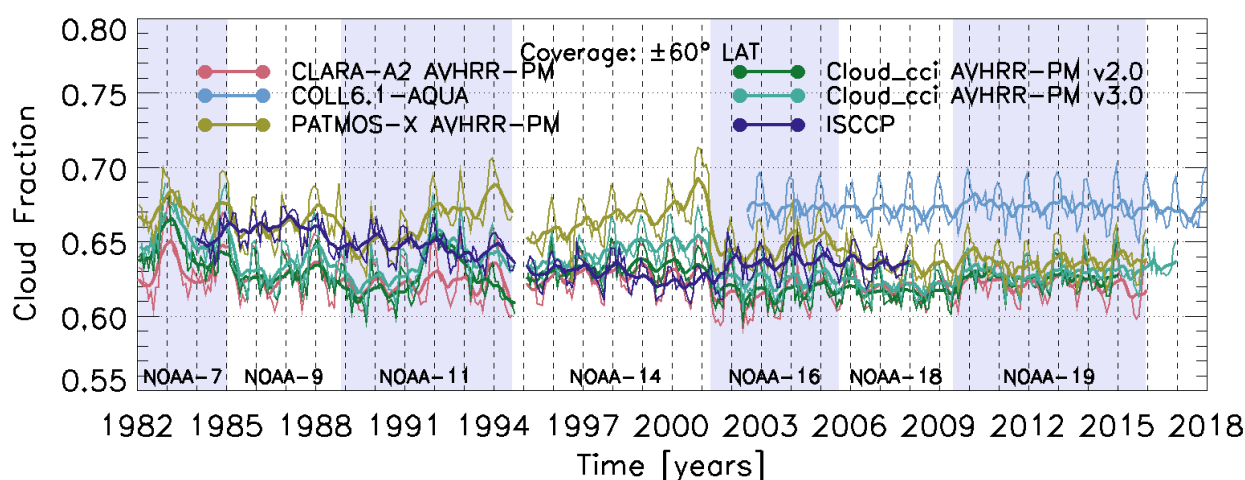


Figure 4-42 Time series of monthly cloud fraction for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

As for the morning satellite based Cloud_cci AVHRR-AMv3 and v2 datasets, also Cloud_cci AVHRR-PMv3 and v2 CFC agrees well with the reference datasets within a 10% range for all parts of the globe within the common period of 2003 to 2011. Only in the polar regions the PATMOS-x reference dataset deviates more than the given range from the others. The datasets Cloud_cci AVHRR-PMv3 as well as Cloud_cci AVHRR-PMv2 are relative stable after 2001, with only small jumps of about 1% at the transition from NOAA-18 to NOAA-19. Before 2001 more inhomogeneities are found for both versions of the Cloud_cci AVHRR-PM dataset. These are caused by satellite drifts of early NOAA satellites, i.e. NOAA-07, -09, -11 and -14, possibly overlaid by some errors introduced by aging of the sensors. Two significant positive anomalies are found for both Cloud_cci AVHRR-PM versions, each in a two to three year period after the severe volcano eruptions of El Chichón (1982) and Pinatubo (1991). These anomalies are not found for CLARA-A2. ISCCP also shows significant inhomogeneities at GEO satellite transitions.

Table 4-13 presents evaluation scores inferred from the comparison of Cloud_cci AVHRR-PMv3 dataset with MODIS C6.1 Aqua. The Cloud_cci bias is below 5% (general underestimation) and bc-RMSD value is between 6 and 7 %. The Cloud_cci AVHRR-PMv3 trends are relative low and resemble to the trends of MODIS Collection 6.1 Aqua.

Table 4-13: Evaluation scores for Cloud_cci v3 Level-3C cloud fraction (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the C6.1 Aqua and the Cloud_cci dataset. This can introduce some small variability in the trend values.

Measure	Cloud_cci v3.0 dataset AVHRR-PM
Bias [%]	-4.83
bc-RMSD [%]	+6.48
Cloud_cci v3.0 trend [% / decade]	+0.46
C6.1-Aqua trend [% / decade]	+0.39
Δ trend [% / decade]	+0.07

4.2.2 Cloud Top Pressure

In this subsection the cloud top pressure of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS Collection 6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-43), zonal mean (Figure 4-44) and standard deviation (Figure 4-45), all for a common time period, and time series plots (Figure 4-46).

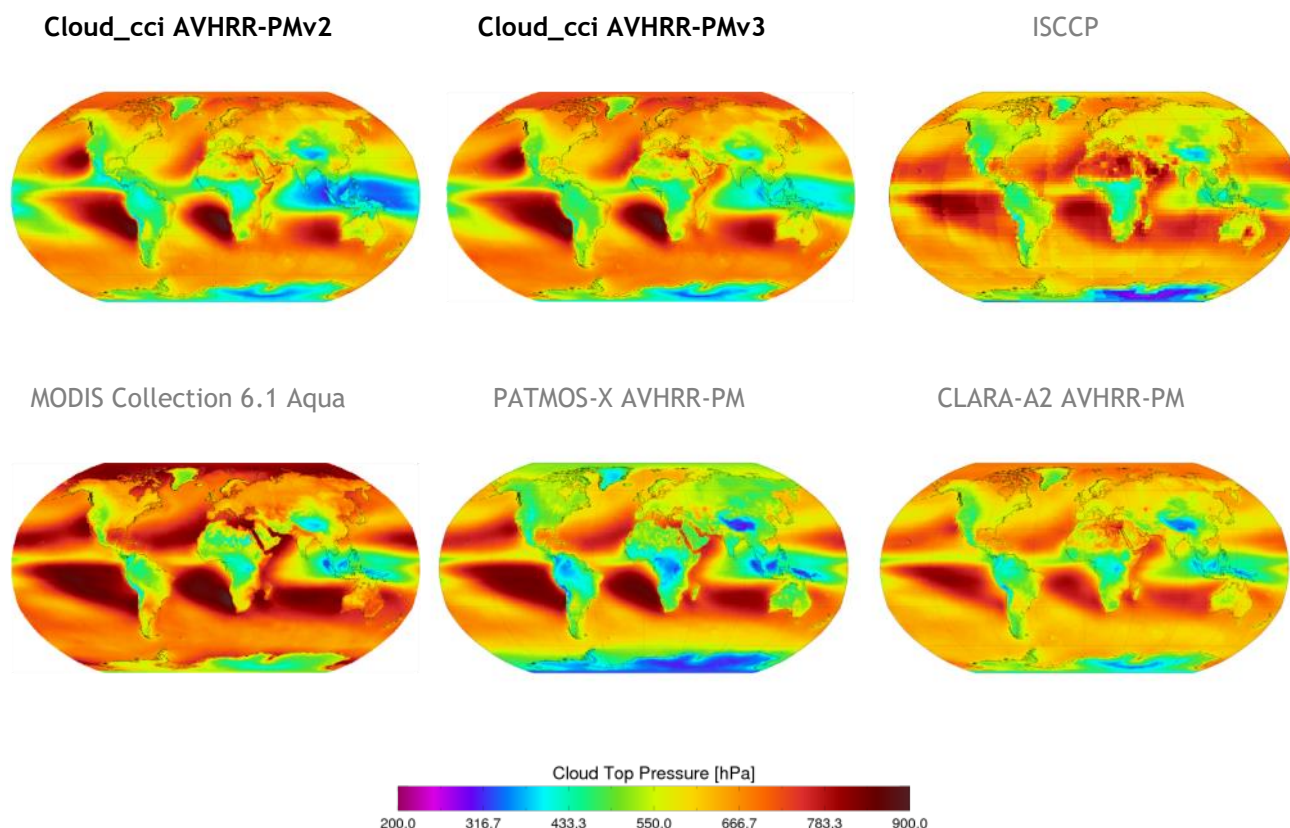


Figure 4-43 Globally gridded means of cloud top pressure for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

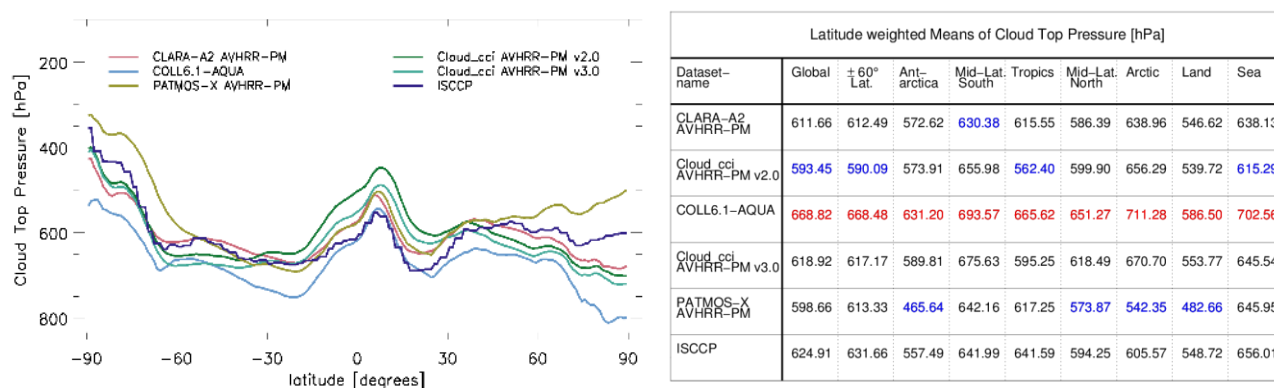



Figure 4-44 Cloud Top Pressure from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. **Highest** and **lowest** values of each region are highlighted.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 71

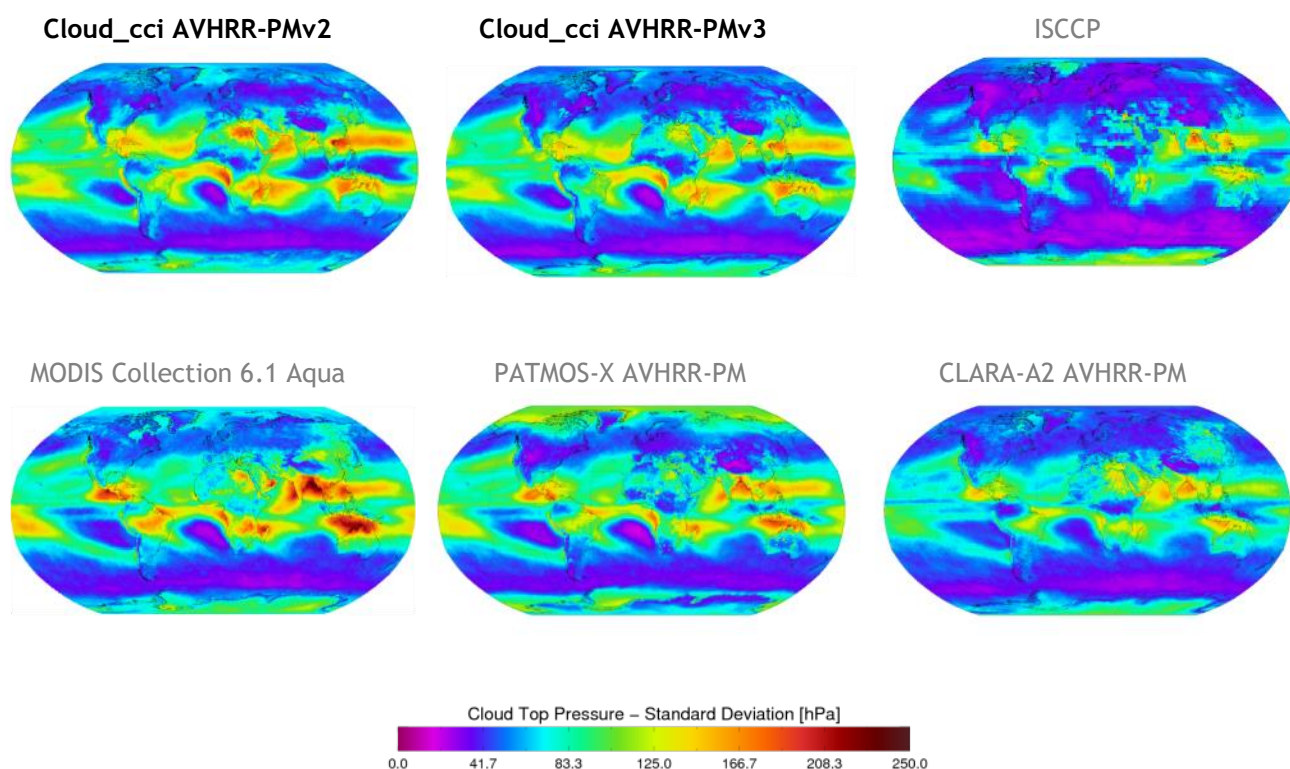


Figure 4-45 Globally gridded standard deviation of cloud top pressure for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- All datasets show the same main features: (1) high mean CTP for the stratocumulus regions, (2) lower mean CTP in the tropics and Antarctica, (3) mean CTP for mid-latitudes are between 600 and 750 hPa (Fig. 4-43). MODIS C6.1 Aqua shows the highest CTP nearly everywhere. Spread among the datasets increases towards the polar regions. Mean maps for ISCCP reveal some stripes which most likely corresponds to the edges of the GEO satellites used.
- Similar temporal variability given in all datasets (MODIS C6.1 Aqua highest and ISCCP lowest values), with highest variability found in outer tropics and lowest variability in the mid-latitudes (Fig. 4-45).
- Time series plots of global mean CTP (60S-60N) show a relative stable behaviour of all datasets with small trends found for the early NOAA satellites in PATMOS-x, CLARA-A2, Cloud_cci AVHRR-PMv3 and v2, leading to jumps at the satellite transitions. For Cloud_cci AVHRR-PMv3 these jumps are noticeable smaller. ISCCP exhibits a positive trend.


	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
		Page 72		

Table 4-14: Relative fraction of low, mid-level, and high clouds calculated from 1-dimensional histograms for a common time period of 2003 to 2011 (excluding the polar regions higher/lower than $\pm 60^\circ$ latitude). Highest values are in red, while lowest values are shown blue. *ISCCP has slightly different definitions of low/mid/high 500/700 hPa instead of 440/680 hPa and does not provide data for the whole period (2003-2007 only).

Dataset	Low clouds	Mid-level clouds	High clouds
CLARA-A2 AVHRR-PM	33.54%	24.90%	41.56%
Cloud_cci AVHRR-PMv2	45.25%	20.48%	34.27%
MODIS Collection 6.1 Aqua	55.61%	10.79%	33.60e%
Cloud_cci AVHRR-PMv3	46.75%	24.31%	28.94%
PATMOS-x AVHRR-PM	46.41%	22.71%	30.88%
ISCCP*	50.56%	22.11%	27.33%

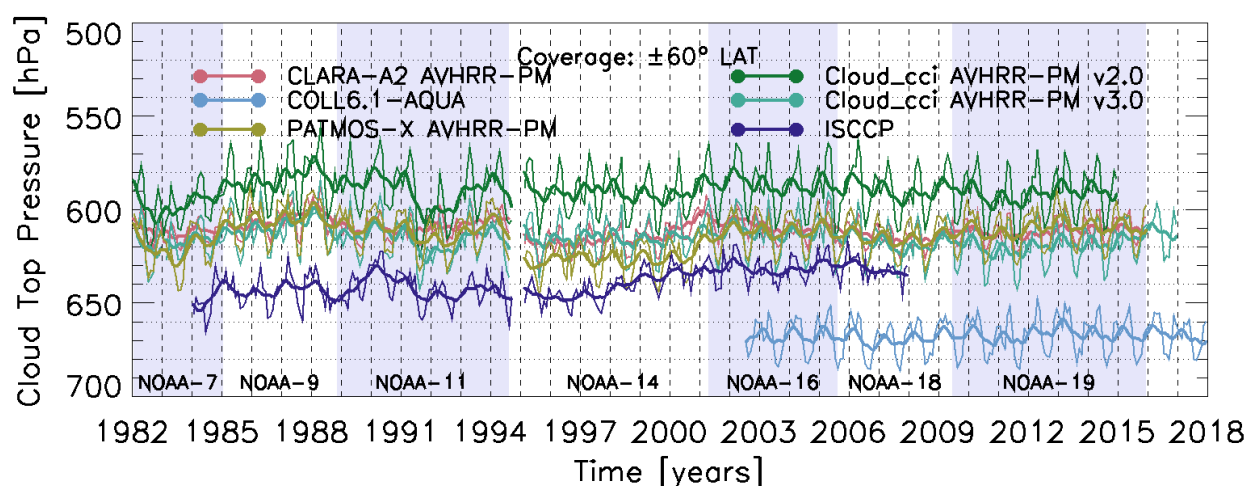


Figure 4-46 Time series of monthly cloud top pressure for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

The Cloud_cci AVHRR-PMv3 CTP data reflects all major global features when compared to the reference data. In the mid latitudes and polar regions, the dataset is more or less in the middle of the spread of all datasets. Since the Cloud_cci AVHRR-PMv2 dataset depicted extreme values with having the lowest CTP on average in the tropic, version 3 dataset lies in the middle of the spread of all datasets. As mean values of CTP give only limited information, due to the bimodal underlying distribution, these comparisons have to be interpreted with caution.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 73

Table 4-14 gives the relative fraction of low, mid- and high-level clouds, which is made by combining CTP and CFC. The relative fraction of the three layers is very similar for Cloud_cci AVHRR-PMv3, Cloud_cci AVHRR-PMv2 and PATMOS-x, while CLARA-A2 has more mid-level and high clouds and MODIS C6.1 Aqua much more low-level clouds. Both Cloud_cci dataset versions are stable in time with two exceptions. After the eruptions of El Chichón (1982) and Pinatubo (1991) positive CTP anomalies are found both lasting for a few years. This is most likely an artefact caused by the positive anomalies in cloud fraction found in these periods, as described in the previous subsection. However, the anomalies are much more visible in version 2 than in version 3. Cloud_cci AVHRR-PMv3 time series agrees much better with the reference datasets than version 2, since there are less high clouds detected in AVHRR-PMv3. Wrongly detected cloudy pixels are also likely to cause high CTP retrievals, which will bias the CTP mean values.

Table 4-15 presents evaluation scores inferred from a comparison of Cloud_cci AVHRR-PMv3 with MODIS C6.1 Aqua. There is a general underestimation of CTP in Cloud_cci dataset compared to MODIS C6.1 Aqua (biases between -55 and -50 hPa). The bc-RMSD values are around 76 to 77 hPa. The Cloud_cci dataset shows a clear positive trend of nearly +13 hPa/decade while MODIS C6.1 Aqua trend is clearly smaller (+1.5 hPa/decade) although of the same sign.

Table 4-15: *Evaluation scores for Cloud_cci v3 Level-3C Cloud Top Pressure (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.*

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM
Bias [hPa]	-51.32
bc-RMSD [hPa]	+76.22
Cloud_cci v3.0 trend [hPa / decade]	+12.70
C6.1-Aqua trend [hPa / decade]	+1.45
Δ trend [hPa / decade]	+11.25

4.2.3 Cloud Optical Thickness - Liquid clouds

In this subsection the cloud optical thickness of liquid clouds (COT_{liq}) of Cloud_cci AVHRR-PM version 2.0 and 3.0, CLARA-A2 (afternoon satellites only), MODIS C6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-47), zonal mean (Figure 4-48) and standard deviation (Figure 4-49), all for a common time period, and time series plots (Figure 4-50). The COTs of the products refer to slightly different wavelengths¹.

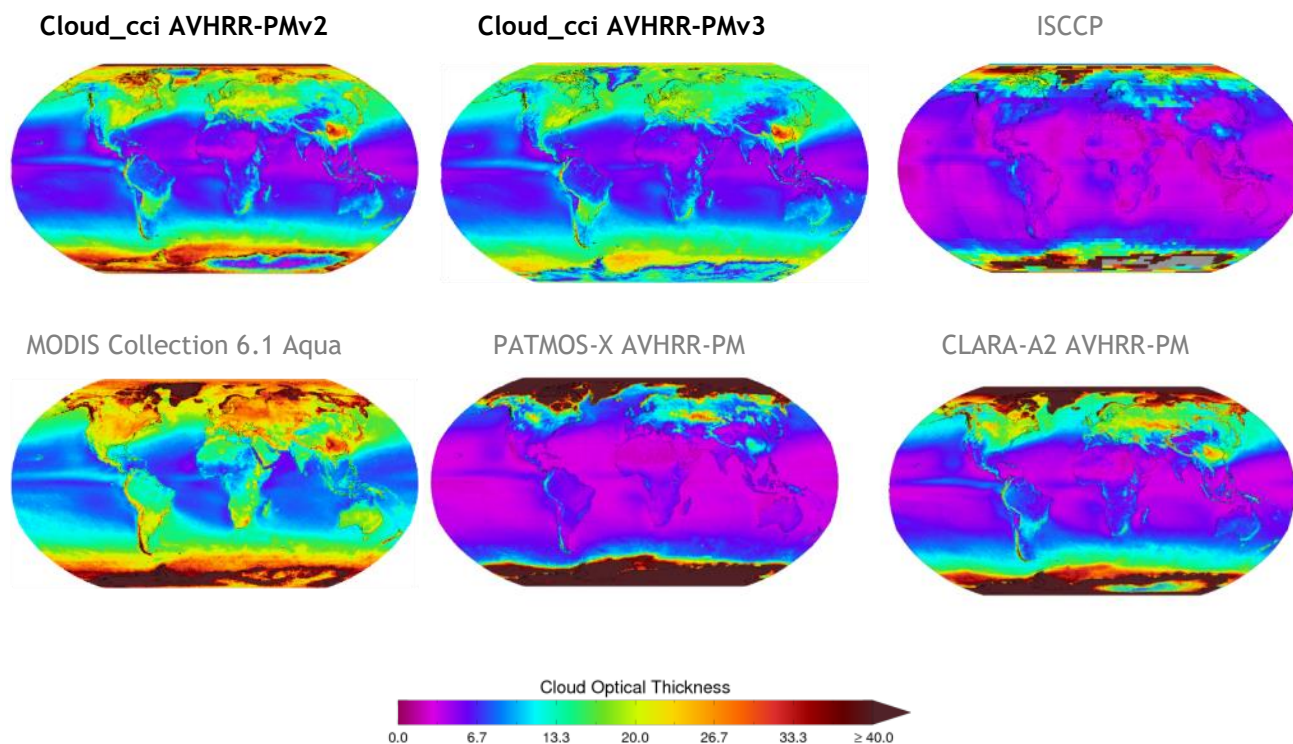


Figure 4-47 Globally gridded means of liquid cloud optical thickness for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

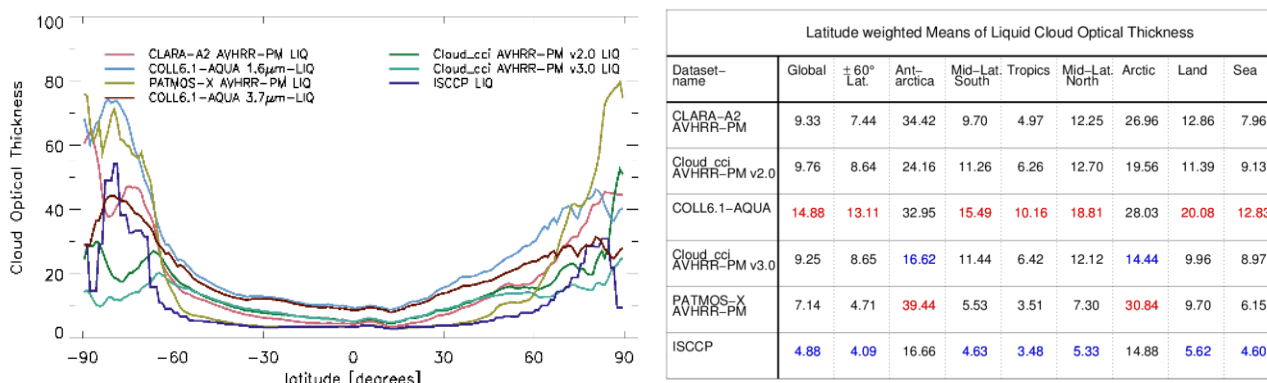



Figure 4-48 Cloud optical thickness - Liquid from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 75

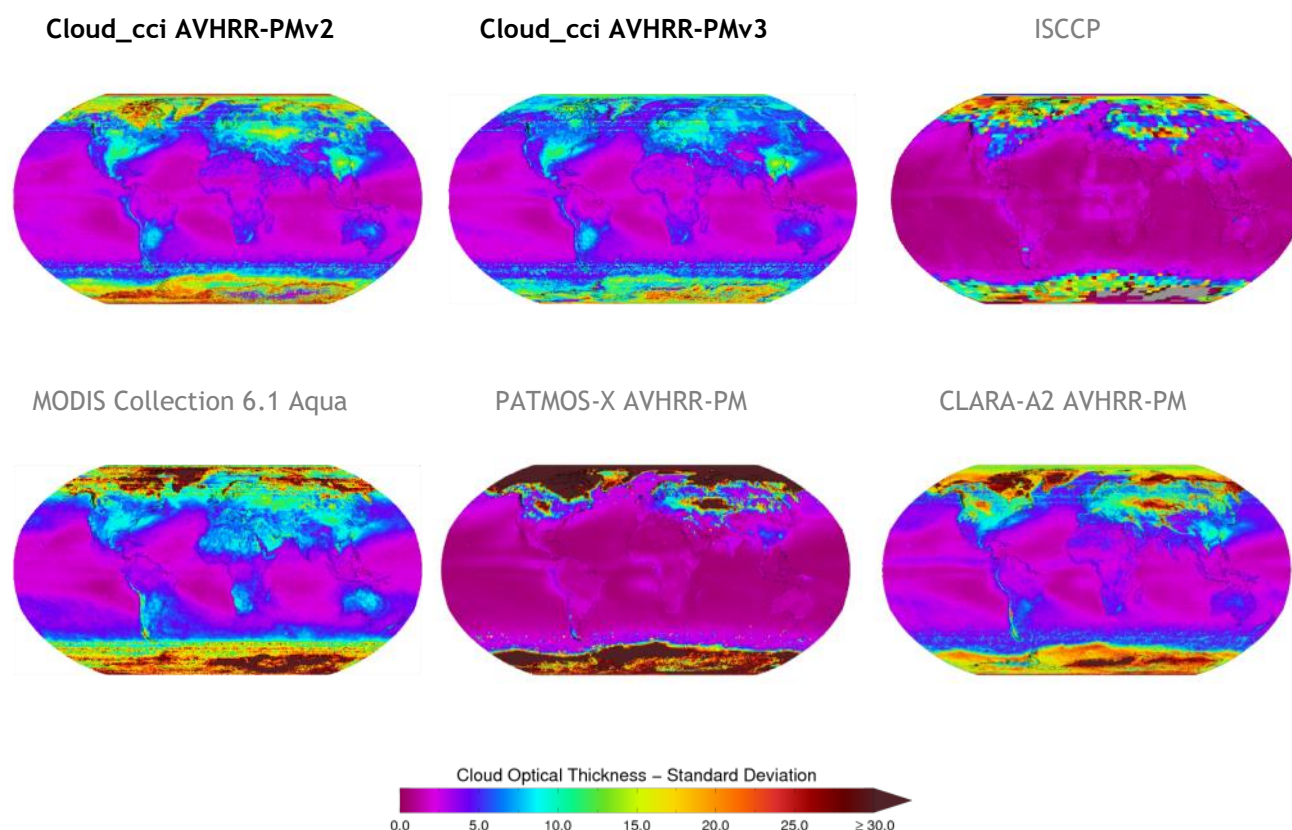



Figure 4-49 Globally gridded standard deviation of liquid cloud optical thickness for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

General findings

- While the main global patterns of COT_{liq} seem similar among all datasets (lower values in the tropics and subtropics, higher values in the higher latitudes), the absolute values are relatively diverse (Fig. 4-47). MODIS C6.1 Aqua COT_{liq} is highest for most parts of the globe except the polar regions in which PATMOS-x and CLARA-A2 provide highest values. Cloud_cci AVHRR-PMv3 as well as Cloud_cci AVHRR-PMv2 are in between, with lower values in v3. Apart from the polar regions, PATMOS-x and ISCCP show much lower mean COT_{liq} than the other datasets.
- The spread among the datasets is largest in the high latitudes. The temporal variability of COT_{liq} has similar spatial patterns as the mean values with highest variability in regions of higher mean COT_{liq} . Also, the differences in temporal variability between the datasets are similar to the difference in the mean values (Fig. 4-49).
- The time series of the MODIS Collection 6.1 Aqua dataset show a very stable behaviour. The AVHRR-based datasets (PATMOS-x, CLARA-A2 and Cloud_cci AVHRR-PM) again suffer from orbital drift, creating a drift in the mean COT_{liq} for the early satellites and jumps at the satellite transitions, which seem most pronounced for PATMOS-x.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 76				

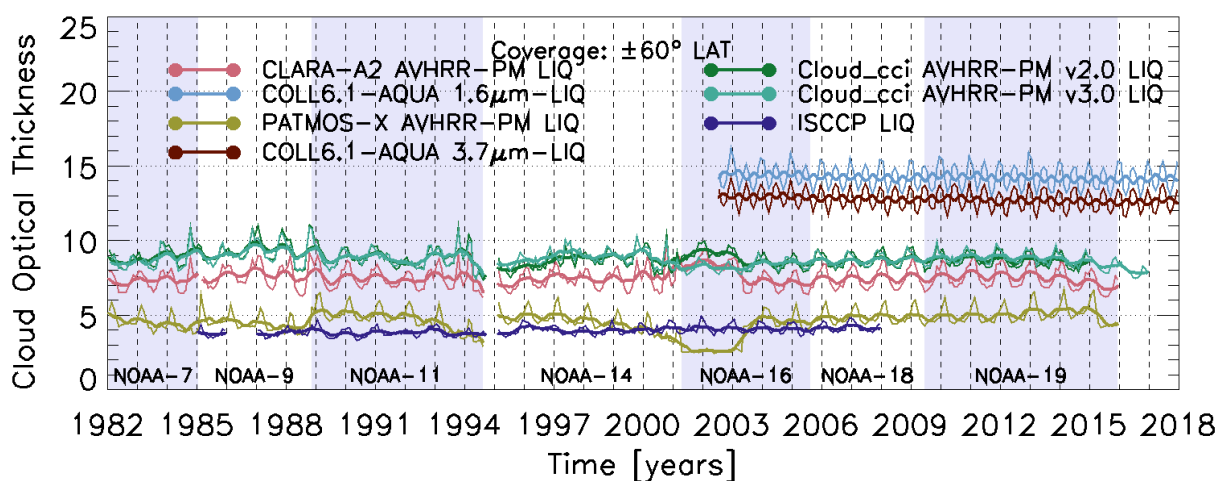


Figure 4-50 Time series of monthly liquid cloud optical thickness for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6µm and 3.7µm products are shown.

Discussion

MODIS C6.1 Aqua COT_{liq} data (1.6 and 3.7µm) present relative high values, while ISCCP and PATMOS-X are at the lower end. The latter is mainly due to the logarithmic averaging done for PATMOS-X in which lower values get higher weights compared to arithmetical averaging. Cloud_cci AVHRR-PM dataset versions and MODIS C6.1 Aqua lie in the middle and agree in most spatial features found on the globe. As the spread among all datasets increases a lot towards the poles, no judgement of the Cloud_cci quality can be made, based on these comparisons in polar regions. However, the globally gridded means and latitude weighted means show, that the new Cloud_cci AVHRR-PMv3 reveal the lowest values in polar regions compared to other datasets. One can only suspect whether higher or lower COT values for liquid clouds in polar regions are realistic or not compared to other regions. As polar regions are often characterized by snow and ice condition at the surface one cannot rule out an impact of those on COT retrievals of semi-transparent and broken clouds if the surface conditions and/or their effects are not well captured.

The Cloud_cci AVHRR-PMv3 and v2 time series of mean COT_{liq} in 60S-60N is relative constant and very similar over the whole time period. The slightly increase in 2001 in version 2.0 marks the 1.6µm channel era of NOAA-16. This slight increase is no longer visible in Cloud_cci AVHRR-PMv3. If the noticeable negative trend in the MODIS C6.1 Aqua dataset might be related to a sensor specific problem, e.g. visible channels degradation, or else needs further research.

Table 4-16 presents evaluation scores inferred from a comparison of Cloud_cci v3 dataset with MODIS C6.1 Aqua. Cloud_cci AVHRR-PMv3 has small negative biases compared to C6.1 Aqua. The bc-RMSD values are slightly smaller than for the morning datasets. The Cloud_cci AVHRR-PMv3 dataset looks relative stable with a small positive trend. In contrast to Cloud_cci AVHRR-PMv3, MODIS C6.1 Aqua shows a persistent, small, negative trend.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55µm, CLARA-A2: 0.6µm, MODIS C6.1 Aqua: 0.65µm (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7µm on COT is assumed to be very minor.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 77

Table 4-16: Evaluation scores for Cloud_cci v3 Level-3C Optical Thickness - Liquid (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias	-3.90
bc-RMSD	+4.77
Cloud_cci v3.0 trend [/ decade]	+0.80
C6.1-Aqua trend [/ decade]	-0.20
Δ trend [/ decade]	+1.00

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

4.2.4 Cloud Optical Thickness - Ice clouds

In this subsection the cloud optical thickness of ice clouds (COT_{ice}) of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS C6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-51), zonal mean (Figure 4-52) and standard deviation (Figure 4-53), all for a common time period, and time series plots (Figure 4-54). The COTs of the products refer to slightly different wavelengths¹.

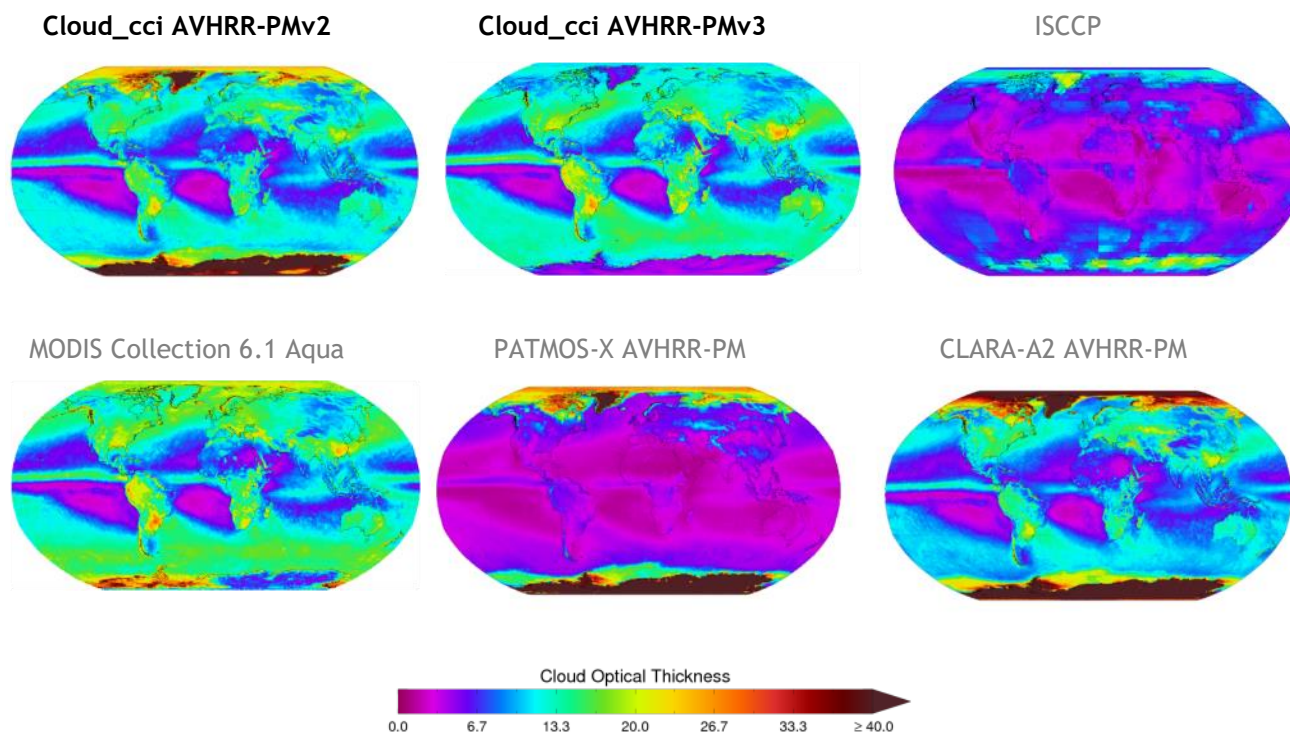


Figure 4-51 Globally gridded means of ice cloud optical thickness for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

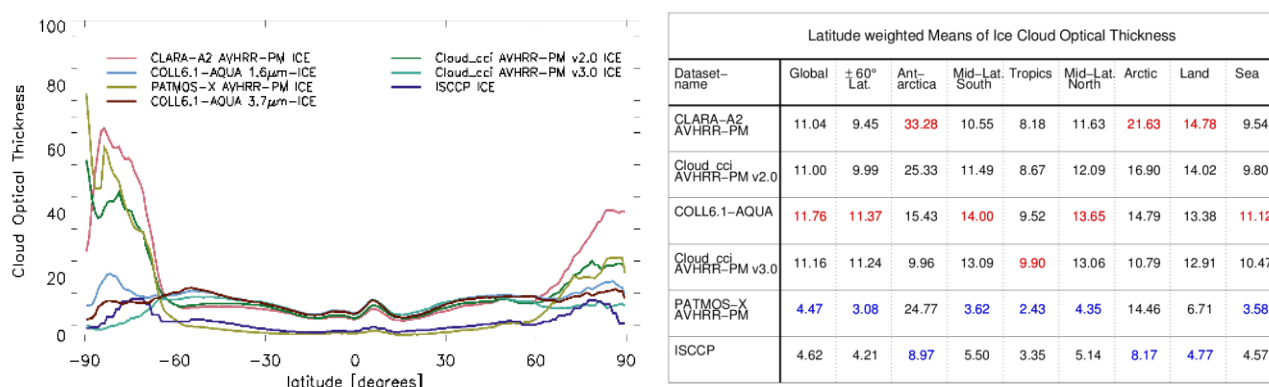



Figure 4-52 Cloud optical thickness - Ice from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 79

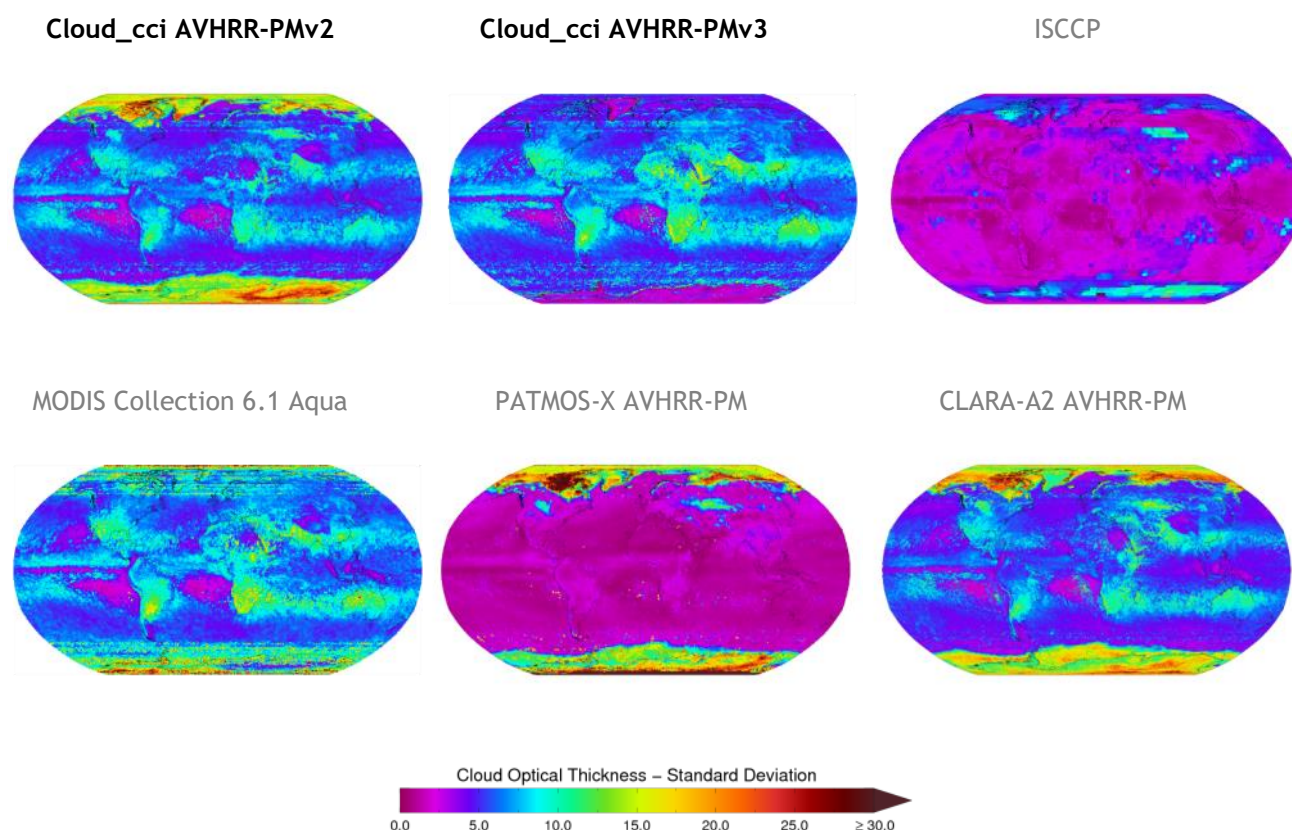



Figure 4-53 Globally gridded standard deviation of ice cloud optical thickness for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

General findings

- In contrast to Cloud_cci COT_{ice} is similar to MODIS C6.1 and CLARA-A2 for most parts of the globe, while differences to ISCCP and PATMOS-x are large. In all datasets, lower values are found in the subtropics, medium values in large parts of the mid- and higher latitude, extremely high values over the snow and ice covered surfaces in the polar regions, where also the spread among the datasets is highest (Fig. 4-51).
- Except PATMOS-x and ISCCP, all datasets agree on: Large temporal variability in COT_{ice} is found in the subtropics outside the stratocumulus regions in which the variability is lowest (Fig. 4-53). The variability is also relative low in the mid-latitudes. In ISCCP and PATMOS-x the temporal variability is very low compared to the other datasets. Polar regions are characterized by high variability in all datasets.
- As for COT_{liq} , the stability of the COT_{ice} time series of the AVHRR-based datasets suffer from the satellite drift of the early NOAA satellites, creating drifts and jumps for PATMOS-x, CLARA-A2 and Cloud_cci AVHRR-PM datasets. The MODIS C6.1 Aqua time series is characterized by a high stability and a small positive trend in the last years.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
		Page 80		

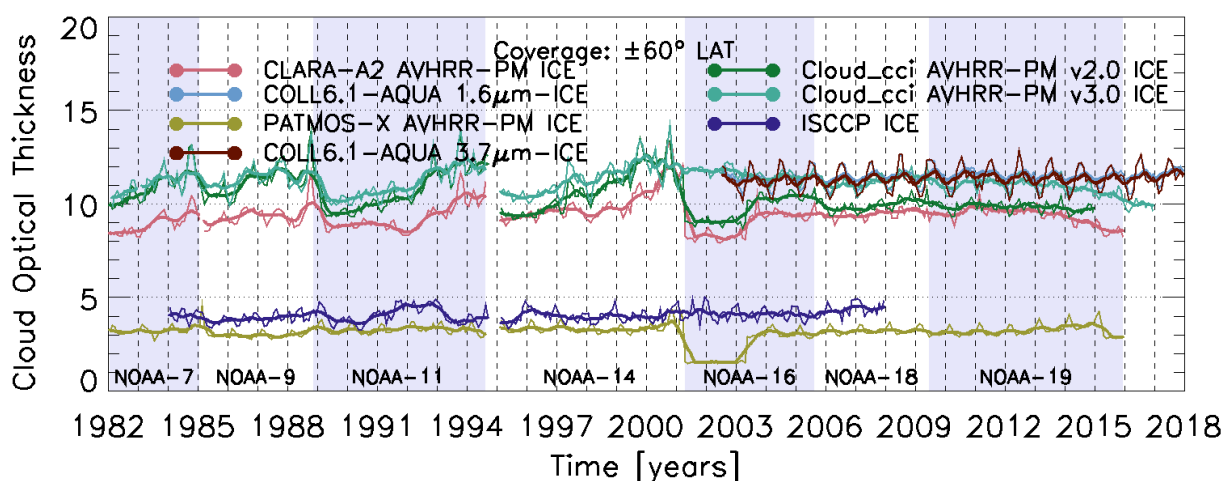


Figure 4-54 Time series of monthly ice cloud optical thickness for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6µm and 3.7µm products are shown.

Discussion

The cloud optical thickness of ice clouds COT_{ice} is very similar for the Cloud_cci dataset AVHRR-PMv2 and v3 compared to the reference datasets MODIS Collection 6.1 Aqua and CLARA-A2 between 60S and 60N. The spread towards the polar regions is high among all datasets, with the Cloud_cci AVHRR-PMv2 being somewhere in the middle, while CLARA-A2 shows highest and Cloud_cci AVHRR-PMv3 lowest COT_{ice} mean values there. Considering the average for 60S-60N, the Cloud_cci AVHRR-PMv2 dataset was 1 to 2 optical thicknesses lower than MODIS C6.1 Aqua, while v3 is very similar.

The time series plots exhibit a strong response of the Cloud_cci AVHRR-PM versions mean COT_{ice} to the satellite drift of NOAA-07, -09, -11 and -14, which also leads to strong jumps in the time series at the ties of satellites transitions. Note that the 1.6µm period in the beginning of NOAA-16 is no longer visible in Cloud_cci AVHRR-PMv3 while it was strongly pronounced in version 2.0. The inhomogeneities due to the satellite drifts appear a bit smaller for PATMOS-x compared to Cloud_cci-AVHRR-PMv3 and CLARA-A2, even though this dataset is also AVHRR-based. The reason for this is not entirely known yet, however, the generally smaller absolute values might attenuate the absolute values of the inhomogeneities. Cloud_cci AVHRR-PMv3 has no abrupt jump due to the beginning of NOAA-16, but a small constant decreasing trend from 2010 until 2016, which can possibly be attributed to satellite drift of NOAA-19, or else needs further research.

Table 4-17 presents evaluation scores inferred from comparisons of the Cloud_cci v3 dataset with MODIS C6.1 Aqua. Bias for Cloud_cci AVHRR-PMv3 to MODIS C6.1 Aqua is very similar (close to 0). The bc-RMSD values are relatively low (values around 5 optical thicknesses). While MODIS Collection 6.1 has a small positive decadal trend, there is a slight negative trend in Cloud_cci AVHRR-PMv3 of -0.3 per decade, which is not explained yet.

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55µm, CLARA-A2: 0.6µm, MODIS C6.1 Aqua: 0.65µm (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7µm on COT is assumed to be very minor.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 81

Table 4-17: Evaluation scores for Cloud_cci v3 Level-3C Optical Thickness - Ice (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias	+0.01
bc-RMSD	+4.70
Cloud_cci v3.0 trend [/ decade]	-0.29
C6.1-Aqua trend [/ decade]	+0.48
Δ trend [/ decade]	-0.77

¹Wavelengths the COT of the individual datasets refer to: Cloud_cci: 0.55 μ m, CLARA-A2: 0.6 μ m, MODIS C6.1 Aqua: 0.65 μ m (assumed to be the same as for C5), ISCCP: unknown. The impact of the representative wave length within 0.5-0.7 μ m on COT is assumed to be very minor.

4.2.5 Cloud Effective Radius - Liquid clouds

In this subsection the cloud effective radius of liquid clouds (CER_{liq}) of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS Collection 6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-55), zonal mean (Figure 4-56) and standard deviation (Figure 4-57), all for a common time period, and time series plots (Figure 4-58). ISCCP is only included for the time series plots as no ISCCP CER_{liq} data is available after the year 2000.

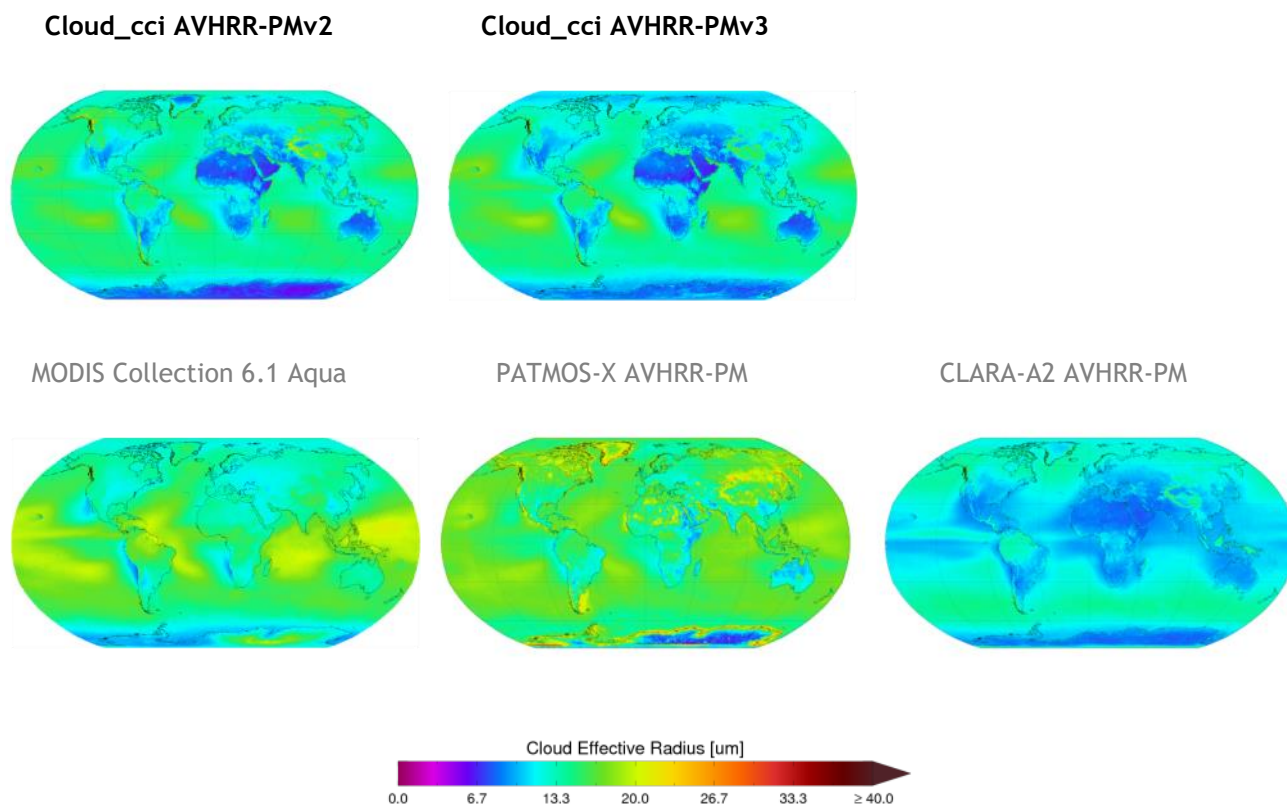


Figure 4-55 Globally gridded means of liquid cloud effective radius for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μm product is plotted.

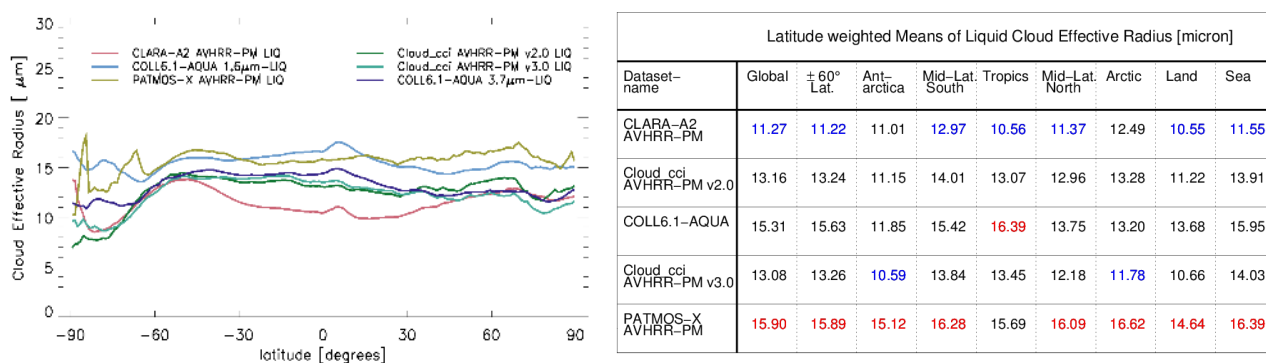

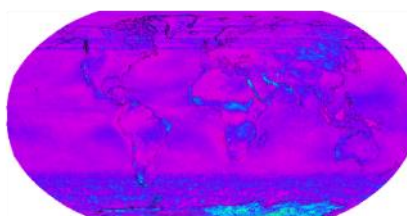
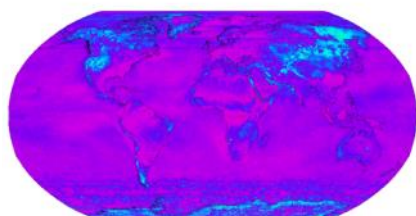


Figure 4-56 Cloud effective radius - Liquid from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μm and 3.7 μm products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μm product is used. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 83

Cloud_cci AVHRR-PMv2

Cloud_cci AVHRR-PMv3



MODIS Collection 6.1 Aqua

PATMOS-X AVHRR-PM

CLARA-A2 AVHRR-PM

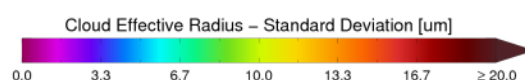
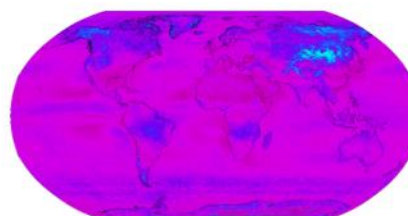
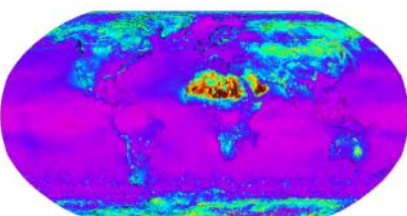
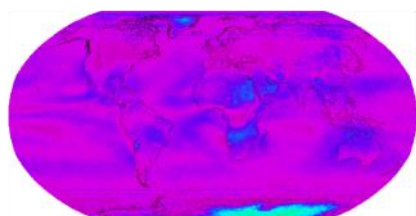



Figure 4-57 Globally gridded standard deviation of liquid cloud effective radius for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

General findings

- The mean cloud effective radius of liquid clouds varies a bit between all datasets with CLARA-A2 CER_{liq} being at the lower end and highest values found for PATMOS-x and MODIS Collection 6.1 Aqua for most parts of the globe. Cloud_cci AVHRR-PM datasets lie in between of the reference datasets but contain the lowest absolute values over many land regions in the subtropics. The spread among the datasets in the polar regions is for CER_{liq} comparable to the rest of the globe (Fig. 4-55).
- The temporal variability is general very low for all datasets, except over Antarctica and some land regions of which the Sahara regions clearly sticks out for PATMOS-x CER_{liq} (Fig. 4-57).
- The time series of the AVHRR-based datasets (and ISCCP) reveal again trends in the course of the lifetime of the early NOAA satellites as well as in homogeneities at satellite transitions resulting from these. In addition, the channel switch to 1.6 μ m in 2000 to 2002 is clearly visible in the AVHRR-based time series. This is reasonable due to the different penetration depths of the 3.7mic and 1.6mic signals.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 84

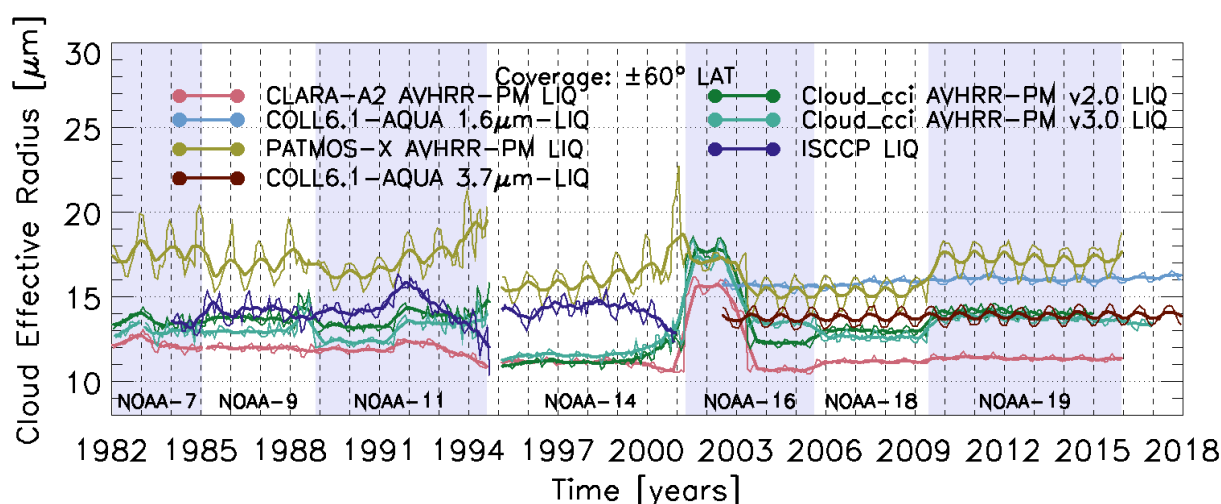


Figure 4-58 Time series of monthly liquid cloud effective radius for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown.


Discussion

The spread in CER_{liq} among all datasets is relative large (about 50% of the mean of all datasets) throughout all latitude bands. Thus, the Cloud_cci AVHRR-PMv3 dataset is difficult to judge based on these comparisons. One can at least conclude that the relative latitudinal dependence of the mean CER_{liq} in Cloud_cci AVHRR-PMv3 is similar to the reference data. For another feature one can also find an agreement between all datasets: the lower CER_{liq} in the stratocumulus regions, especially near the coasts, compared to the outer regions of the stratocumulus further westwards and trade cumulus regions. The Cloud_cci data shows a rather strong land-sea contrast in mean CER_{liq} (lower values over land compared to sea) which is not reflected in the reference datasets to the same extent.

The time series of Cloud_cci AVHRR-PMv3 shows some peculiarities: relative constant values for certain periods, but significant jumps at the times of satellite transitions. The low values for NOAA-14 are very unusual. The higher values at the beginning for NOAA-16 are most likely due to the 1.6 μ m channel being available and used, compared to the 3.7 μ m channel available and used at all other times.

Table 4-18 presents evaluation scores inferred from comparison of the Cloud_cci v3 dataset with MODIS C6.1 Aqua. Compared to the 3.7 μ m CER_{liq} of MODIS C6.1 Aqua, Cloud_cci AVHRR-PMv3 has a small negative bias. The bc-RMSD values are also very low with values of 2.1 μ m. Cloud_cci AVHRR-PMv3 and MODIS C6.1 Aqua do not show a significant trend per decade, but being both slightly positive. The time series plots indicate that sudden increases (jumps) from NOAA-16 to NOAA-18 and from NOAA-18 to NOAA-19 are probably responsible for the small trend found in Cloud_cci AVHRR-PMv3.

Table 4-18: Evaluation scores for Cloud_cci v3 Level-3C Effective Radius - Liquid (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 85

dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias [μm]	-0.51
bc-RMSD [μm]	+2.10
Cloud_cci v3.0 trend [μm / decade]	+0.09
C6.1-Aqua trend [μm / decade]	+0.18
Δ trend [μm / decade]	-0.10

4.2.6 Cloud Effective Radius - Ice clouds

In this subsection the cloud effective radius of ice clouds (CER_{ice}) of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS C6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-59), zonal mean (Figure 4-60) and standard deviation (Figure 4-61), all for a common time period, and time series plots (Figure 4-62). ISCCP is only included for the time series plots as no ISCCP CER_{ice} data is available after the year 2000.

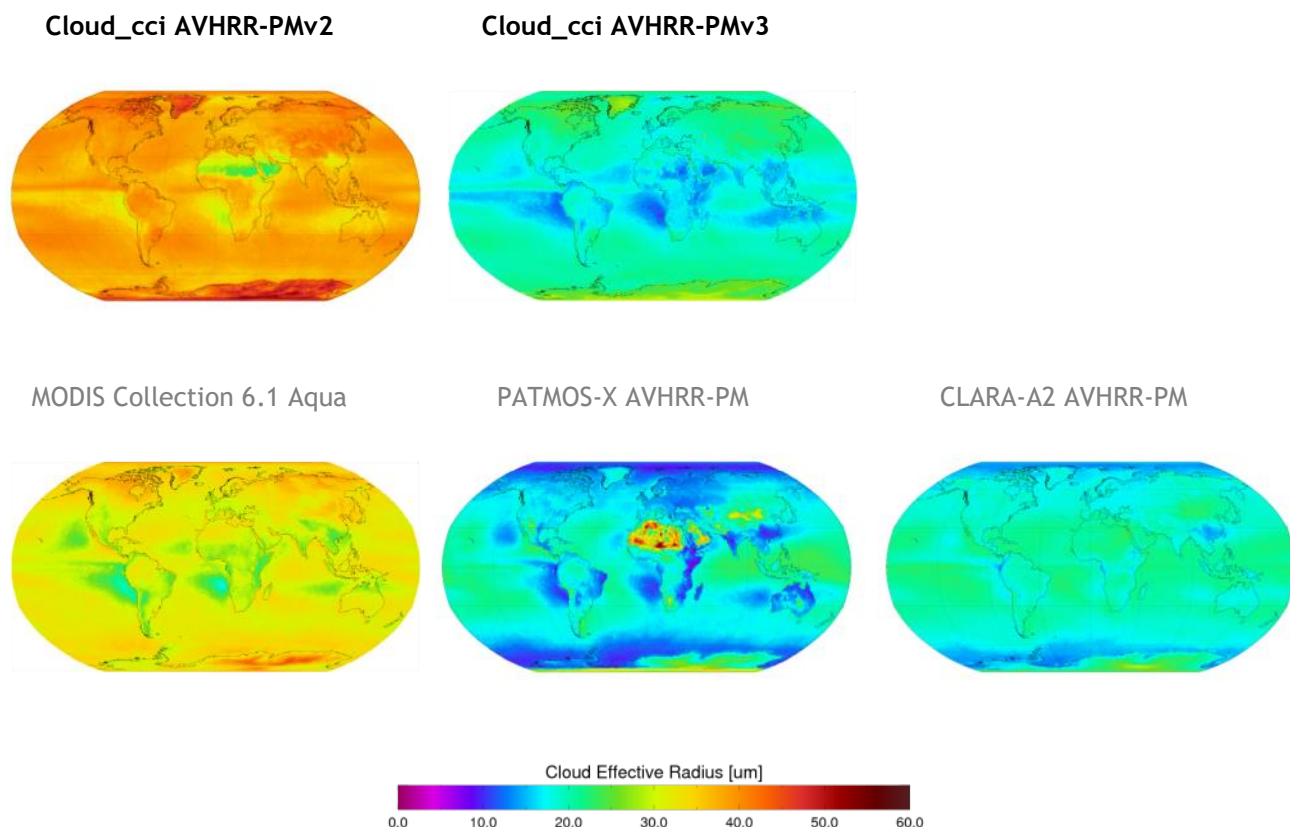


Figure 4-59 Globally gridded means of ice cloud effective radius for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μm product is plotted.

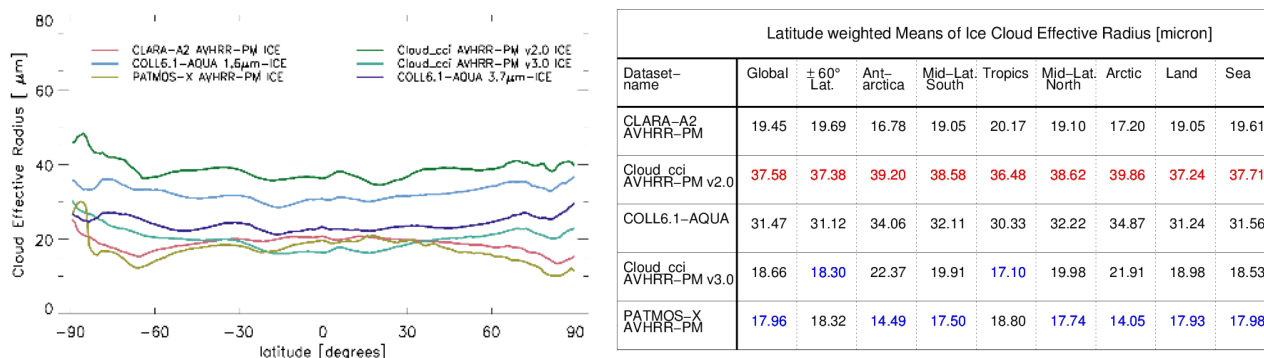



Figure 4-60 Cloud effective radius - Ice from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μm and 3.7 μm products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μm product is used. **Highest** and **lowest** values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 87

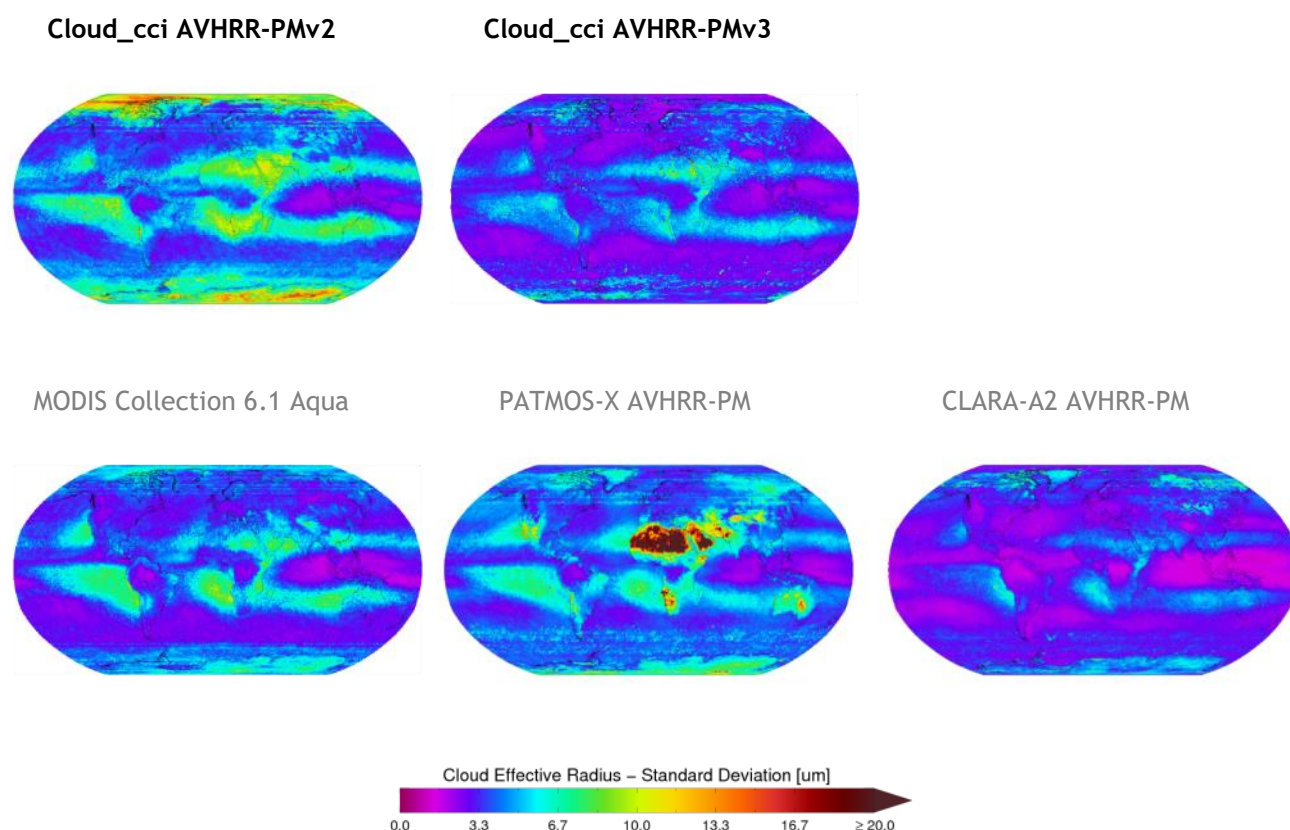


Figure 4-61 Globally gridded standard deviation of ice cloud effective radius for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μm product is plotted.

General findings

- Very high mean CER_{ice} were found for the Cloud_cci AVHRR-PMv2 dataset, while much lower values are found for Cloud_cci AVHRR-PMv3 now agreeing much better with PATMOS-x and CLARA-A2 (Fig. 4-59). No strong latitudinal dependence is found for mean CER_{ice} .
- Highest temporal variabilities of CER_{ice} are found in the subtropics, especially in the stratocumulus regions, which are characterized by very low ice clouds frequencies of occurrences. The temporal variability is, as the mean values, highest for Cloud_cci AVHRR-PMv2 as well as for PATMOS-x (Fig. 4-61). The variability is relatively low for AVHRR-PMv3 and CLARA-A2.
- The MODIS-based datasets are stable in time throughout the MODIS era. AVHRR-based datasets show again jumps at the satellite transitions of the early, significantly drifting NOAA satellite. Another feature is found in the period of 2000 to 2002 for which the 1.6 μm channel was switched on and used from NOAA-16, while the rest of the time period always the 3.7 μm was switched on and thus used.

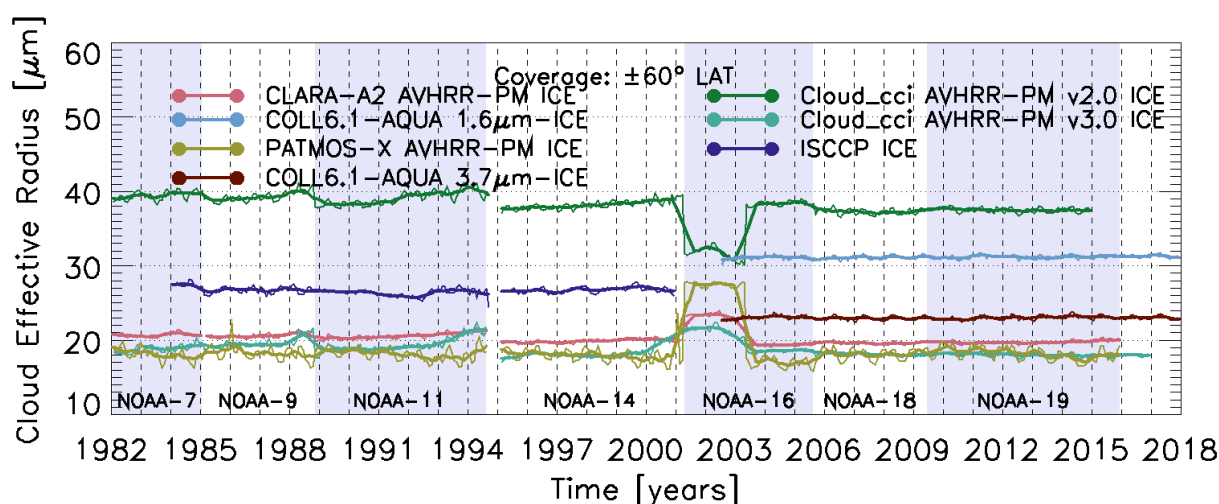


Figure 4-62 Time series of monthly ice cloud effective radius for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6 μm and 3.7 μm products are shown.

Discussion

The spread in CER_{ice} among the datasets is even larger than for CER_{liq} which prevents drawing too many conclusions from this intercomparison. On the other hand, as already mentioned when comparing the morning satellite datasets, the CC4CL scheme used wrong LUTs for ice cloud properties in Cloud_cci AVHRR-PMv2, which led to a very strong overestimation of CER_{ice} . This is also visible in this section for Cloud_cci AVHRR-PMv2 showing mean CER_{ice} values of about 40 μm . Keeping this in mind, there was only little confidence in the quality of CER_{ice} in AVHRR-PMv2. The corresponding bug has been fixed now, which led to CER_{ice} of much better quality in Cloud_cci AVHRR-PM dataset version v3 and a higher agreement with the reference datasets.

Table 4-19 presents evaluation scores inferred from the comparison of the Cloud_cci v3 dataset with MODIS C6.1 Aqua. For Cloud_cci AVHRR-PMv3 a relative small negative bias around -5 μm is found when compared against MODIS C6.1 Aqua. The bc-RMSD value is moderate with values around 4 to 5 μm . Cloud_cci AVHRR-PMv3 shows a relative large negative trend of -1.03 $\mu\text{m} / \text{decade}$, which is mainly caused by a few months of 1.6 μm CER_{ice} retrievals being included in the beginning of the common 2003-2011 time period (see Figure 4-62). Apart from that, all time series look very stable with only small jumps at satellite transitions for Cloud_cci AVHRR-PM version 3.0 and the other AVHRR-based datasets.

Table 4-19: Evaluation scores for Cloud_cci v3 Level-3C Effective Radius - Ice (afternoon satellites) based on comparison to MODIS Collection 6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias [μm]	-4.66
bc-RMSD [μm]	+4.10
Cloud_cci v3.0 trend [$\mu\text{m} / \text{decade}$]	-1.03
C6.1-Aqua trend [$\mu\text{m} / \text{decade}$]	-0.26
Δ trend [$\mu\text{m} / \text{decade}$]	-0.77

4.2.7 Liquid Water Path

In this subsection the cloud liquid water path (LWP) of Cloud_cci AVHRR-PMv2 and v3, CLARA-A2 (afternoon satellites only), MODIS Collection 6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-63), zonal mean (Figure 4-64) and standard deviation (Figure 4-65), all for a common time period, and time series plots (Figure 4-66 and Figure 4-67).

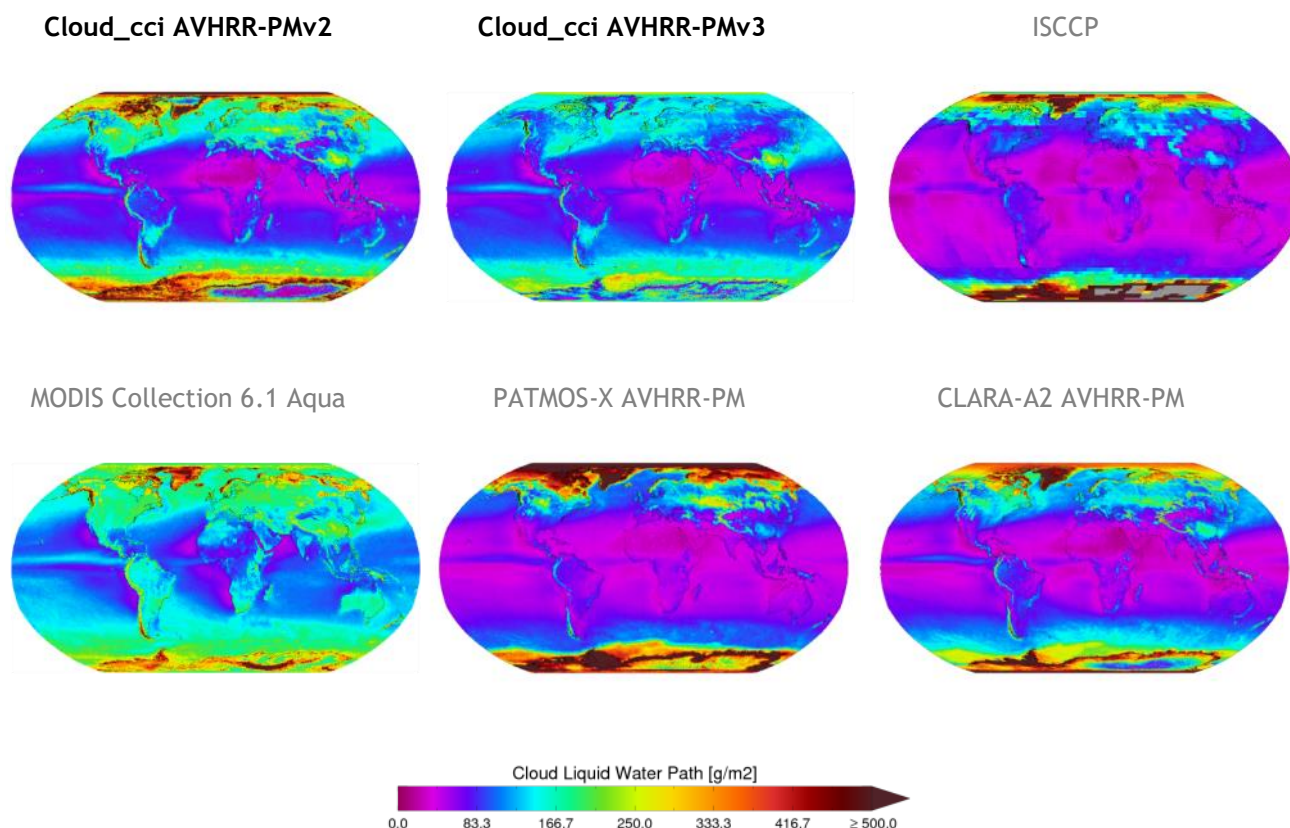


Figure 4-63 Globally gridded means of cloud liquid water path for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μm product is plotted.

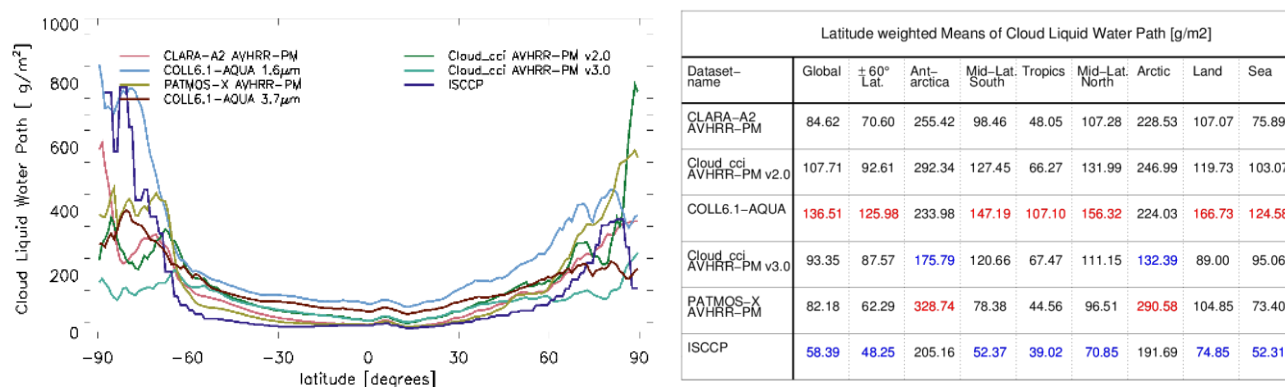



Figure 4-64 Cloud liquid water path from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μm and 3.7 μm products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μm product is used. Highest and lowest values of each region are highlighted.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 90

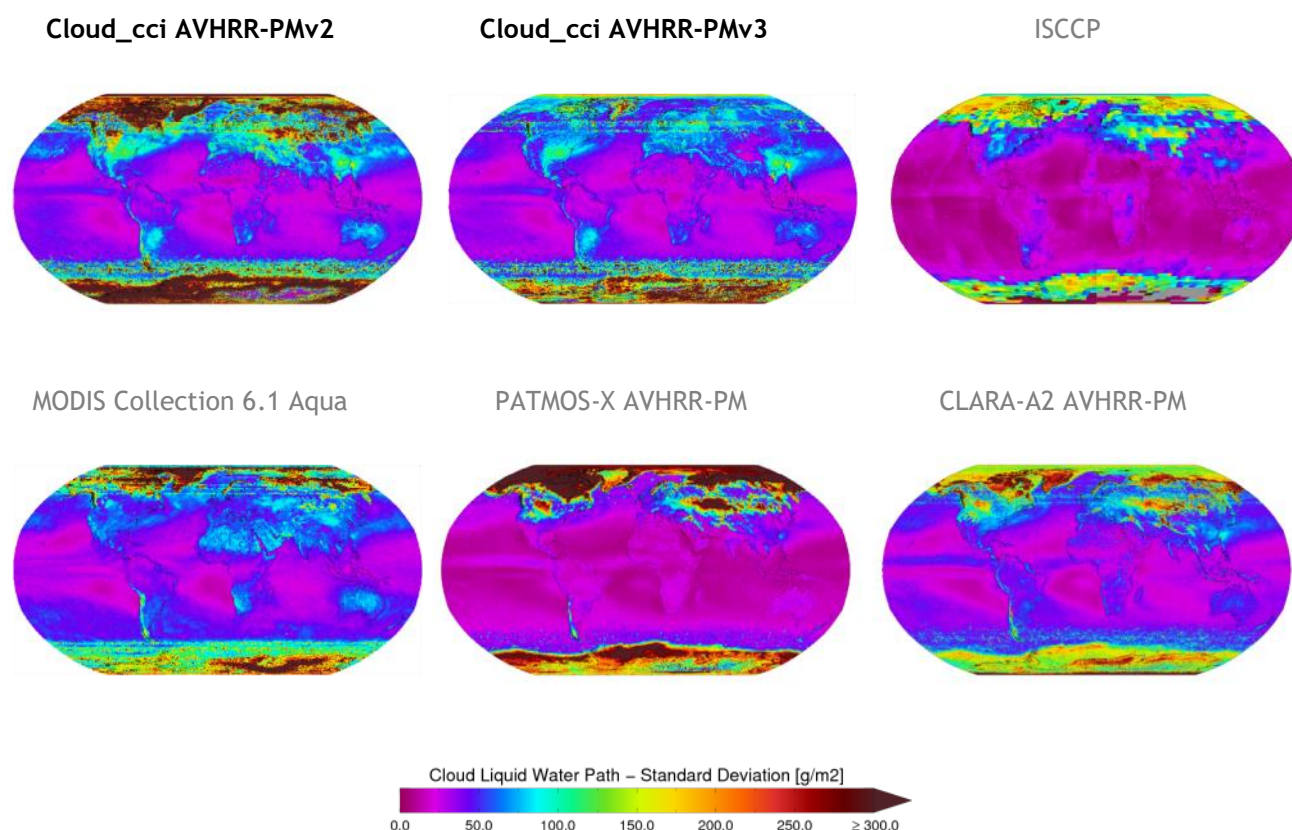


Figure 4-65 Globally gridded standard deviation of cloud liquid water path for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

General findings

- The mean LWP in all datasets shows similar spatial patterns with generally lower values in the tropics (a small local maximum near the ITCZ), increasing towards the mid- and high latitudes and (partly) reaching very high values in the polar regions, where also the spread among the datasets is highest (Fig. 4-63). Apart from the polar regions, MODIS C6.1 Aqua exhibits highest and ISCCP lowest mean values among all datasets.
- The temporal variability of LWP shows similar pattern: low in the tropics and subtropics and high in the high latitudes (Fig. 4-65).
- Time series of mean LWP show a significant seasonal cycle. The stability of the LWP time series of Cloud_cci AVHRR-PMv3 and v2, CLARA-A2 and PATMOS-x are again affected changing local observations time among the NOAA satellites and the by satellite drift of the early NOAA satellites that does enlarge this effect. For the later AVHRR-carrying satellites and MODIS, the time series seem relative stable. In the all-sky time series plots the spread among the datasets is reduced. Also, the channel switching within the AVHRR-based times series is less pronounced.

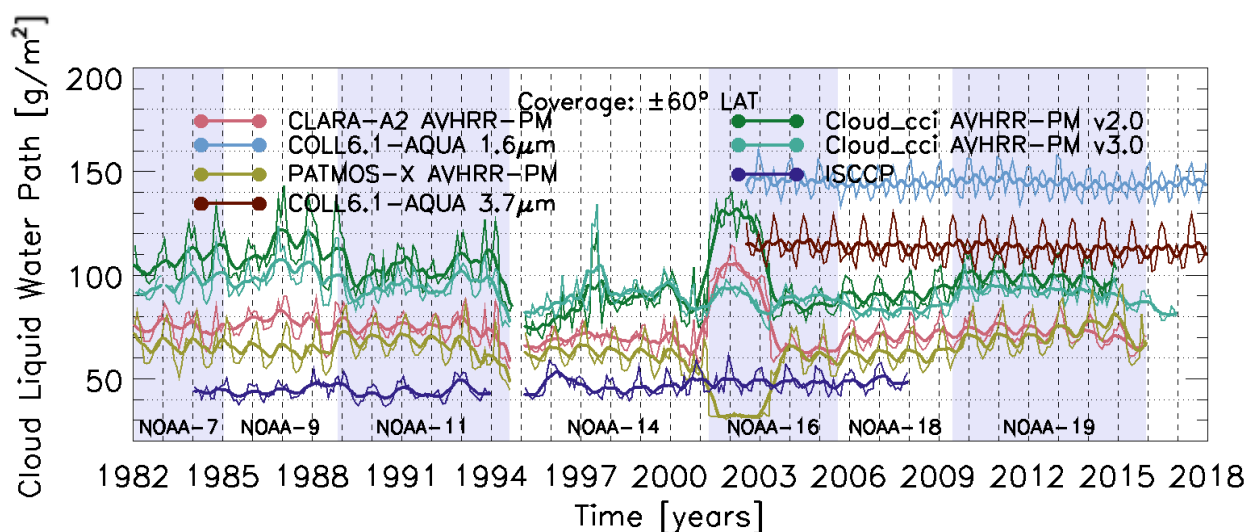


Figure 4-66 Time series of monthly “in-cloud” cloud liquid water path for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown.

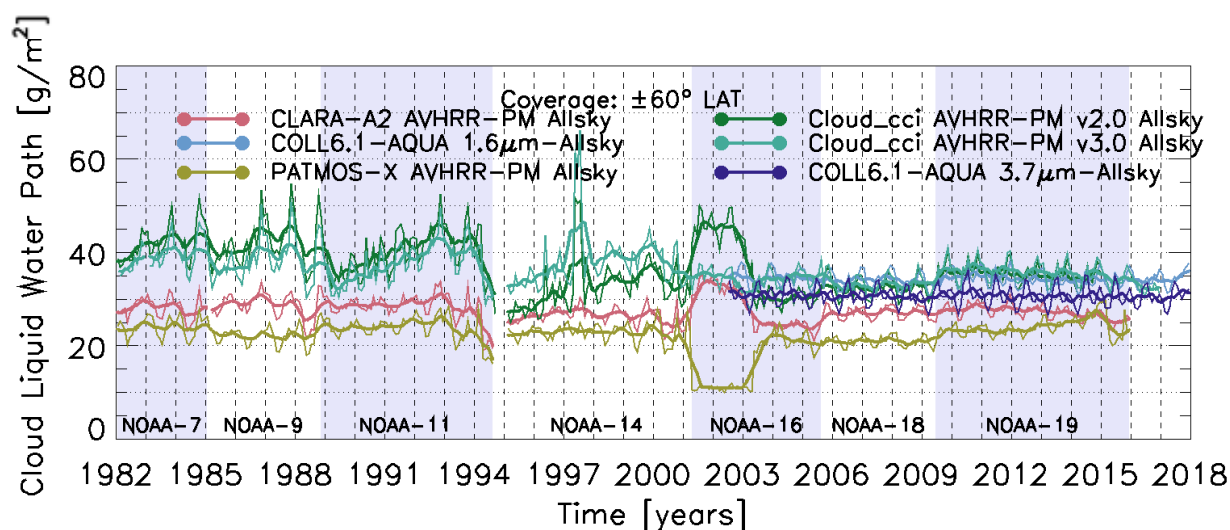



Figure 4-67 Time series of monthly “all-sky” cloud liquid water path for all afternoon satellite retrievals. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 92

Discussion

For a large portion of the globe (60S-60N) the Cloud_cci v3 mean LWP lies in the middle between MODIS C6.1 Aqua data (highest) and ISCCP and PATMOS-x (lowest). As Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3 agrees with the other datasets on increasing in-cloud LWP from tropical regions towards the mid-latitudes. They also agree on smaller LWP values in the eastern parts of the stratocumulus regions, compared to the western parts and compared to trade cumulus regions. Another regional feature, the LWP ‘hotspot’ in South-east Asia is well represented in the Cloud_cci AVHRR-PMv3 dataset.

The time series plots again present the difficulties that the drifting, early satellites cause in AVHRR-PM datasets. LWP values increase with life time and ‘fall back’ at the transition to the next sensor. However, it should be possible to increase the stability by applying a proper drift correction. Also found is the 2 year period in which the NOAA-16 1.6 μ m channel is switched on and used, causing significantly higher mean LWP values in this period. In the all-sky time series plot, all datasets seem to agree a bit better, which indicates that difference in CFC play a significant role in the in-cloud comparisons.

Table 4-20 presents evaluation scores inferred from the comparison of Cloud_cci AVHRR-PMv3 with MODIS C6.1 Aqua. The bias with respect to C6.1 Aqua for Cloud_cci AVHRR-PMv3 is negative and relative moderate. The bc-RMSD value is similar around 52 g/m². While the MODIS C6.1 Aqua time series shows a slight negative trend, Cloud_cci AVHRR-PMv3 has a clear positive trend, which is mainly caused by the significant jump for LWP between NOAA-18 and NOAA-19.

Table 4-20: Evaluation scores for Cloud_cci v3 Level-3C Liquid Water Path (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias [g/m ²]	-23.73
bc-RMSD [g/m ²]	+52.68
Cloud_cci v3.0 trend [g/m ² / decade]	+9.78
C6.1-Aqua trend [g/m ² / decade]	-0.99
Δ trend [g/m ² / decade]	+10.77

4.2.8 Ice Water Path

In this subsection the cloud ice water path (IWP) of Cloud_cci AVHRR-PMv2, Cloud_cci AVHRR-PMv3, CLARA-A2 (afternoon satellites only), MODIS C6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-68), zonal mean (Figure 4-69) and standard deviation (Figure 4-70), all for a common time period, and time series plots (Figure 4-71 and Figure 4-72).

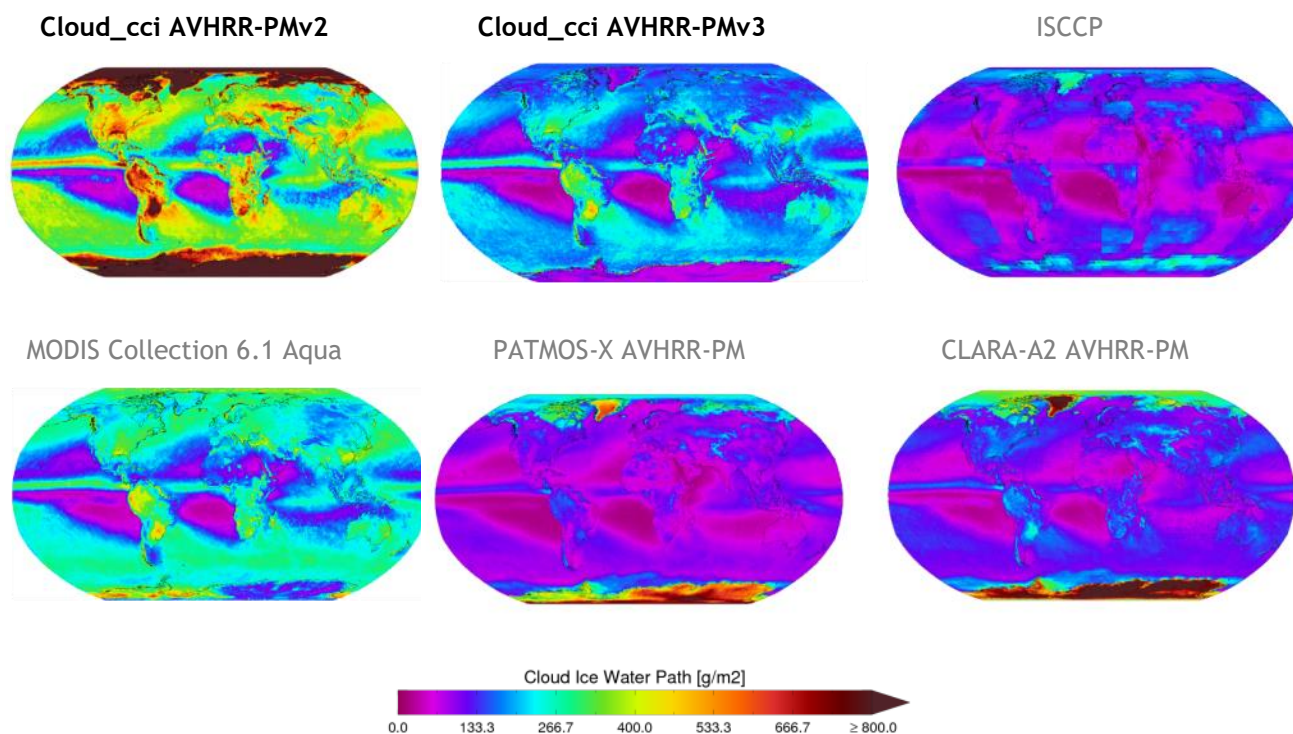


Figure 4-68 Globally gridded means of cloud ice water path for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

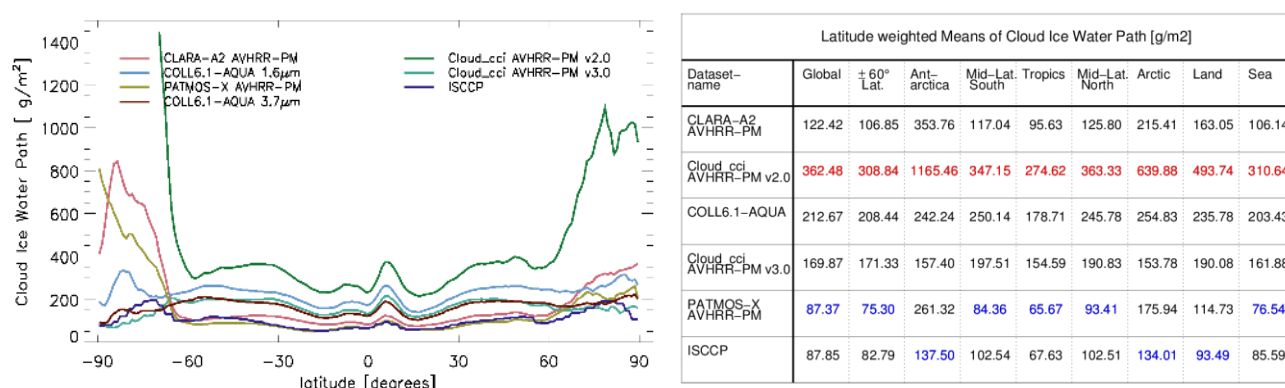



Figure 4-69 Cloud Ice water path from 2003 to 2011. Left: Zonal means. For MODIS Collection 6.1 Aqua the 1.6 μ m and 3.7 μ m products are shown. Right: Latitude weighted means for different regions. For MODIS Collection 6.1 Aqua the 2.1 μ m product is used. Highest and lowest values of each region are highlighted.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 94

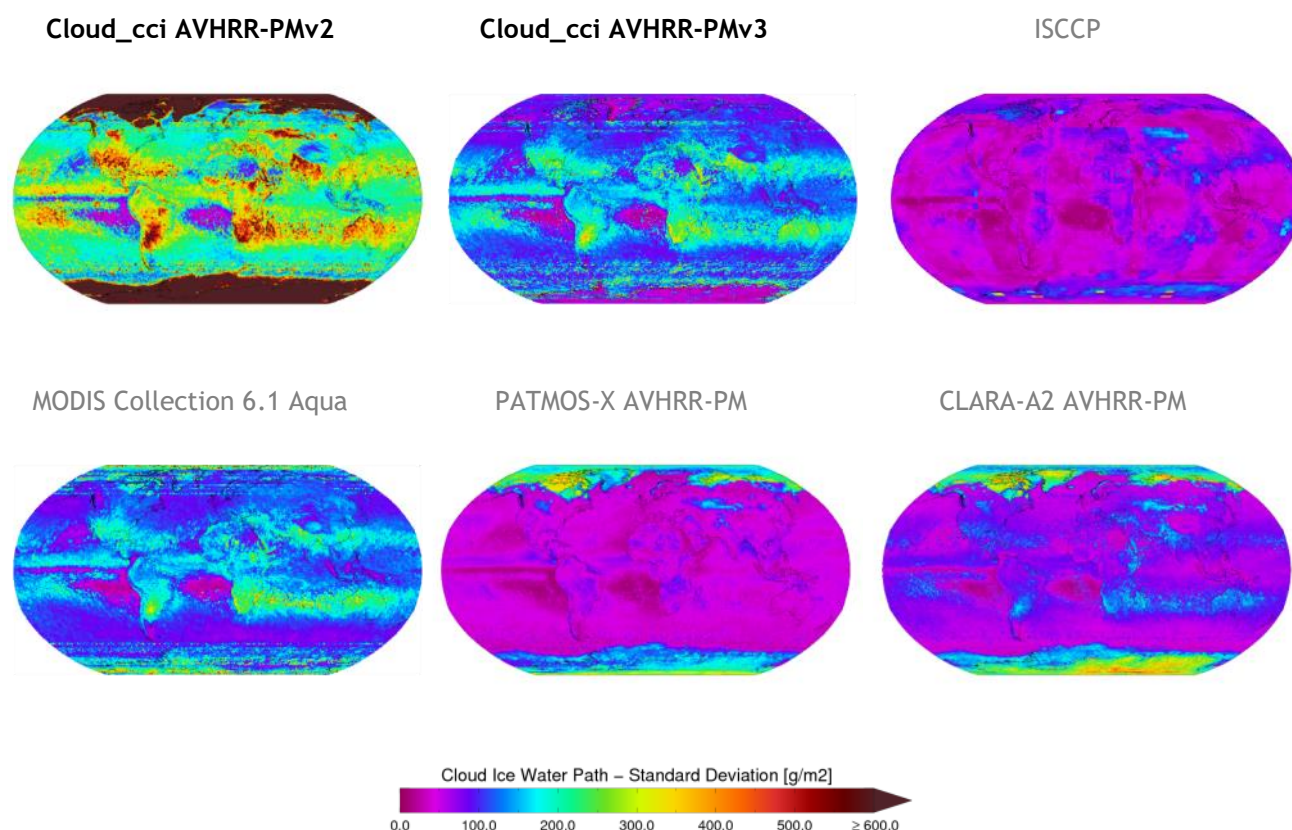


Figure 4-70 Globally gridded standard deviation of cloud ice water path for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey. For MODIS Collection 6.1 Aqua the 2.1 μ m product is plotted.

General findings

- The mean IWP is high for in the ITCZ for all datasets, decreasing in the subtropics and increasing again towards the mid-latitudes with higher values over land (Fig. 4-68). Polar regions partly show extremely high values, in particular the Cloud_cci AVHRR-PMv2 dataset, which also have highest values among the datasets for the rest of the globe.
- Apart from the polar regions, highest temporal variability for IWP is found tropical and subtropical land regions (Fig. 4-70).
- In the time series plots, Cloud_cci AVHRR-PMv2 exhibits a very high sensitivity to the local observation time creating trends and jumps in IWP during the life time of the early NOAA satellites and at their transitions. This effect is smaller (possibly due to the significantly lower absolute values for IWP) for CLARA-A2, ISCCP and PATMOS-x. In Cloud_cci AVHRR-PMv3 these trends and jumps are also much less distinctive. A significant jump is found in 2000 to 2002 for all AVHRR datasets because of the NIR channel switch of NOAA-16 (1.6 μ m instead of 3.7 μ m).

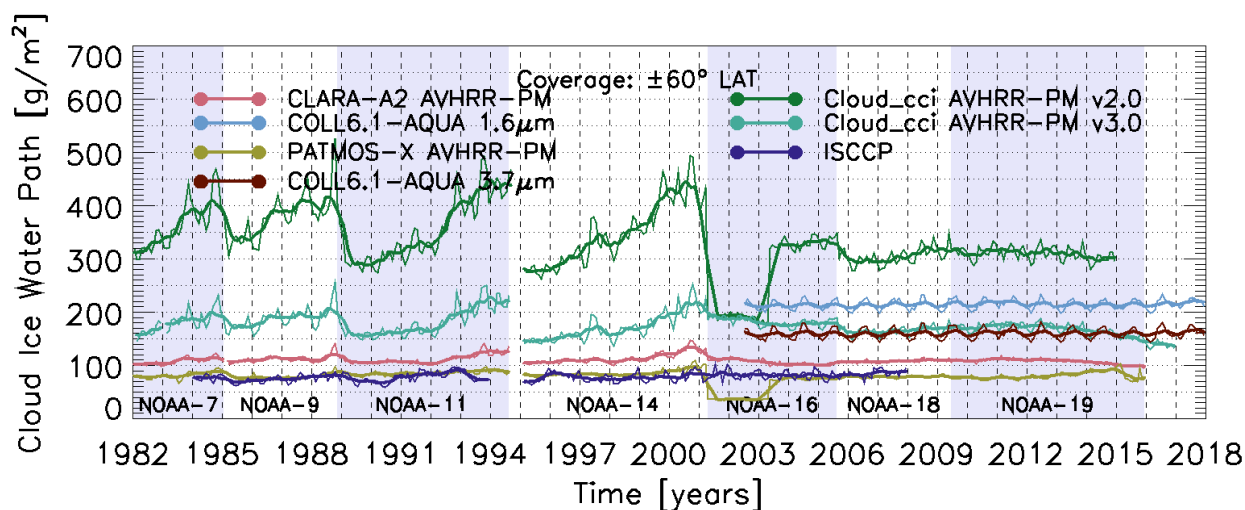


Figure 4-71 Time series of monthly “in-cloud” cloud ice water path for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6µm and 3.7µm products are shown.

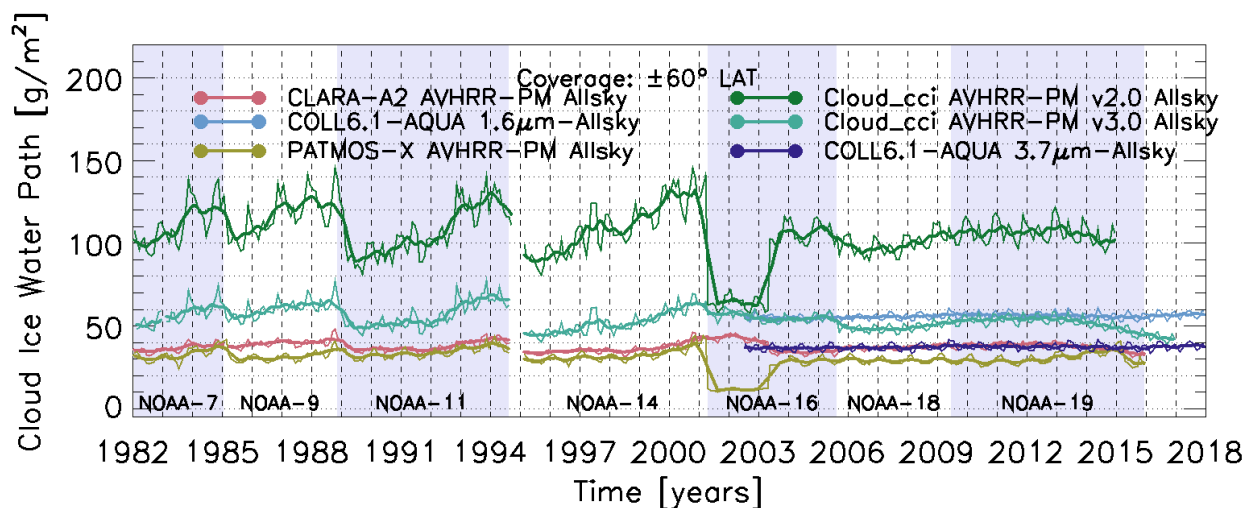



Figure 4-72 Time series of monthly “all-sky” cloud ice water path for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines. For MODIS Collection 6.1 Aqua the 1.6µm and 3.7µm products are shown.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 96

Discussion

Although the Cloud_cci AVHRR-PMv2 dataset reflected the spatial pattern of mean IWP, it presented values that were about 2 times higher than the reference datasets. Since this was mainly due to the LUT bug described before, that caused in particular too high CER_{ice} . As IWP is diagnosed using COT_{ice} and CER_{ice} this overestimation had a direct effect on the IWP in Cloud_cci data v2. For Cloud_cci AVHRR-PMv3 this bug is fixed and the resulting mean IWP in AVHRR-PMv3 shows lower values with many similarities to MODIS Collection 6.1 Aqua.

The drift of the early NOAA satellites has a significant impact on the stability of the Cloud_cci AVHRR-PM datasets, but is higher for Cloud_cci AVHRR-PMv2. In version 3 the stability and agreement with other datasets is noticeably better.

Table 4-21 presents evaluation scores inferred from the comparison of Cloud_cci AVHRR-PMv3 with MODIS C6.1 Aqua. For Cloud_cci AVHRR-PMv3 (using the $3.7\mu m$ channel and is compared against MODIS C6.1 Aqua $3.7\mu m$) a positive bias is found. The bc-RMSD value lies between 80 and $85 g/m^2$. Like MODIS C6.1 Aqua $3.7\mu m$ data has nearly no trend in the considered period, Cloud_cci AVHRR-PMv3 shows also only small trends, but with a negative sign. For Cloud_cci AVHRR-PMv3 it is possibly due to drifts of the early NOAA satellites, while trends in MODIS C6.1 Aqua have to be further researched. Overall, it is visible, that the LUT bug of Cloud_cci AVHRR-PMv2 (Bias: $+149 g/m^2$; bc-RMSD: $+183 g/m^2$) is fixed in version 3.0 resulting in higher stability and moderate biases.

Table 4-21: Evaluation scores for Cloud_cci v3 Level-3C Ice Water Path (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM _{3.7}
Bias [g/m^2]	+12.28
bc-RMSD [g/m^2]	+81.95
Cloud_cci v3.0 trend [$g/m^2 / decade$]	-2.17
C6.1-Aqua trend [$g/m^2 / decade$]	+4.49
Δ trend [$g/m^2 / decade$]	-6.66

4.2.9 Liquid Cloud Fraction

In this subsection the liquid cloud fraction LCF of Cloud_cci AVHRR-PMv2 and v3, CLARA-A2 (afternoon satellites only), MODIS Collection 6.1 Aqua, ISCCP and PATMOS-x (afternoon satellites only) are compared by means of multi-annual mean (Figure 4-73), zonal mean (Figure 4-74) and standard deviation (Figure 4-75), all for a common time period, and time series plots (Figure 4-76).

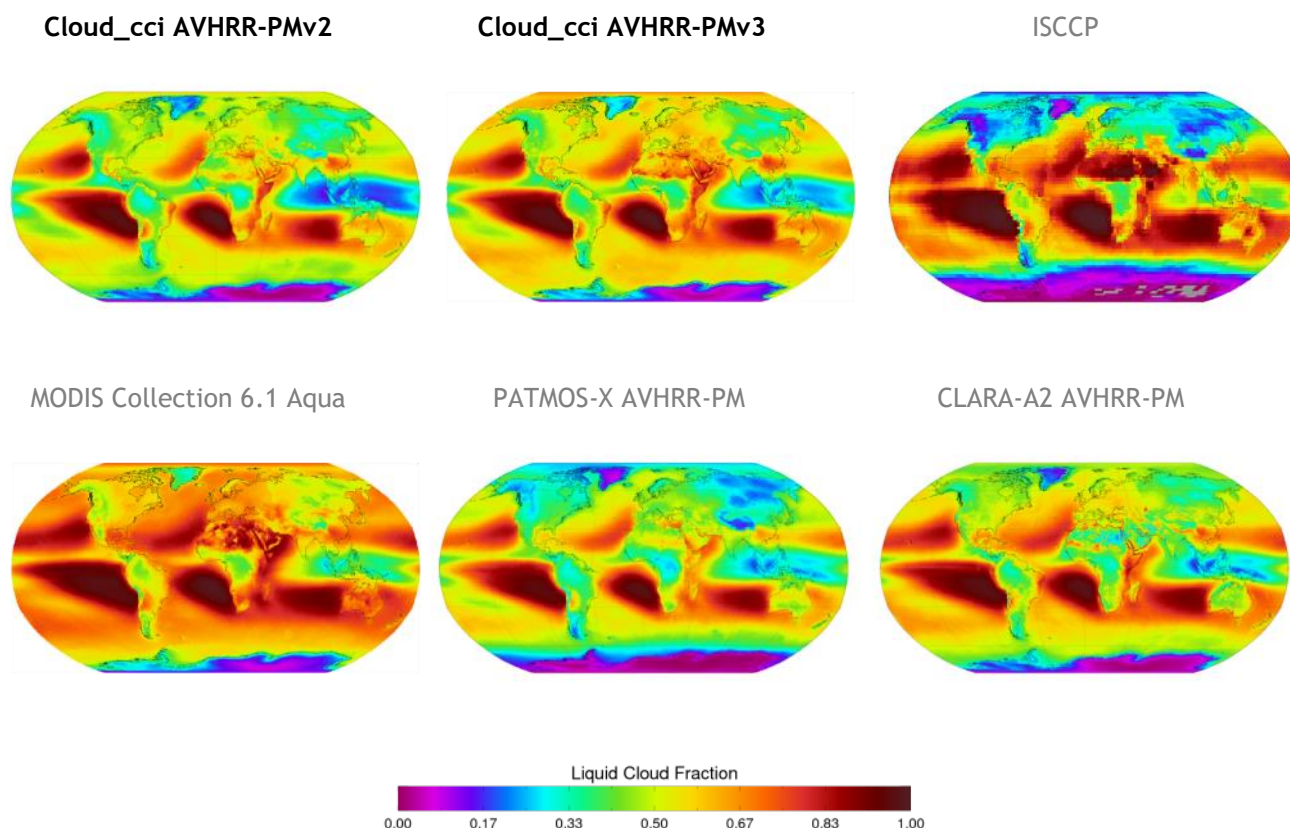


Figure 4-73 Globally gridded means of liquid cloud fraction for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

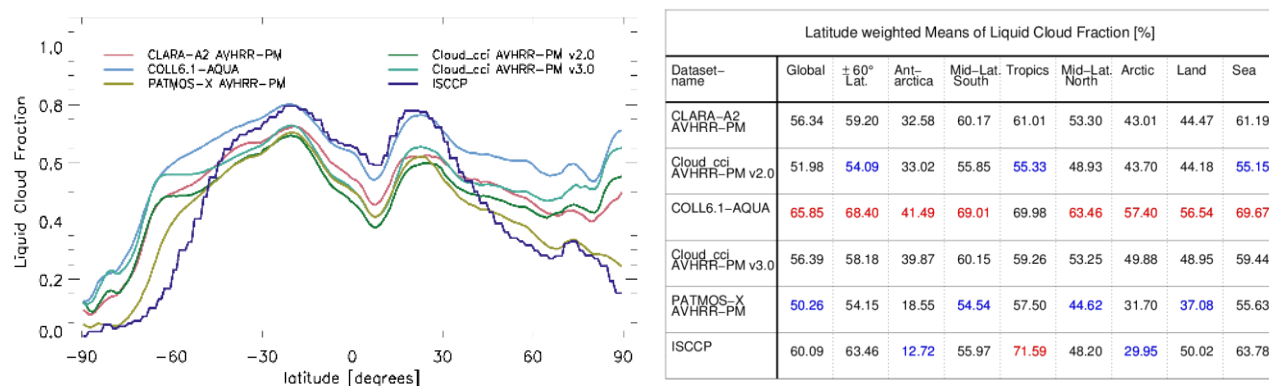


Figure 4-74 Liquid cloud Fraction from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. Highest and lowest values of each region are highlighted.

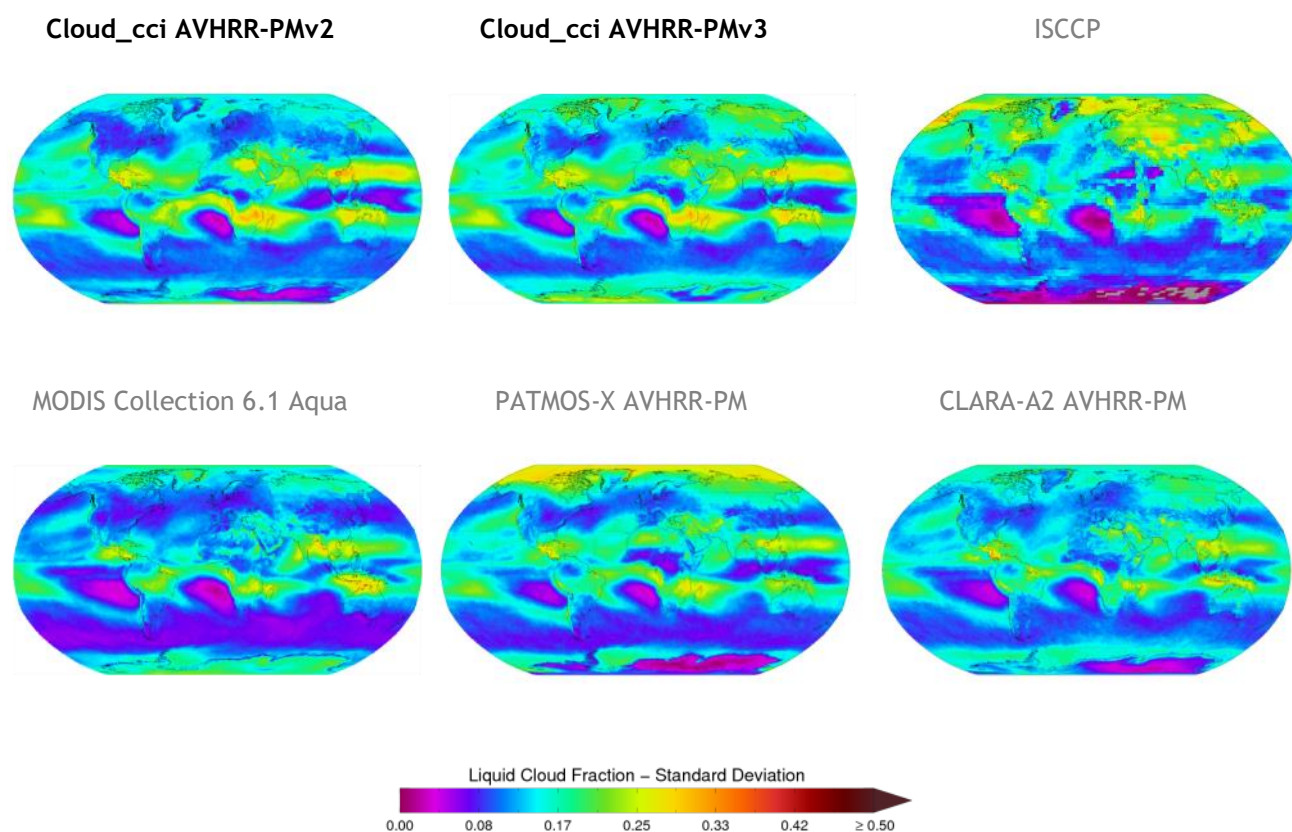


Figure 4-75 Globally gridded standard deviation of liquid cloud fraction for all afternoon satellite retrievals averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- The LCF of all datasets show the characteristic patterns with lower LCF in the inner tropics and very high LCF for the stratocumulus regions (Fig. 4-73). The LCF is decreasing again towards the higher latitudes. Extremely low values are found for Antarctica. Highest LCF is found for MODIS C6.1 Aqua, lowest for PATMOS-x and ISCCP outside the tropics and for Cloud_cci AVHRR-PMv2 inside the tropics.
- The stratocumulus regions are also the regions with the lowest temporal variability in LCF (Fig. 4-75). Relative large variability is found in the outer tropics / inner subtropics over land and over oceans without persistent stratocumulus regimes. The temporal variability in the mid-latitudes is relative low. All datasets seem to agree more or less on these findings.
- The time series plots of LCF show significant seasonal cycles in the global (60S-60N) mean in all datasets. Cloud_cci AVHRR-PMv2 shows largest discontinuities among the AVHRR-based datasets for the early NOAA satellites.

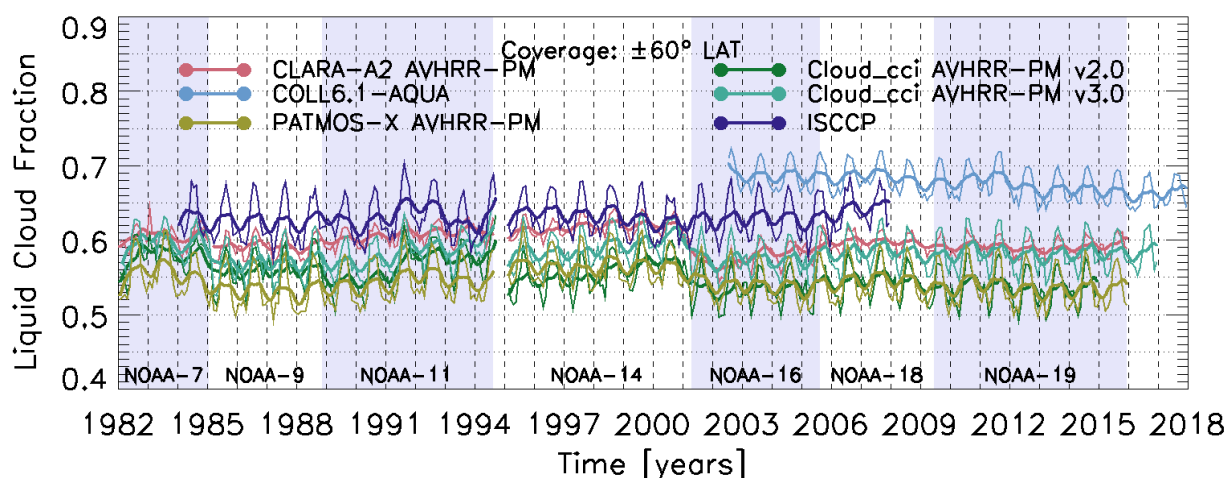


Figure 4-76 Time series of monthly liquid cloud fraction for all afternoon satellite retrievals (thin lines). Running averages with a time window of 8 months are shown as bold lines.


Discussion

The LCF of Cloud_cci AVHRR-PMv3 shows, similar to the morning satellite datasets, a high level of agreement with the reference data, in particular in relative terms, e.g. global spatial patterns. In terms of absolute LCF values, MODIS C6.1 Aqua has about 20% higher LCF nearly everywhere except in the southern high latitudes; ISCCP has lowest LCF anywhere outside the tropical region values. In terms of long-term stability, Cloud_cci AVHRR-PMv3 is always within 5%, however with slight, short-term trends for the early NOAA satellites and two positive anomalies related to the volcano eruptions of El Chichón (1982) and Pinatubo (1991), for which also positive CFC and CTP anomalies were found. Falsely detected low clouds would cause positive LCF anomalies, which indicate that these two anomalies are artefacts. A Pinatubo eruption related anomaly can also be found for ISCCP around 1991.

Table 4-22 presents evaluation scores inferred from the comparison of Cloud_cci AVHRR-PMv3 with MODIS C6.1 Aqua. As seen for MODIS C6.1 Terra, MODIS C6.1 Aqua also has a very high LCF, significantly higher than Cloud_cci or other AVHRR-based datasets. The bias is around -10 %. The bc-RMSD value amounts to 9.3 %. Trend for Cloud_cci AVHRR-PMv3 is small and positive; the trend for MODIS Collection 6.1 Aqua is small and negative. All in all, the datasets are relative stable and comparable.

Table 4-22: Evaluation scores for Cloud_cci v3 Level-3C Liquid Cloud Fraction (afternoon satellites) based on comparison to MODIS C6.1 Aqua from 2003-2011. The scores were calculated separately for the Cloud_cci dataset by including all valid data points pairwise in the MODIS and the Cloud_cci dataset. This can introduce some small variability in the trend values for MODIS.

Measure	Cloud_cci v3.0 dataset
	AVHRR-PM
Bias [%]	-10.21
bc-RMSD [%]	+9.28
Cloud_cci v3.0 trend [% / decade]	+1.69
C6.1-Aqua trend [% / decade]	-0.37
Δ trend [% / decade]	+2.06

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 100

5. Intercomparison of Cloud_cci radiation products with other radiation datasets

In this section the Cloud_cci radiation datasets are compared to other well-established datasets. The basis for all comparisons are monthly mean values (Level-3C in ESA CCI notation).

The comparisons are further stratified into the individual radiative flux components. For each of the coming subsections, global maps of multi-annual means (of a common time period) and multi-annual (monthly) standard deviations are shown, together with zonal mean plots. In addition, time series plots of latitude-weighted global mean values (for a latitude band of 60S-60N) are given without restriction to the common time period. Using CERES EBAF-TOA Ed4.0 and CERES EBAF-Surface Ed4.0 radiation flux data as reference, Level 3C scores (Bias, bc-RMSD) are calculated for each Cloud_cci dataset for the common period (2003 to 2011) and within 60S-60N. The stability of Cloud_cci datasets is investigated by calculating the linear trends and comparing these again to CERES EBAF Ed4.0.

In this section the following radiation datasets are used:

- Cloud_cci AVHRR-AMv3 (see Section 2.2 for details)
- Cloud_cci AVHRR-PMv3 (see Section 2.2 for details)
- CERES EBAF-TOA Ed4.0 (see Section A.6 for details)
- CERES EBAF-Surface Ed4.0 (see Section A.6 for details)
- CLARA-A2 (see Section A.8 for details)
- NOAA HIRS CDR OLR Version 2.7 (see Section A.12 for details)
- ERA-Interim reanalysis (see Section A.13 for details)

The intercomparison is performed for the radiation fluxes (1) top of the atmosphere upwelling solar radiation, (2) top of the atmosphere upwelling thermal radiation, (3) bottom of the atmosphere downwelling solar radiation and (4) bottom of the atmosphere downwelling thermal radiation. Note that not all compared datasets contain all four radiation fluxes. However, Cloud_cci, CERES EBAF Ed4.0 and ERA-Interim include all investigated fluxes and form therefore the basis for the intercomparison. Please note, that for Cloud_cci AVHRR-based radiation fluxes the minimum number of processed orbits must be at least 300 per month. Otherwise the monthly means were blacklisted.

Table 5-1 Time periods of the utilized datasets compared in this section. Note that not all compared datasets contain all for radiation fluxes and the available period of contained radiation fluxes may differ.

Dataset	Cloud_cci AVHRR-AMv3	Cloud_cci AVHRR-PMv3	Cloud_cci ATSR2-AATSRv3	CERES EBAF Edition 4.0	CLARA-A2	NOAA HIRS CDR Version 2.7	ERA-Interim reanalysis
Period	1991-2016	1982-2016	1995-2012	2000-2016	1982-2015	1979-2017	1979-2017

5.1 Top of the Atmosphere Upwelling Solar Radiation

In this subsection the top of the atmosphere (TOA) upwelling solar radiation of Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-PMv3, Cloud_cci ATSR2-AATSRv3, CERES EBAF-TOA Ed4.0 and ERA-Interim are compared by means of multi-annual mean (Figure 5-1) , zonal mean (Figure 5-2) and standard deviation (Figure 5-3), all for a common time period, and time series plots (Figure 5-4).

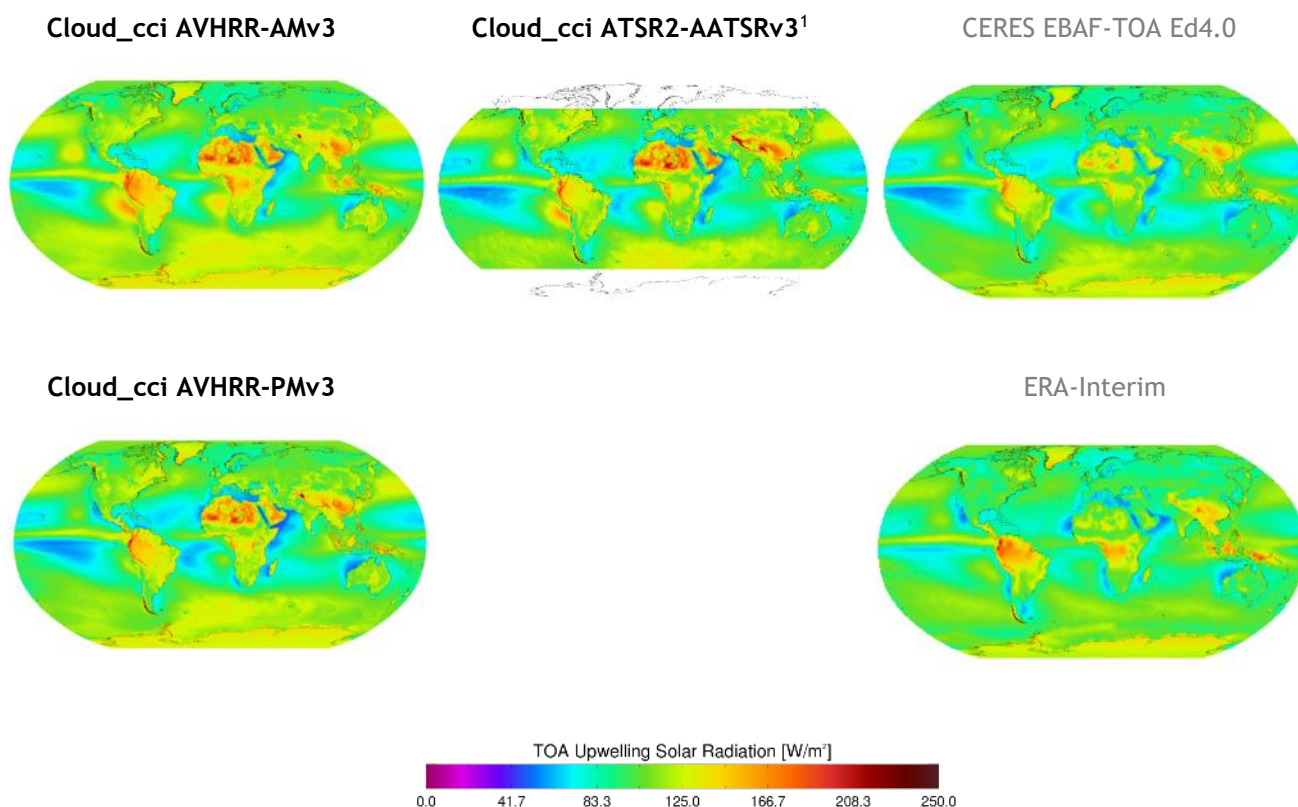


Figure 5-1 Globally gridded means of TOA Upwelling Solar Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

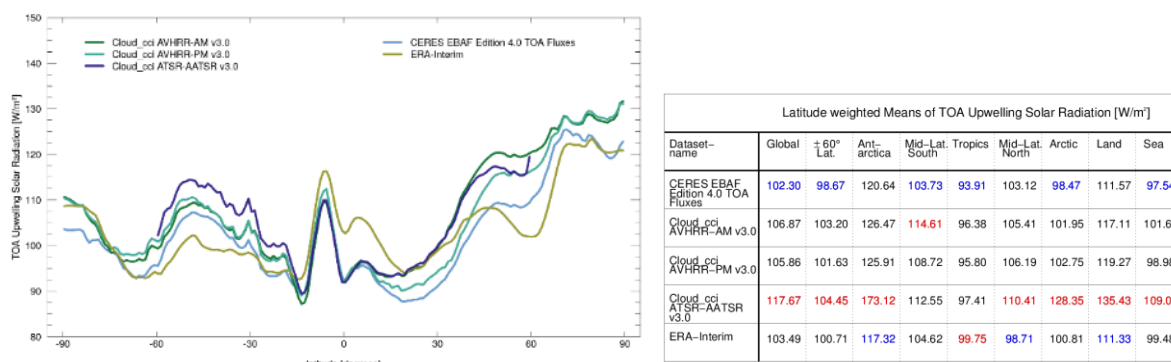



Figure 5-2 TOA Upwelling Solar Radiation from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. Highest and lowest values of each region are highlighted.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 102

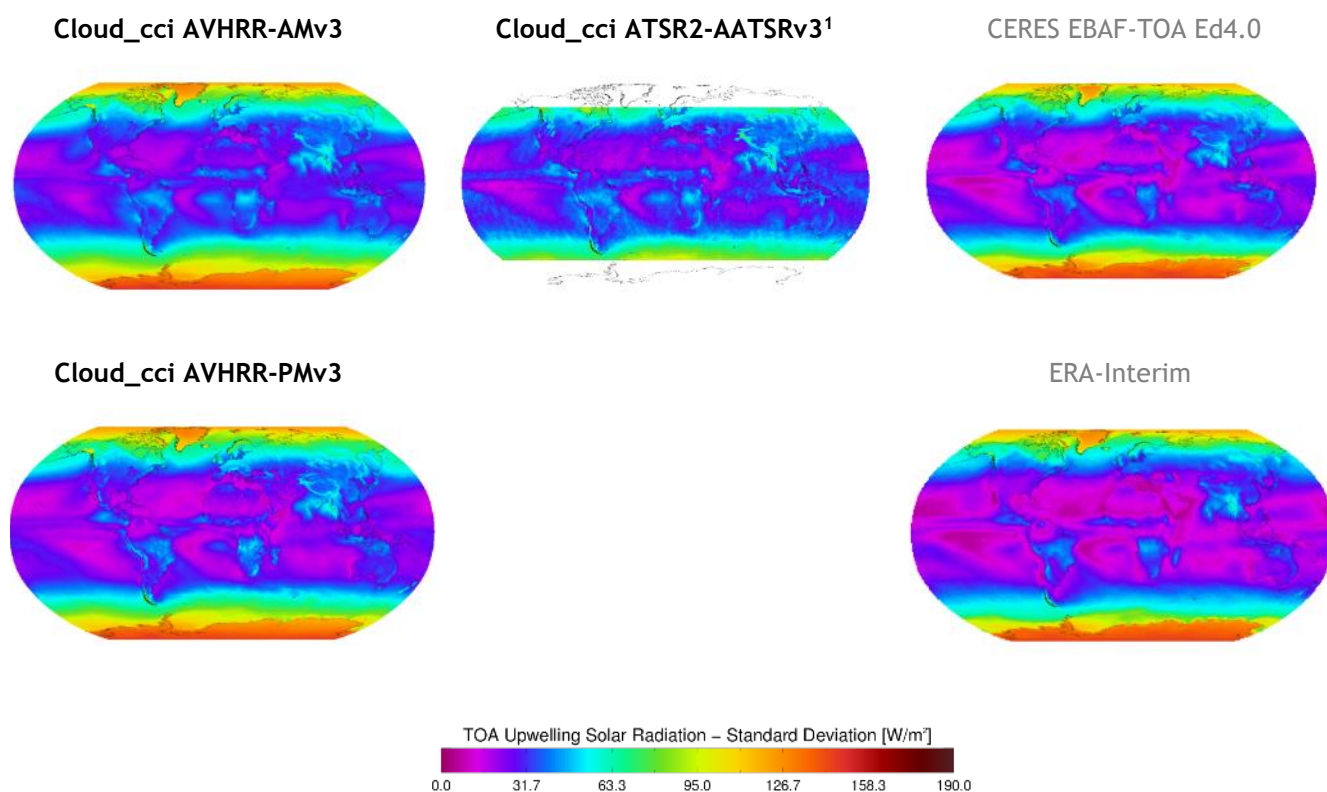


Figure 5-3 Globally gridded standard deviation of TOA Upwelling Solar radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- All datasets show the main characteristic patterns of the global TOA upwelling solar radiation distribution. Highest mean TOA upwelling solar radiation is found in the subtropics over land, lowest mean TOA Upwelling solar radiation is also found in the subtropics over the ocean.
- Generally, all multi-annual averages of the datasets compare well with each other; regionally higher differences are visible. Largest spread among Cloud_cci and reference datasets is found in the subtropics and mid-latitudes.
- Temporal variability is highest for Cloud_cci AVHRR-AMv3 and lowest for ERA-Interim. However, other datasets lie in between, show very similar values and seem relative consistent with showing higher variability over land areas. Largest temporal variability is found for the polar regions, which is reflected in all datasets.
- Time series plots reveal jumps in the time series for Cloud_cci AVHRR-AMv3 and AVHRR-PMv3 at the satellite transitions from 1982 to 2003. After 2003 all datasets are in a good agreement. CERES EBAF Ed4.0 and ERA-Interim are relatively stable, but show small negative trend over their periods.
- The time series plots of TOA upwelling solar radiation show significant seasonal cycles in the global (60S-60N) mean with higher values in boreal winter and lower values in boreal summer for all datasets.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

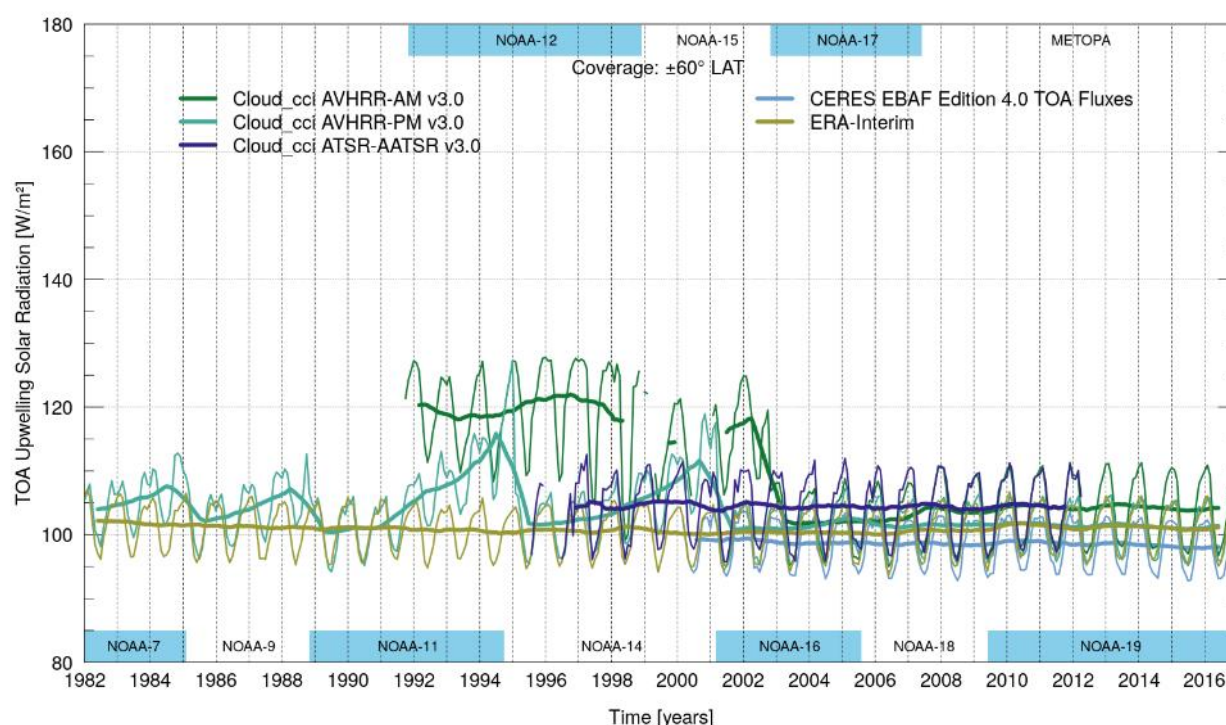


Figure 5-4 Time series of monthly TOA Upwelling Solar Radiation for satellite retrievals and reanalysis (thin lines). Running averages with a time window of 8 months are shown as bold lines.


Discussion

All three Cloud_cci datasets are comparable to the reference datasets for most parts of the globe. An exception here is the Sahara region where Cloud_cci datasets have significantly higher TOA upwelling solar radiation. In general, Cloud_cci ATSR2-AATSR gives the highest TOA upwelling solar radiation for large parts of the globe, which might be due to the different equator crossing times. Further reasons might be algorithm based differences in cloud detection (CERES EBAF Edition 4.0) or fundamental differences in the generation of the dataset e.g. the ERA-Interim reanalysis.

Looking at the time series of AVHRR-AM, problems of early AVHRRs (in particular NOAA-12) become evident. As mentioned earlier NOAA-12 has a difficult twilight orbit. In twilight conditions the derived cloud properties (that are used for calculation of the fluxes), are characterized by higher uncertainty. The AVHRR-PM time series of TOA upwelling solar radiation is affected by the drift of the afternoon satellite, for which we have seen similar features for cloud properties in the previous section. After 2003, all Cloud_cci datasets agree very well with the references. Due to the low frequency sampling of single polar-orbiting AVHRR sensor and the diurnal cycle of sun illumination, a diurnal cycle correction based on the solar zenith angle and the atmospheric path length is applied to the shortwave radiation fluxes to make the data represent 24 hour averages.

For the comparison period from 2003 to 2011 trends for the Cloud_cci datasets are slightly positive. The trend for the morning satellites is stronger than for the afternoon satellites, which is probably due to the pronounced satellite transition from NOAA-17 to METOP-A. CERES EBAF Ed4.0 as well as ERA-Interim show similar behavior as Cloud_cci AVHRR-PM with only very slight trends and high stability. Looking at the whole time series of CERES EBAF Ed4.0 and ERA-Interim small decreasing trends for TOA upwelling shortwave radiation are visible. Overall, after the year 2000 the Cloud_cci and reference datasets are very similar. It should be noted that the gaps in the Cloud_cci AVHRR-AMv3 dataset around the year 2000 is due to blacklisting a large number of AVHRR NOAA-15 observations due to AVHRR scan motor errors. Monthly means are excluded

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 104

from this evaluation when the number of processed orbits was below 300 (normally at about 450) over the entire month.

Table 5-2 presents evaluation scores inferred from comparisons of the Cloud_cci datasets with CERES EBAF TOA Edition 4.0. All Cloud_cci datasets show slightly positive TOA upwelling shortwave radiation biases compared to CERES EBAF Ed4.0 up to 4.5 W/m². The bc-RMSD values are very low with around 1.7 W/m² for Cloud_cci AVHRR-AMv3 and 0.73 for Cloud_cci AVHRR-PMv3. Decadal trend in 2003-2011 is very small positive for Cloud_cci AVHRR-PM and in good agreement to CERES EBAF Ed4.0. For Cloud_cci AVHRR-AM the positive decadal trend is stronger, mainly due to the satellite transition from NOAA-17 to Metop-A in 2007.

Table 5-2: Evaluation scores for Cloud_cci Level-3C AVHRR-AMv3, AVHRR-PMv3 and ATSR2-AATSRv3 based on comparison to CERES EBAF-TOA Ed4.0 from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the CERES and the Cloud_cci dataset. This can introduce some small variability in the trend values for CERES.

Measure	Cloud_cci v3.0 dataset		
	AVHRR-AM	AVHRR-PM	ATSR2-AATSR
Bias [W/m ²]	+4.48	+2.92	+5.72
bc-RMSD [W/m ²]	+1.70	+0.73	+1.64
Cloud_cci v3.0 trend [W/m ² / decade]	+3.44	+0.03	-0.09
CERES EBAF Edition 4.0 trend [W/m ² / decade]	+0.06	+0.06	+0.06
Δ trend [W/m ² / decade]	+3.38	-0.04	-0.15

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

5.2 Top of the Atmosphere Upwelling Thermal Radiation

In this subsection the top of the atmosphere upwelling thermal radiation of Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-PMv3, Cloud_cci ATSR2-AATSRv3, CERES EBAF-TOA Ed4.0, NOAA HIRS OLR Climate Data Record and ERA-Interim are compared by means of multi-annual mean (Figure 5-5), zonal mean (Figure 5-6) and standard deviation (Figure 5-7), all for a common time period, and time series plots (Figure 5-8).

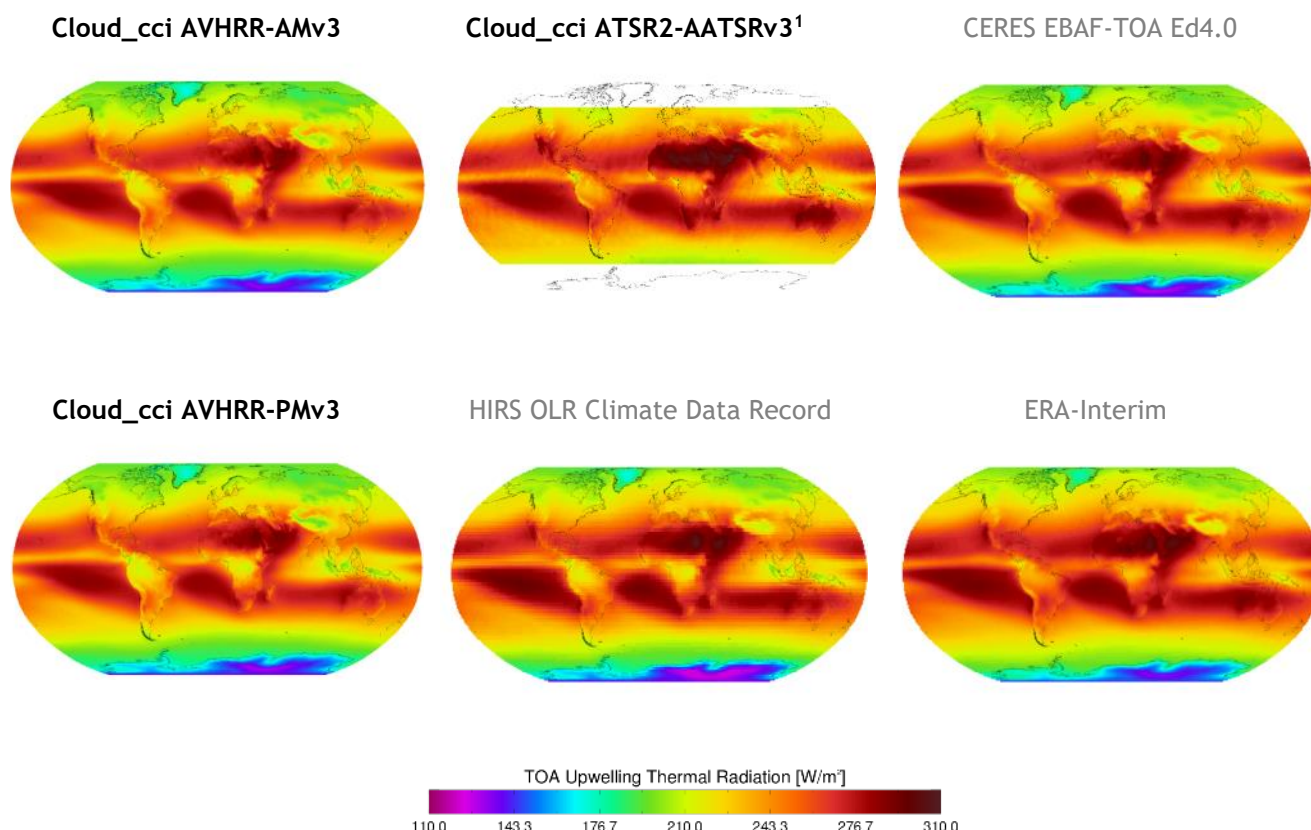


Figure 5-5 Globally gridded means of TOA Upwelling Thermal Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

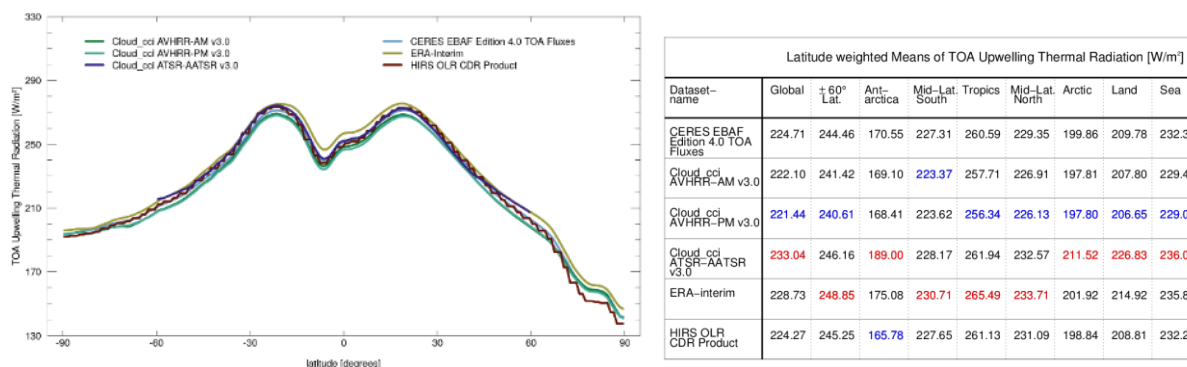



Figure 5-6 TOA Upwelling Thermal Radiation from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. **Highest** and **lowest** values of each region are highlighted.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 106

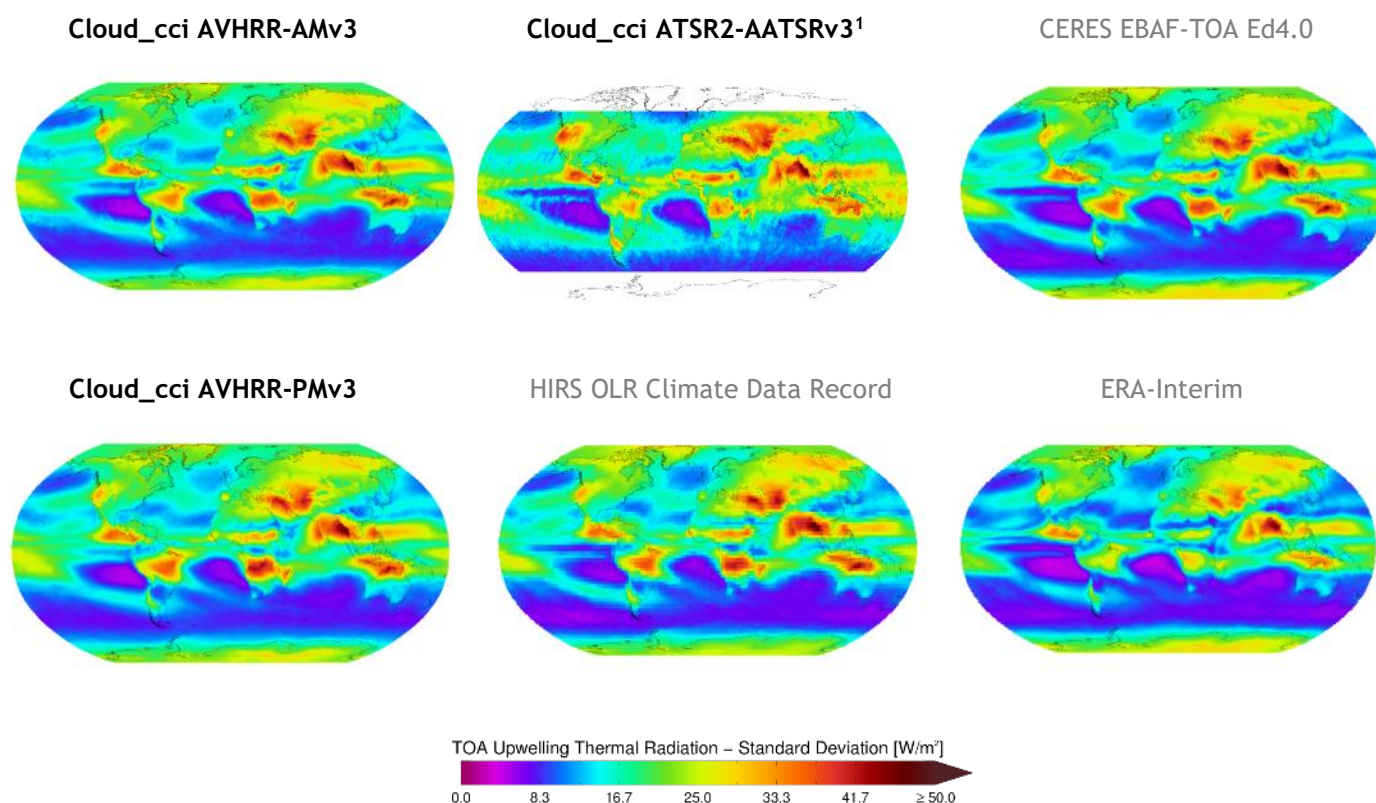



Figure 5-7 Globally gridded standard deviation of TOA Upwelling Thermal radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- Mean global TOA upwelling thermal radiation are lowest over Antarctica and highest over the subtropics. Despite visible differences in the spatial resolution, all datasets show very similar global patterns and comparable mean values.
- Stratocumulus regions are strongly pronounced with the highest TOA upwelling thermal radiation means. In case of the eastern Pacific, a strong gradient between the sea and land surface is noticeable. Stratocumulus regions around Africa and Australia show less differences between land and sea, which is probably due to the different topographic conditions.
- Higher temporal variability is found over land than over the ocean for all datasets. Subtropical land areas show the largest temporal variabilities, e.g. Southeast Asia. The stratocumulus regions and southern hemispheric storm track region show the lowest temporal variability. ERA-Interim dataset has the overall lowest variability, while all other datasets are very similar.
- Global (60S-60N) mean values are highest for ERA-Interim (approx. 248 W/m²) and lowest for Cloud_cci AVHRR-PMv3 with approx. 240 W/m², but all datasets are stable beyond 2003. Satellite transitions and drifts caused jumps and gaps in the time series for Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-PMv3 and NOAA HIRS OLR CDR. Unusual is the opposite trend for the time series of Cloud_cci AVHRR-PM compared with NOAA HIRS OLR CDR in the period from 1998 to 2002.
- In addition, time series plots of TOA upwelling thermal radiation highlight a significant seasonal cycle in the global (60S-60N) mean with maximum values in boreal summer and minimum values in boreal winter for all datasets.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 107				

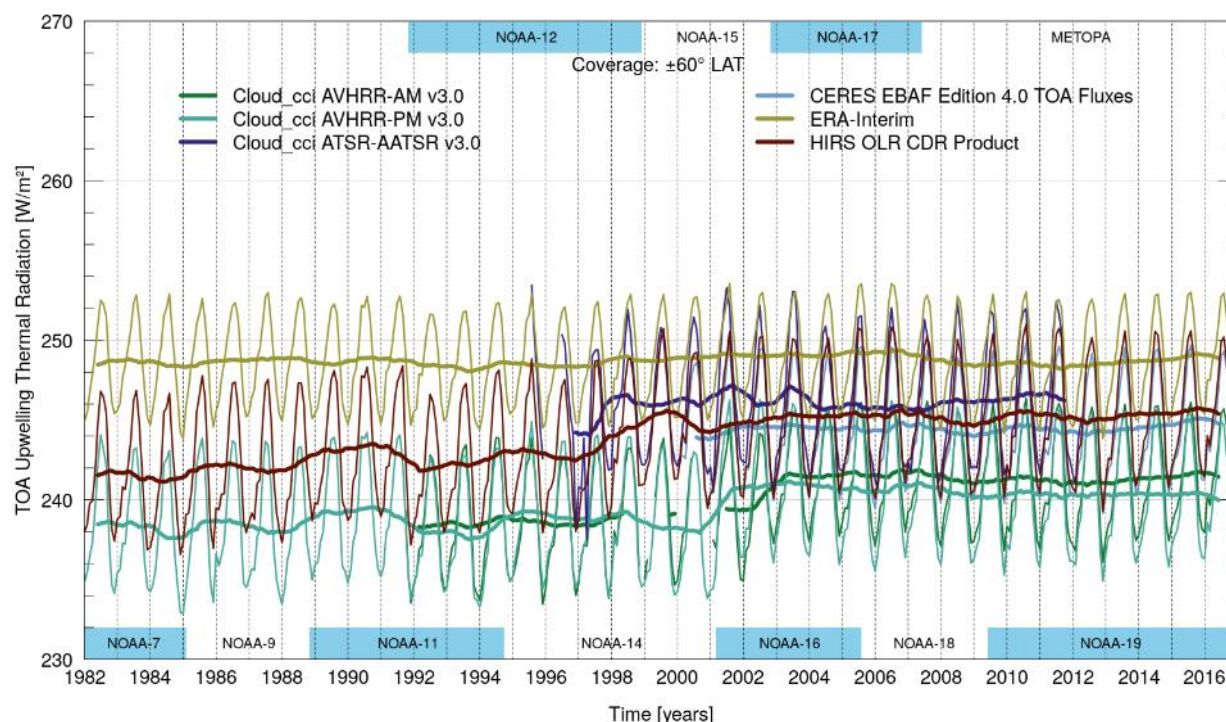



Figure 5-8 Time series of monthly TOA Upwelling Thermal Radiation for satellite retrievals and reanalysis (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

The Cloud_cci AVHRR-AM and PM L3C datasets for TOA upwelling thermal radiation show the characteristic global patterns as seen in the reference datasets. Note that this time the reference datasets CERES EBAF Ed4.0 and ERA-Interim are supplemented by the AVHRR-based HIRS OLR CDR Product, which is only available for the TOA upwelling thermal radiation. All datasets show very similar globally gridded means and even regional characteristics are very comparable. In contrast to the TOA upwelling solar radiation, Cloud_cci AVHRR-AMv3 and AVHRR-PMv3 are almost identical in case of the TOA upwelling thermal radiation, which is probably due to the lower variability of the thermal radiation fluxes. When looking at the zonal means, the ERA-Interim dataset shows the slightly highest values for the globe. In some Figures the relative coarse spatial resolution of the NOAA HIRS product is apparent.

The time series plot reveals a good agreement among all datasets within a range of ± 10 W/m². The Cloud_cci AVHRR-AMv3 and AVHRR-PMv3 datasets are slightly below the reference datasets. The time series show that the differences between morning and afternoon satellite datasets are very small. Again problems of early AVHRRs are visible, but in case of TOA upwelling thermal radiation they are less pronounced. In the Cloud_cci AVHRR-PMv3 dataset decreasing trends are noticeable over the life span of NOAA-7, NOAA-9, NOAA-11 and NOAA-14 satellites, resulting in smaller jumps at the satellite transitions. The strongest jumps are found at the satellite transition from NOAA-14 to NOAA-16 for Cloud_cci AVHRR-PM in 2001 and from NOAA-15 to NOAA-17 for Cloud_cci AVHRR-AM in 2002. After 2003, the Cloud_cci datasets are relatively stable. The strong trends and jumps are the result of steadily drifting orbits of early NOAA AVHRRs, which leads to a strongly reduced data accuracy. Further, jumps, trends and gaps in both AVHRR datasets are the result of twilight or night observations and the blacklisting of observations due to erroneous AVHRR scan motors (NOAA-15). When looking at the reference data, the NOAA HIRS OLR CDR product is very similar to Cloud_cci datasets, which is most likely due to the fact that the HIRS product is also AVHRR-based. Besides a constantly positive bias of about 5 W/m², HIRS shows almost all the trends and jumps found in Cloud_cci. The only exception can be found for the observations of NOAA-14. Here, Cloud_cci shows a decrease of TOA upwelling thermal radiation with the advancing age of NOAA-14 while HIRS has a strong positive trend. Possible reasons for these variations

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 108				

are differences in the calibration coefficients applied on NOAA-14 observations. However, when considering the other radiation fluxes mentioned in this report, the decreasing trends for NOAA-14 are consistent within the Cloud_cci AVHRR-PMv3 dataset.

The two other reference datasets CERES EBAF Edition 4.0 and ERA-Interim are relatively stable over time. CERES EBAF Ed4.0 shows similar but slightly lower values than the NOAA HIRS OLR CDR product, while ERA-Interim shows the constantly highest values of all five datasets for TOA upwelling thermal radiation. Compared to both Cloud_cci AVHRR datasets, ERA-Interim is constantly by about 10 W/m² higher and compared to ATSR2-AATSRv3, HIRS and CERES constantly about 5 W/m² higher. For the comparison period from 2003 to 2011 significant trends or anomalies are barely visible. All datasets are relatively stable during this period and show the same small variations in the TOA upwelling thermal radiation. The main reason for the observed systematic differences are most probably the contrasting observation and generation methods for TOA upwelling thermal radiation.

Table 5-3 presents evaluation scores inferred from comparisons of the Cloud_cci datasets with CERES EBAF Edition 4.0. The Cloud_cci datasets show slightly negative TOA upwelling thermal radiation biases between -3.0 and -4.0 W/m² compared to CERES EBAF Ed4.0. All bc-RMSD values are very low and similar around 0.29 W/m² for Cloud_cci AVHRR-AMv3 and 0.73 W/m² for Cloud_cci AVHRR-PMv3. The decadal trends in 2003-2011 are very small negative for Cloud_cci AVHRR-AM (-0.27 W/m²) and also small negative for Cloud_cci AVHRR-PM (-0.70 W/m²). All trends are in good agreement to CERES EBAF Ed4.0 trend for TOA upwelling thermal radiation, which is also very slightly negative (-0.01 W/m²).

Table 5-3: Evaluation scores for Cloud_cci Level-3C AVHRR-AMv3, AVHRR-PMv3 and ATSR2-AATSRv3 based on comparison to CERES EBAF-TOA Ed4.0 from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the CERES and the Cloud_cci dataset. This can introduce some small variability in the trend values for CERES.

Measure	Cloud_cci v3.0 dataset		
	AVHRR-AM	AVHRR-PM	ATSR2-AATSR
Bias [W/m ²]	-3.02	-3.84	+1.72
bc-RMSD [W/m ²]	+0.29	+0.50	+1.12
Cloud_cci v3.0 trend [W/m ² / decade]	-0.27	-0.70	+0.51
CERES EBAF Edition 4.0 trend [W/m ² / decade]	-0.01	-0.01	-0.01
Δ trend [W/m ² / decade]	-0.25	-0.69	+0.52

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

5.3 Bottom of the Atmosphere Downwelling Solar Radiation

In this subsection the bottom of the atmosphere (BOA) downwelling solar radiation of Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-PMv3, Cloud_cci ATSR2-AATSRv3, CERES EBAF-SURFACE Ed4.0, CLARA-A2 and ERA-Interim are compared by means of multi-annual mean (Figure 5-9), zonal mean (Figure 5-10) and standard deviation (Figure 5-11), all for a common time period, and time series plots (Figure 5-12).

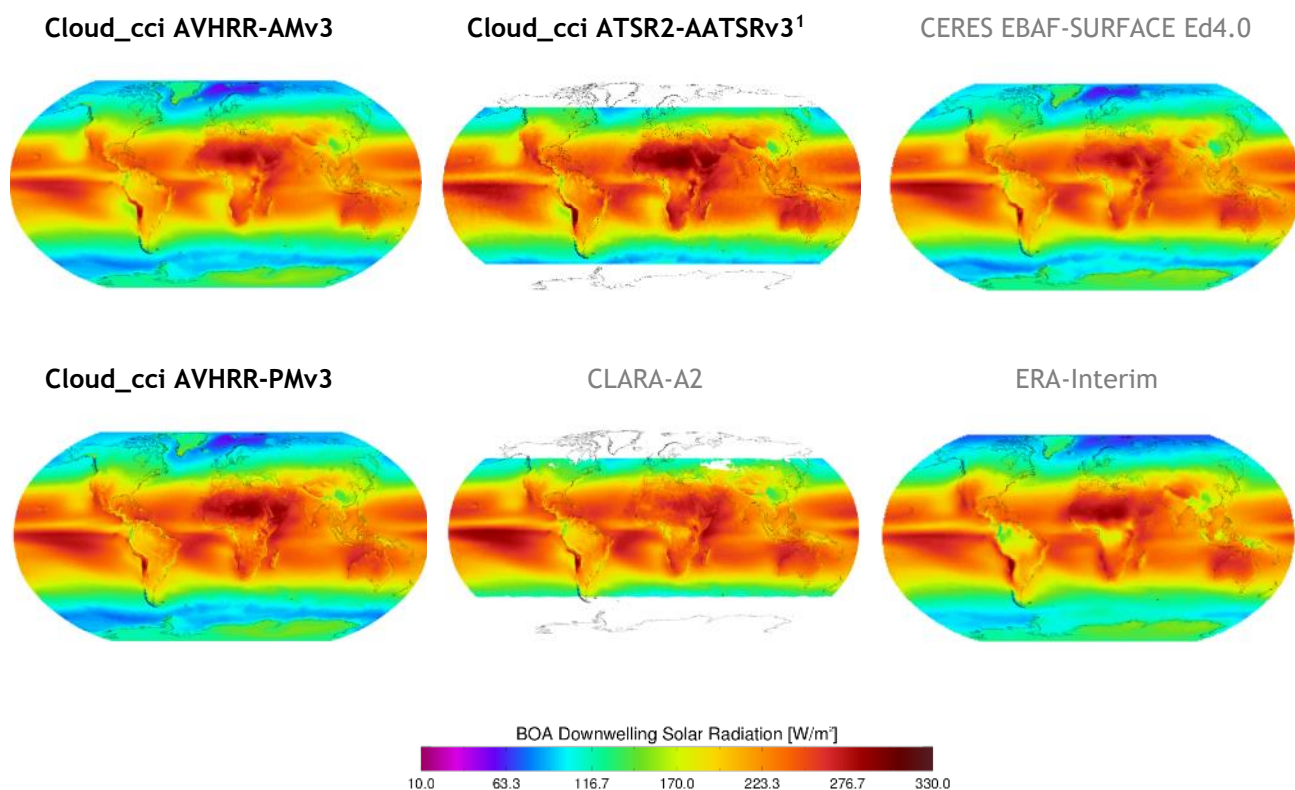


Figure 5-9 Globally gridded means of BOA Downwelling Solar Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

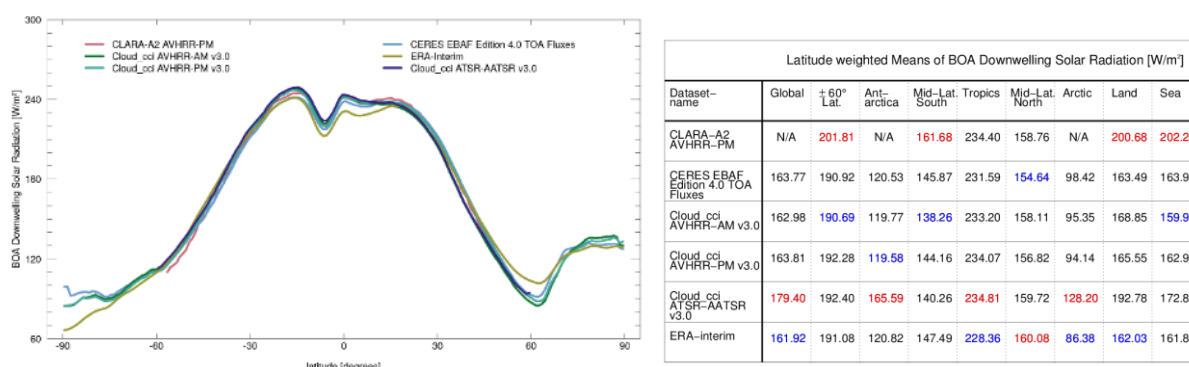



Figure 5-10 BOA Downwelling Solar Radiation from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. **Highest** and **lowest** values of each region are highlighted.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 110

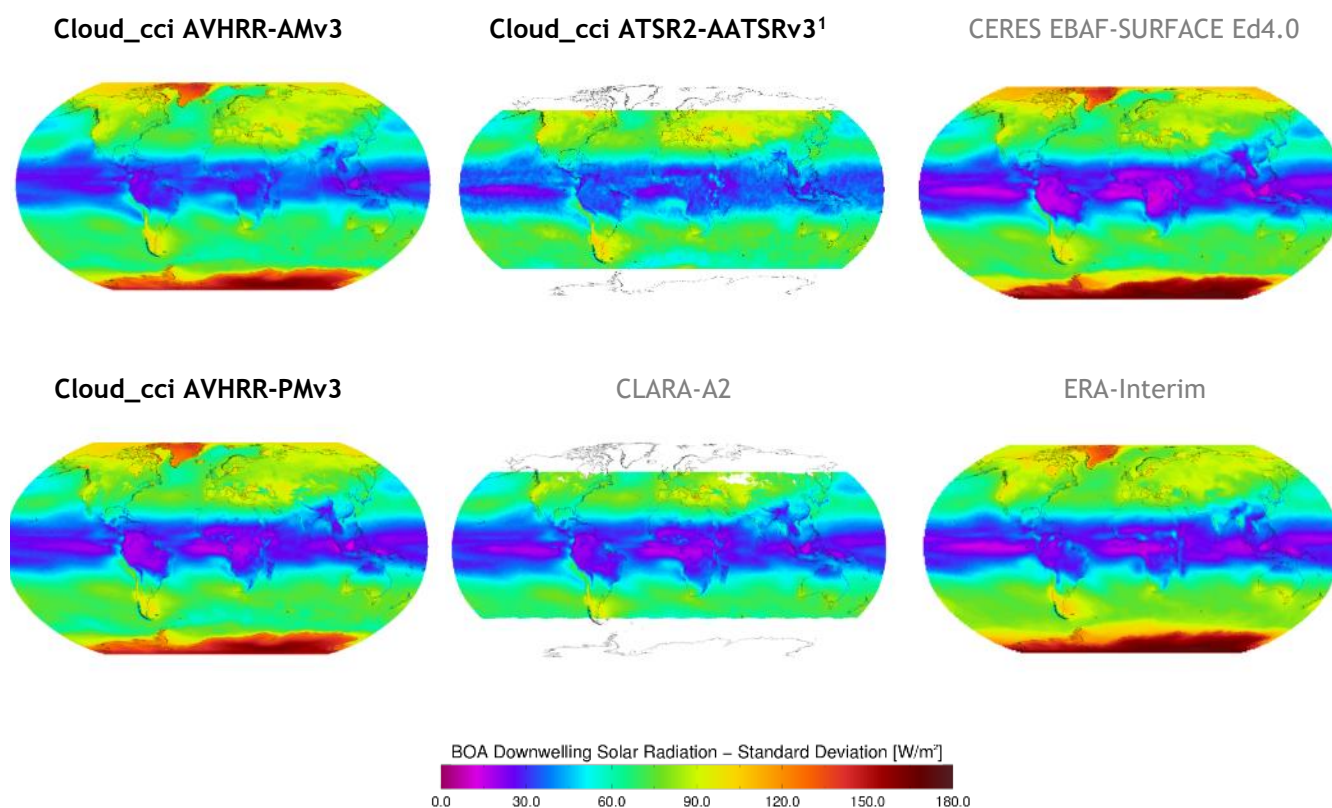



Figure 5-11 Globally gridded standard deviation of BOA Downwelling Solar Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- The datasets show very similar patterns of the global mean bottom of the atmosphere (BOA) downwelling solar radiation. Larger values are found for the subtropics; the maximum is located in the Atacama desert. Lowest mean BOA downwelling solar radiation is found over the polar regions.
- Spread among the datasets is highest for polar regions and high latitudes. Note, that the CLARA-A2 dataset is restricted to an area from 50° south to 60° north. However, for the existing area CLARA-A2 shows the highest mean BOA downwelling solar radiation. Again, also the stratocumulus regions are clearly noticeable in the datasets.
- All datasets represent similar patterns of temporal variability. The tropics show the lowest variability over time, while polar landmasses show the highest. CERES EBAF-SURFACE Ed4.0 contains both the highest and lowest values for temporal variability, but other datasets are alike.
- The time series plots reveal large jumps in the time series for Cloud_cci AVHRR-AMv3 and CLARA-A2 due to satellite drifts as well as satellite transitions before 2003. CLARA-A2 shows an increase for BOA downwelling solar radiation, while in contrast Cloud_cci AVHRR-AMv3 shows a clear decrease for drifting periods. Other datasets are relatively stable in time, with small jumps being visible in the Cloud_cci AVHRR-PMv3 time series before 2003.
- After 2003, CLARA-A2 indicates a positive systematic error about 5 W/m² compared to all other datasets. In addition, it seems that the existing seasonal cycle for BOA downwelling solar radiation is slightly delayed for CLARA-A2, which is due to the limited area of available data.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 111				

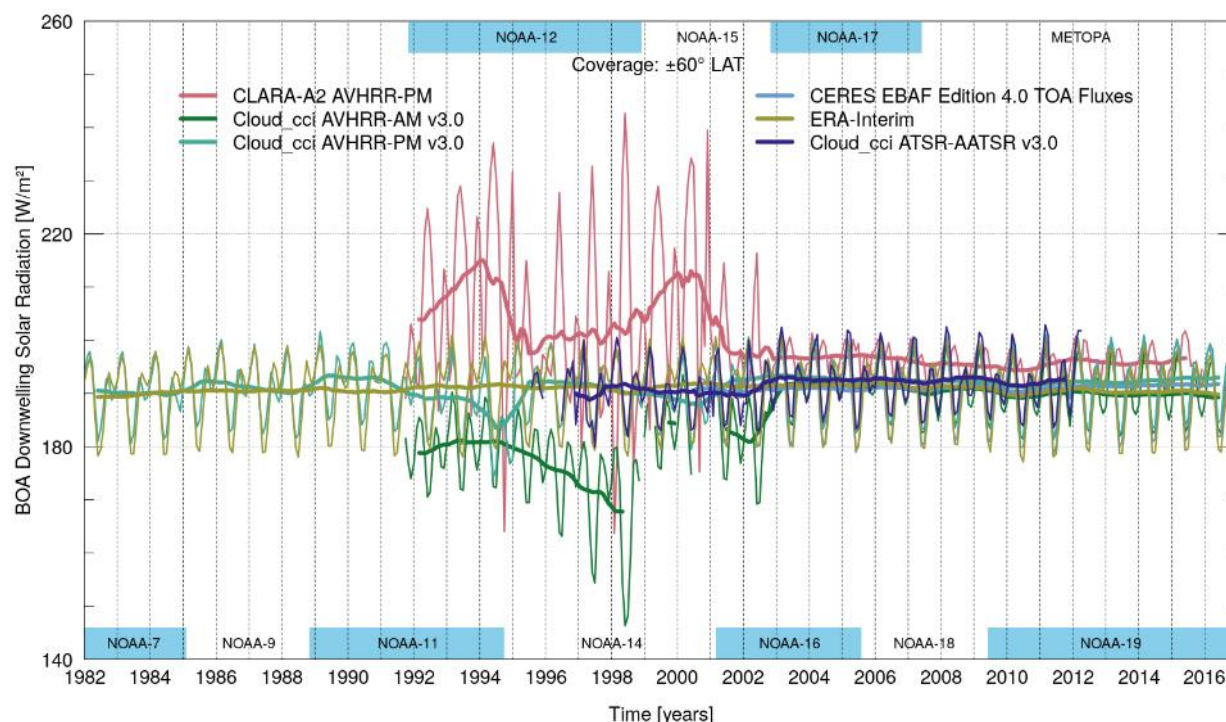



Figure 5-12 Time series of monthly BOA Downwelling Solar Radiation for satellite retrievals and reanalysis (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

The Cloud_cci datasets show very similar patterns of the global mean BOA downwelling solar radiation compared to the reference datasets. In case of the surface radiation fluxes (bottom of the atmosphere - BOA) the AVHRR-based CLARA-A2 dataset is added to the intercomparison. Due to missing values and low numbers of processed orbits, the CLARA-A2 dataset is restricted to the period from 1991 to 2015 and to the area from 50° south to 60° north. The considered datasets have their maximum located in the subtropics. In particular, desert regions like the Atacama or Sahara desert show very high BOA downwelling solar radiation fluxes, which is most likely due to the geographical location as well as the frequent clear sky conditions. On the basis of the globally gridded means or the zonal means, it is not exactly possible to determine the dataset with the lowest or highest values. Overall, the spread between the datasets is very small and only slightly increases for higher latitudes.

The time series of Cloud_cci AVHRR-AM reveals some problems of early AVHRRs, especially at the observations of NOAA-12. As mentioned in Figure 5-5, the orbits of NOAA-12 and other early AVHRRs is steadily drifting, which leads to a shift in the local observation time from morning to dawn with twilight or even night illumination conditions. The result is a significantly reduced accuracy and in case of the Cloud_cci AVHRR-AM BOA downwelling solar radiation, a strong decreasing trend with anomalous low radiation values for the life span of NOAA-12. In addition, jumps at the satellite transitions are noticeable, most striking in 2002 at the transition from NOAA-15 to NOAA-17. In contrast, the time series of Cloud_cci AVHRR-PM is much more stable. Although smaller drifting effects and transition jumps are still visible, these errors are much lower and the time series stays very close to the reference datasets. However, after 2003 all Cloud_cci datasets agree very well with the reference data. Cloud_cci AVHRR-PMv3 is slightly above CERES EBAF Ed4.0 and ERA-Interim, while Cloud_cci AVHRR-AMv3 is slightly below the two references. Due to the low sampling of a single polar-orbiting AVHRR sensor and the diurnal cycle of sun illumination on the earth surface, a diurnal cycle correction is applied to the BOA downwelling solar radiation fluxes. With the correction the running mean and the monthly mean values of Cloud_cci datasets are in a good agreement with reference datasets, especially after 2003. Looking at the CLARA-A2 dataset time series, again problems with early AVHRR

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 112				

satellites are noticeable. The effects are very similar to Cloud_cci AVHRR-AMv3, but instead CLARA-A2 is based on AVHRR afternoon satellites. Contrary to the negative trends seen for the morning satellites, the CLARA-A2 AVHRR-PM time series is affected by strong positive trends, especially at the observations of NOAA-11 and NOAA-14. In addition, significant jumps at the satellite transitions from NOAA-11 to NOAA-14 and from NOAA-14 to NOAA-16 are visible. After 2003, CLARA-A2 is much more stable, which is again similar to other AVHRR-based products. For the comparison period from 2003 to 2011, CLARA-A2 is about 5 W/m² higher than the other datasets, which is possibly due to the spatial restriction from 50° south to 60° north. Performed tests also proved that the spatial restriction is the reason for the slightly shifted yearly cycle seen in the monthly means of CLARA-A2. The other reference datasets CERES EBAF SURFACE Edition 4.0 and ERA-Interim are very stable over time. For the period from 2003 to 2011 no significant trends or anomalies are determinable. Compared datasets are relatively stable during this period and only show small variations in the BOA downwelling solar radiation.

Table 5-4 presents evaluation scores inferred from comparisons of the Cloud_cci datasets with CERES EBAF SURFACE Edition 4.0. The Cloud_cci datasets show slightly negative (Cloud_cci AVHRR-AM: -0.18 W/m²) as well as positive (Cloud_cci AVHRR-PM: 1.40 W/m²) TOA upwelling shortwave radiation biases compared to CERES EBAF Ed4.0. The bc-RMSD values are low with values around 2.85 W/m² for Cloud_cci AVHRR-AMv3 and 1.43 for Cloud_cci AVHRR-PMv3. All Cloud_cci decadal trends in 2003-2011 are negative and range from -3.6 to -2.6 W/m². The CERES EBAF Ed4.0 datasets also shows a slightly negative trend about -0.76 W/m² for BOA downwelling solar radiation.

Table 5-4: Evaluation scores for Cloud_cci Level-3C AVHRR-AM, AVHRR-PMv3 and ATSR2-AATSRv3 based on comparison to CERES EBAF-SURFACE Ed4.0 from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the CERES and the Cloud_cci dataset. This can introduce some small variability in the trend values for CERES.

Measure	Cloud_cci v3.0 dataset		
	AVHRR-AM	AVHRR-PM	ATSR2-AATSR
Bias [W/m ²]	-0.18	+1.40	+1.53
bc-RMSD [W/m ²]	+2.85	+1.43	+3.18
Cloud_cci v3.0 trend [W/m ² / decade]	-3.54	-2.68	-1.73
CERES EBAF Edition 4.0 trend [W/m ² / decade]	-0.76	-0.76	-0.76
Δ trend [W/m ² / decade]	-2.78	-1.93	-0.97

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

5.4 Bottom of the Atmosphere Downwelling Thermal Radiation

In this subsection the bottom of the atmosphere downwelling thermal radiation of Cloud_cci AVHRR-AMv3, Cloud_cci AVHRR-PMv3, Cloud_cci ATSR2-AATSRv3, CERES EBAF-SURFACE Ed4.0, CLARA-A2 and ERA-Interim are compared by means of multi-annual mean (Figure 5-13), zonal mean (Figure 5-14) and standard deviation (Figure 5-15), all for a common time period, and time series plots (Figure 5-16).

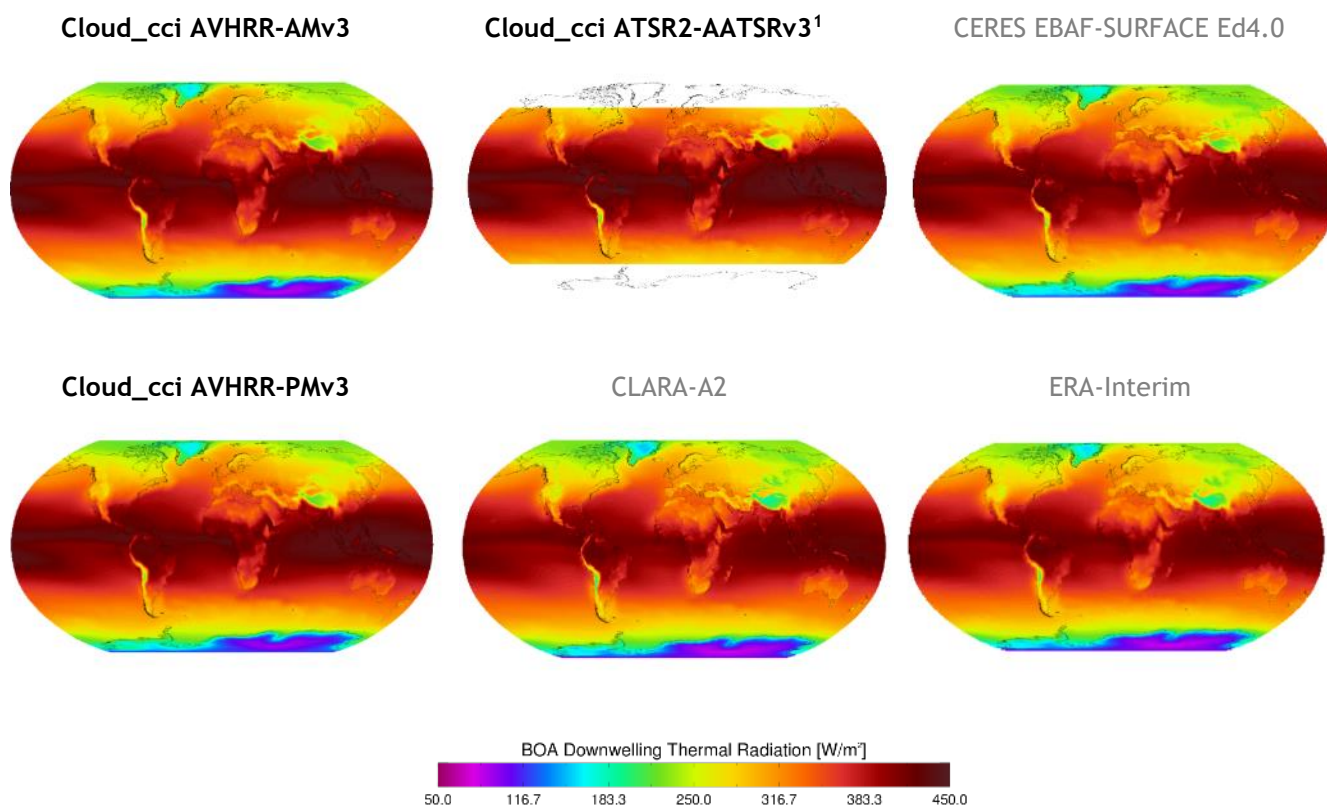


Figure 5-13 Globally gridded means of BOA Downwelling Thermal Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

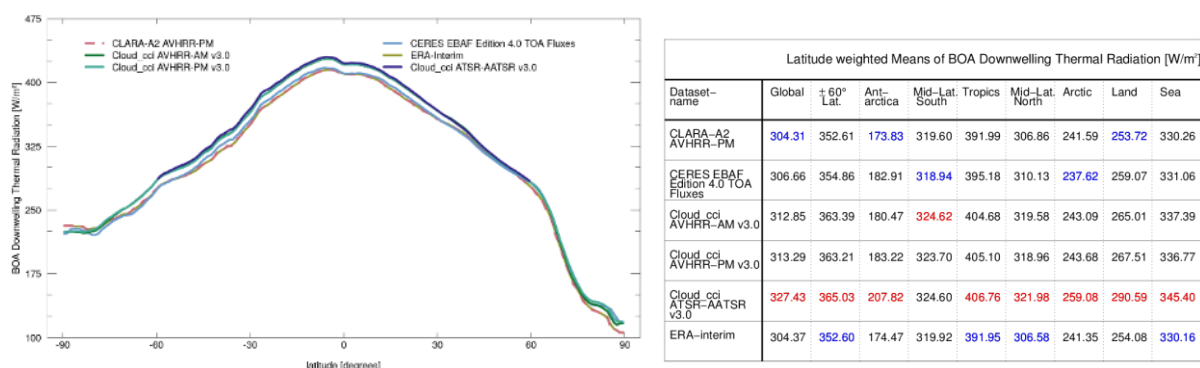



Figure 5-14 BOA Downwelling Thermal Radiation from 2003 to 2011. Left: Zonal means. Right: Latitude weighted means for different regions. Highest and lowest values of each region are highlighted.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 114

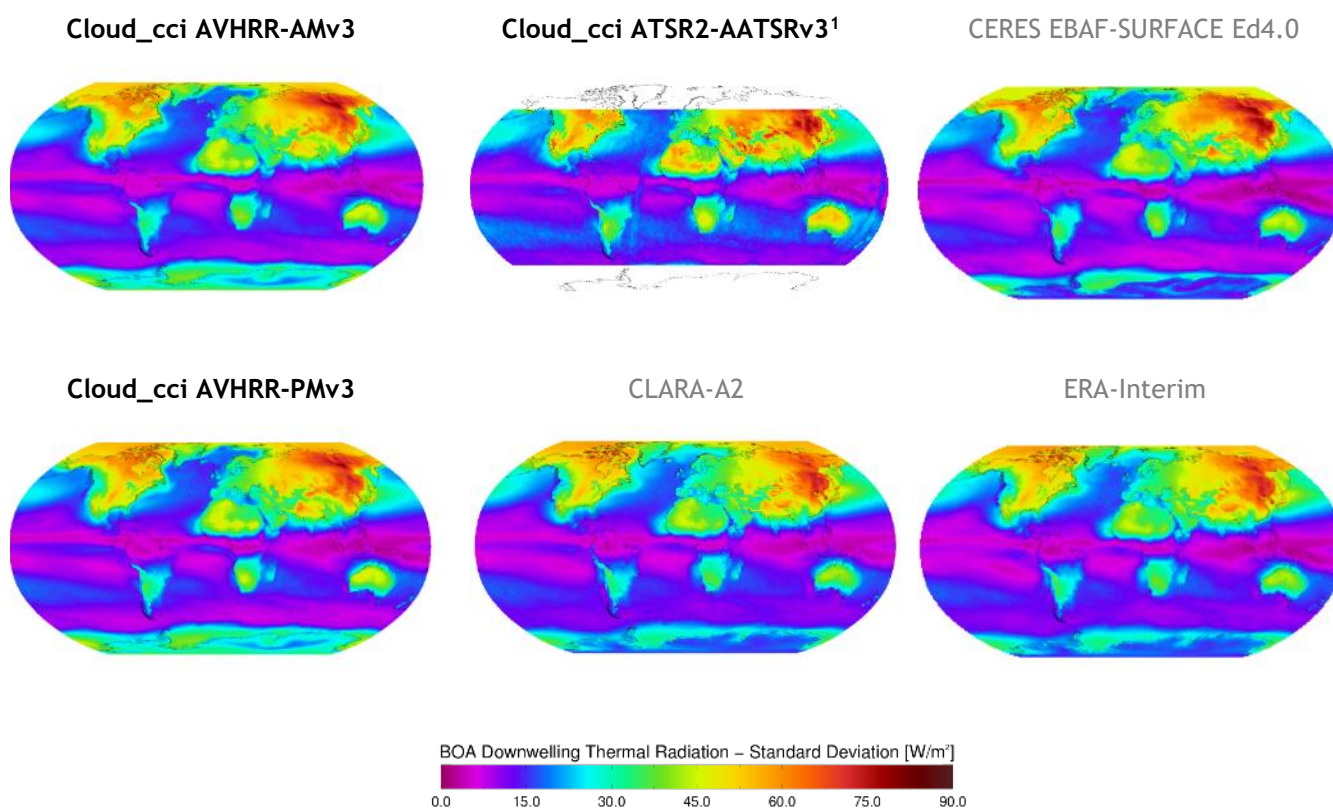


Figure 5-15 Globally gridded standard deviation of BOA Downwelling Thermal Radiation for satellite retrievals and reanalysis averaged over the years 2003 to 2011. Reference data named in grey.

General findings

- Multi-annual global means of all datasets compare very well with each other and hardly any larger differences are visible. The highest BOA downwelling thermal radiation is found over the tropics and subtropics, lowest values are found in Antarctica. Note that CLARA-A2 uses the ERA-Interim BOA downwelling thermal radiation data and is therefore identical to this dataset.
- With the exception of the inner tropics, there are slight differences between sea and land BOA downwelling thermal radiation in all datasets. Higher values are measured over the ocean than over land. Over land, mountains such as the Himalayas or the Andes are noticeable due to their clearly lower values.
- Similar to the global mean, the temporal variability of all datasets also shows a good agreement. In addition, the variability over land is significantly higher than over the ocean, with the exception of a narrow band in the inner tropics. The highest variability is found in East Asia.
- Time series plots reveal differences between the datasets, which are not directly visible in the maps. Cloud_cci AVHRR-AMv3 and Cloud_cci AVHRR-PMv3 have monthly means which are about 10 W/m² higher than the time series of CERES EBAF-Surface Ed4.0 or ERA-Interim/CLARA-A2. Compared to other time series, BOA downwelling thermal radiation contains no clear jumps, but periods of 1-2 years with more or less downwelling thermal radiation.
- Strong seasonal cycles are visible and equally pronounced for all datasets with higher values in boreal summer and lower values in boreal winter. All time series also contain a small positive trend.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

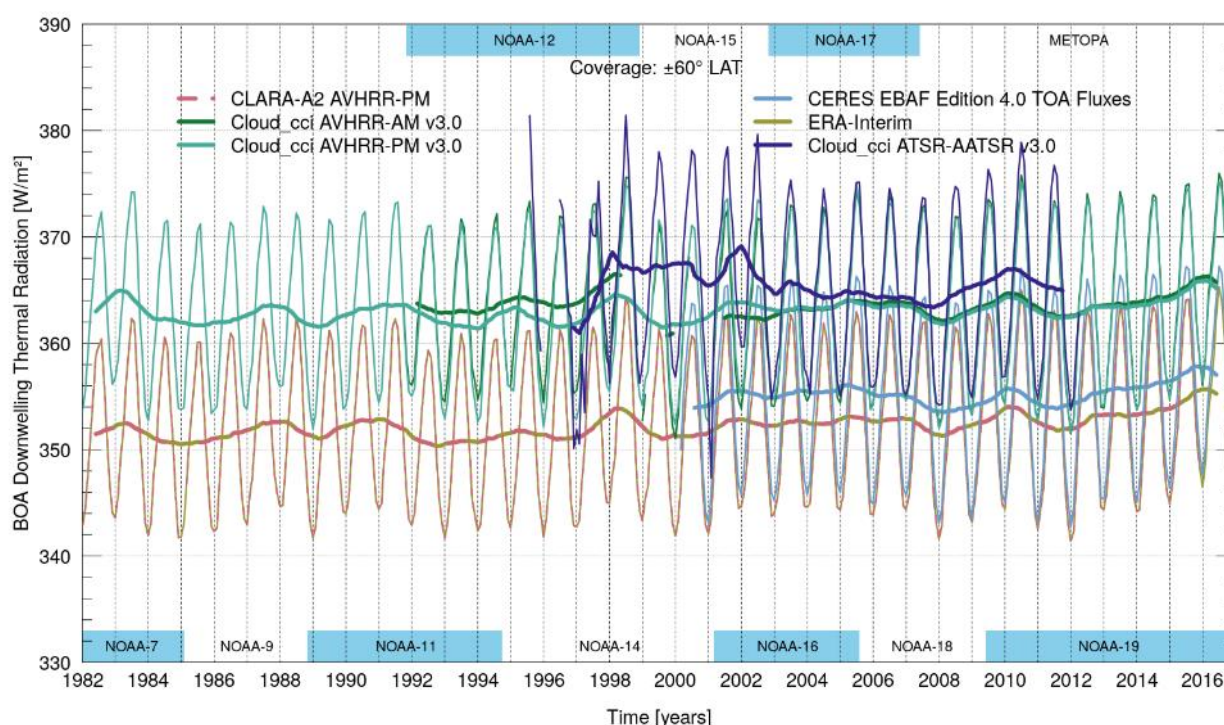



Figure 5-16 Time series of monthly BOA Downwelling Thermal Radiation for satellite retrievals and reanalysis (thin lines). Running averages with a time window of 8 months are shown as bold lines.

Discussion

All Cloud_cci AVHRR L3C datasets for BOA downwelling thermal radiation agree very well with the reference datasets for all parts of the globe. In general, the datasets represent very similar globally gridded means. The highest radiation fluxes are located in the tropics and values decrease with increasing latitude. The zonal mean plot shows that all Cloud_cci AVHRR datasets are slightly higher than the reference datasets. However, the spread among the datasets is very low, even for higher latitudes. Comparable to the TOA upwelling thermal radiation, the globally gridded and zonal means of Cloud_cci AVHRR-AM and AVHRR-PM are almost identical. As mentioned before, this is most likely due to the low variation of the thermal fluxes during the day, or more specific, between the morning and afternoon observation time. In addition, note that in case of the BOA downwelling thermal radiation CLARA-A2 uses the equivalent ERA-Interim fluxes, which means that both datasets are identical in this comparison.

The time series of Cloud_cci AVHRR-AM and Cloud_cci AVHRR-PM show a slightly but systematic positive bias compared to the reference datasets. The differences between Cloud_cci AVHRR-AM and AVHRR-PM are very small and especially after 2003 both time series display almost equal values. Those gaps are the results of satellite transitions as well as the blacklisting of NOAA-15 observations, due to erroneous AVHRR scan motors. Compared to other radiation variables, BOA downwelling thermal radiation shows no problems with the satellite drift of early AVHRRs. Instead, all Cloud_cci time series show several local maxima, peaks as well as short-term trends, for example in 1998 and 2010. The same maxima and minima can be found in all reference datasets, which suggests natural causes by for example large-scaled weather phenomena such as El-Niño Southern Oscillation (ENSO). In addition, the Cloud_cci AVHRR datasets show a significantly positive trend in the recent years from 2011 to 2016. In comparison to Cloud_cci AVHRR, the reference datasets display slightly lower values. In detail, CERES EBAF SURFACE Ed4.0 is constantly about 8 W/m² lower, while ERA-Interim is constantly about 10 W/m² lower. However, the references generally show the same patterns seen in the time series of Cloud_cci AVHRR. For example, CERES shows the significant peak in 2010 and the recent positive trends for BOA downwelling thermal radiation. ERA-Interim also displays these characteristics and also the peaks in the years from 1982 to 2000.

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 116				


When looking at the comparison period from 2003 to 2011 all datasets show very similar patterns for BOA downwelling thermal radiation. Besides the systematically higher values for Cloud_cci AVHRR, the datasets display the same local maxima and minima, with a striking peak in 2010. In the period from 2003 to 2007, small fluctuations are consistently visible for all datasets. Overall, the plotted time series reveal a generally good agreement among all datasets in range of $\pm 10 \text{ W/m}^2$. The many common features and minor differences suggest that the technical characteristics of the datasets are less important in case of the BOA downwelling thermal radiation.

Table 5-5 shows evaluation scores inferred from comparisons of the Cloud_cci datasets with the CERES EBAF SURFACE Edition 4.0 dataset. All Cloud_cci datasets show very similar positive biases between 8 and 10.5 W/m^2 for the BOA downwelling thermal radiation. In addition, the bc-RMSD values are also very similar with values about 0.64 W/m^2 for Cloud_cci AVHRR-AMv3 and 0.67 W/m^2 for Cloud_cci AVHRR-PMv3; ATSR2-AATSRv3 showing slightly larger bc-RMSD. The decadal trends in 2003-2011 are small positive for Cloud_cci AVHRR-AM (0.60 W/m^2) and for Cloud_cci AVHRR-PM (0.18 W/m^2) and a bit larger for ATSR2-AATSR (+1.9 W/m^2). All Cloud_cci AVHRR trends are in contrast to the slightly negative trend (-0.86 W/m^2) of CERES EBAF Ed4.0 for BOA downwelling thermal radiation.

Table 5-5: Evaluation scores for Cloud_cci Level-3C AVHRR-AM & AVHRR-PMv3 based on comparison to CERES EBAF-SURFACE Ed4.0 from 2003-2011. The scores were calculated separately for each Cloud_cci dataset by including all valid data points pairwise in the CERES and the Cloud_cci dataset. This can introduce some small variability in the trend values for CERES.

Measure	Cloud_cci v3.0 dataset		
	AVHRR-AM	AVHRR-PM	ATSR2-AATSR
Bias [W/m^2]	+8.53	+8.35	+10.17
bc-RMSD [W/m^2]	+0.64	+0.67	+1.26
Cloud_cci v3.0 trend [W/m^2 / decade]	+0.60	+0.18	+1.90
CERES EBAF Edition 4.0 trend [W/m^2 / decade]	-0.86	-0.86	-0.86
Δ trend [W/m^2 / decade]	+1.47	+1.04	+2.76

¹The Cloud_cci ATSR2-AATSRv3 data contain Fill_Values in polar regions during polar night. These regions were excluded from this evaluation, despite the fact that good data are available during polar day.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 117

6. Uncertainty validation

In this section the pixel-based random uncertainties associated with the OE-based cloud properties CTP, LWP and IWP are validated. Mandatory for such a validation is a very accurate reference source. For CTP, the CALIOP data is used, for LWP an AMSR-E product and for IWP the DARDAR product. All Cloud_cci data used in this section are from the AVHRR-PM dataset, which is expected to perform similar to all other datasets in terms of pixel-based uncertainty characterizations. For the cloud properties CER and COT no reliable reference data source is available which would allow a similar uncertainty validation as presented for CTP, LWP and IWP. However, as LWP and IWP (and their uncertainties) are derived from CER and COT retrievals (and their uncertainties) the uncertainty validation for LWP and IWP are considered to implicitly contain a CER and COT uncertainty validation. For cloud mask CMA the best reference data (CALIOP) has been used to statistically determine the CMA uncertainty, thus an additional CMA uncertainty validation is neither possible nor needed.

The reported pixel-based uncertainties X_{unc} for a given variable X (both being a result of the optimal estimation technique) represent the 68% confidence interval that the true value is within $X \pm X_{unc}$. Given this, it can be assumed that for 68% of all pixels the truth is within $X \pm X_{unc}$. We will use this rationale in the following to analyse a large set of collocated Cloud_cci and reference observations for the validation of the uncertainties. As the reference data are also not perfect, the reference data uncertainty is also considered in this exercise. In Section 6.1 the CTP is considered (separately for liquid and ice clouds: Section 6.1.1 and 6.1.2, respectively) and Sections 6.2 and 6.3 give the results for LWP and IWP, respectively.

It needs to be noted that the presented framework considers random uncertainties only. For this reason, occurring systematic biases between Cloud_cci and the reference data are removed if significant.

In all equations of this Section i is used as pixel index, d for the pixel-based difference between Cloud_cci and the reference, u for the pixel-based total uncertainty (consisting of Cloud_cci uncertainty and reference data uncertainty) and β as pixel-based ratio between the absolute value of d and u .

6.1 Cloud top pressure


As mentioned above CALIOP data is used as best estimate of the truth for CTP. Unfortunately, no uncertainty measure exists for the CALIOP CTP. For this reason we set all CALIOP CTP uncertainties to 10hPa arbitrarily. In addition, to account for the lack of sensitivity of AVHRR-type sensors to semi-transparent cloud layers, the uppermost cloud layers (until an top-down optical thickness of 1) were removed from CALIOP profiles in this evaluation.

Equations 6-1 to 6-3 list the calculations of d , u and β as defined above. The CTP uncertainty (CTP_{unc}) validation is separated into liquid and ice clouds, for which both phase retrievals (Cloud_cci and CALIOP) had to agree.

$$d_i = CTP_i^{cci} - CTP_i^{caliop} \quad \text{Equation 6-1}$$

$$u_i = \sqrt{(CTP_{unc_i}^{cci})^2 + (CTP_{unc_i}^{caliop})^2} \quad \text{Equation 6-2}$$

$$\beta_i = \frac{|d_i|}{u_i} \quad \text{Equation 6-3}$$

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 118

6.1.1 Liquid clouds

Figure 6-1 shows boxplots for CTP_{unc}^{cci} and CTP_{unc}^{caliop} as function of CTP. The majority of the CTP_{unc}^{cci} does now show a dependence on the CTP^{cci} retrieval itself and lies between 10 and 50 hPa. Exceptions are the low-level clouds below 850hPa for which the median of CTP_{unc}^{cci} becomes smaller, along with an increase in spread. Few outliers exist in each bin with very small or very high uncertainties. As mentioned above no CTP_{unc}^{caliop} information are provided along with CTP^{caliop} , which were in turn set to 10 hPa in this exercise (see right hand side of Figure 6-1).

A small systematic bias is found between CTP^{cci} and CTP^{caliop} for liquid clouds. After analysing this, 12 hPa is subtracted from CTP^{cci} data to centre the frequency distribution of d around 0 (not shown). From here onwards, we consider all remaining deviations to be of random nature only.

Figure 6-2 shows a 2-dimensional histogram over all CTP^{cci} , CTP^{caliop} pairs for all liquid cloud pixels used. In addition a histogram over β is given. Ideally, the portion of all cases that have a β value of 1 or below (β_1 -portion hereafter), should amount to 68%. Thus 68% should fall into the first bin. However, the histogram only reports 47% with $\beta \leq 1$. In the second bin ($1 < \beta \leq 2$) about 21% are found. The numbers decrease further with increasing bin. This investigation suggests too small values for CTP_{unc}^{cci} given that the assumed CTP_{unc}^{caliop} is not unrealistically small. In the following it is investigated if the appropriateness of CTP_{unc}^{cci} is a function of CTP itself. For this, the portion of cases with $\beta \leq 1$ with respect to all cases within certain CTP bins is investigated (Figure 6-3 to Figure 6-5).

Two features are prominent in Figure 6-3: (1) the increase of the β_1 -portion for high clouds reaching values near 70% (although there only few liquid clouds at these altitudes), and (2) the decrease for low clouds, reaching values below 40% for the lowest clouds (high CTPs). The second feature is significant in a sense that it represents a large set of pixels (see right hand side of Figure 6-3). Looking at the results stratified for CTP^{caliop} bins (Figure 6-4), the β_1 -portion is between 35% and 55% for almost all clouds. Exception here is that for the (very few) high level liquid clouds, the β_1 -portion goes down. Figure 6-5 only reveals that the β_1 -portion is large enough when the differences between CTP^{cci} and CTP^{caliop} are small anyway.

Table 6-1 reports correlation coefficients of d with (a) CTP_{unc}^{cci} , (b) with the combined uncertainty u and (c) with β . The correlation coefficients for (a) and (b) are very close to zero. Correlation (c) is 0.76.

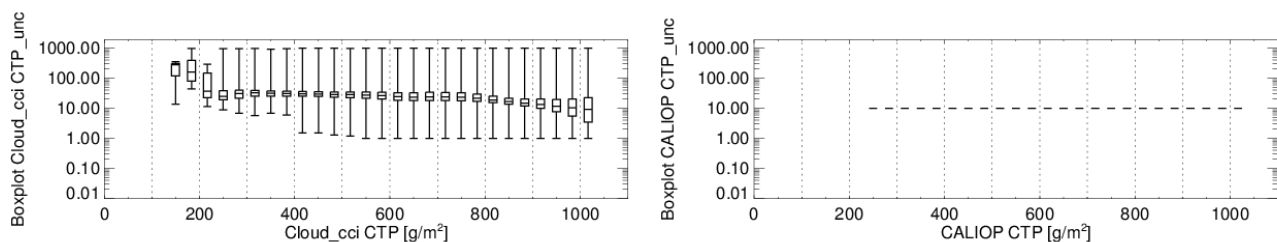



Figure 6-1 Boxplot of CTP uncertainties (CTP_{unc}) shown as function of CTP for Cloud_cci (left) and CALIOP (right) for liquid clouds. Boxes reflect minimum and maximum as well as 25%, 50% and 75% percentiles of CTP_{unc} distribution in each CTP bin. As no CTP uncertainty is available to CALIOP, all values have been set to 10hPa. CTP uncertainties below 1hPa or above 1000hPa were omitted.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 119

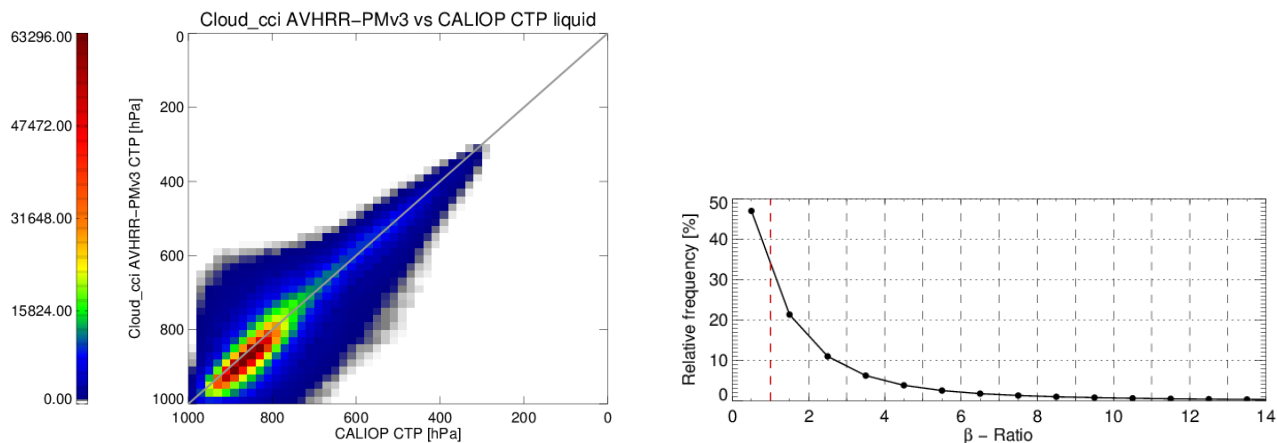


Figure 6-2 Left: 2-dimensional frequency histogram of colocated *Cloud_cci* and *CALIOP* CTP pairs for *liquid* clouds before applying bias correction. Right: Histogram of β -ratio (as defined in Equation 6-3) over all collocations. Vertical dashed lines reflect bin borders (bin width is 1).

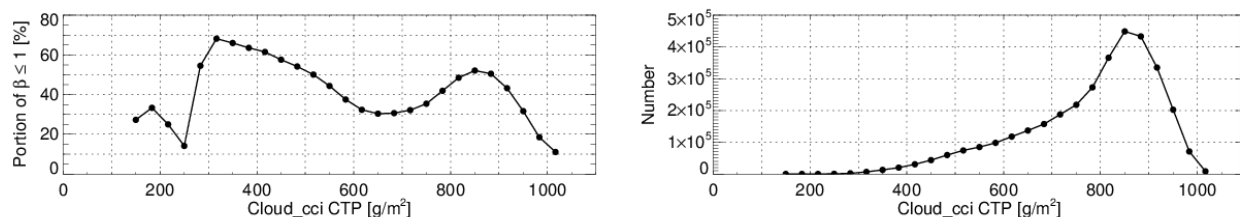


Figure 6-3 Left: Portion of all cases for which $\beta \leq 1$ (β_1 -portion) shown as function of *Cloud_cci* CTP for *liquid* clouds. Right: Corresponding number of pixels per *Cloud_cci* CTP bin.

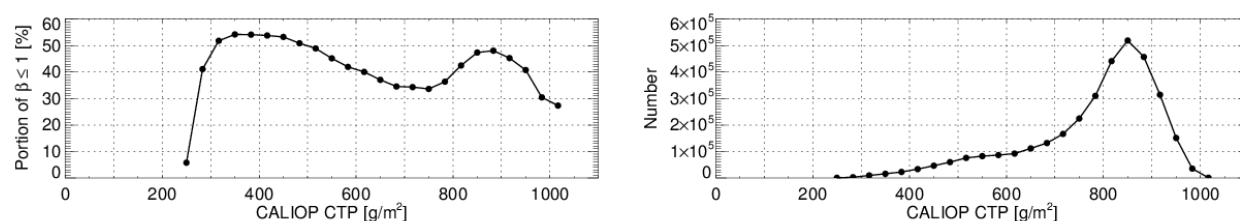


Figure 6-4 As Figure 6-3 but as function of *CALIOP* CTP.

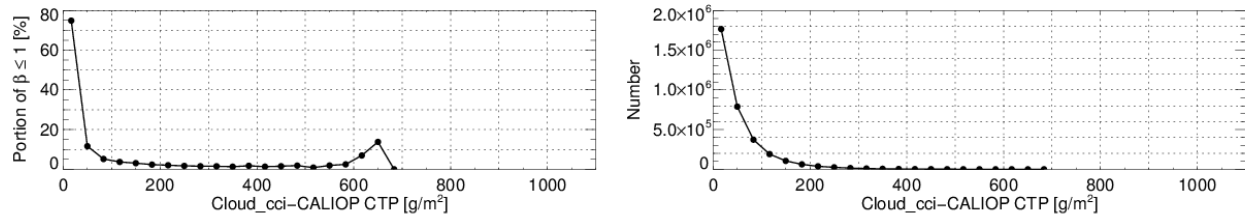


Figure 6-5 As Figure 6-3 but as function of the absolute difference *Cloud-cci CTP - CALIOP CTP*.

Table 6-1 Correlation coefficients for liquid clouds. See text for details.

(a)	Correlation ($ d $, CTP_{unc}^{cci})	0.06
(b)	Correlation ($ d $, u)	0.06
(c)	Correlation ($ d $, β)	0.765

6.1.2 Ice clouds

Figure 6-6 shows boxplots for CTP_{unc}^{cci} and CTP_{unc}^{caliop} as function of CTP. The majority of CTP_{unc}^{cci} does now show a strong dependence on the CTP^{cci} retrieval itself and lies between 10 and 50 hPa. Exceptions are firstly the low-level ice clouds below 750hPa for which the spread becomes larger. Secondly, for high level ice clouds with CTPs below 200 hPa the CTP_{unc}^{cci} becomes larger. Few outliers exist with very small or very high uncertainties. As mentioned above for liquid clouds, no CTP_{unc}^{caliop} are available, which were also set to 10hpa in this exercise for ice clouds.

A moderate systematic bias is found between CTP^{cci} and CTP^{caliop} for ice clouds. After analysing the frequency distribution of d , 60 hPa is subtracted from CTP^{cci} data to centre the distribution around 0 (not shown). From here onwards we consider all remaining deviations to be of random nature only.

Figure 6-7 shows a 2-dimensional histogram over all CTP^{cci} , CTP^{caliop} pairs for all ice cloud pixels used. In addition a histogram over β is given. The histogram only reports 43% with $\beta \leq 1$. In the second bin ($1 < \beta \leq 2$) about 18% are found. These values are a bit lower than for liquid clouds. This investigation suggests too small values for CTP_{unc}^{cci} also for ice clouds given that the assumed CTP_{unc}^{caliop} is not unrealistically small. In the following it is investigated if the appropriateness of CTP_{unc}^{cci} is a function of CTP^{cci} itself. For this, the portion of cases with $\beta \leq 1$ with respect to all cases within certain CTP bins is investigated (Figure 6-8 to Figure 6-10).

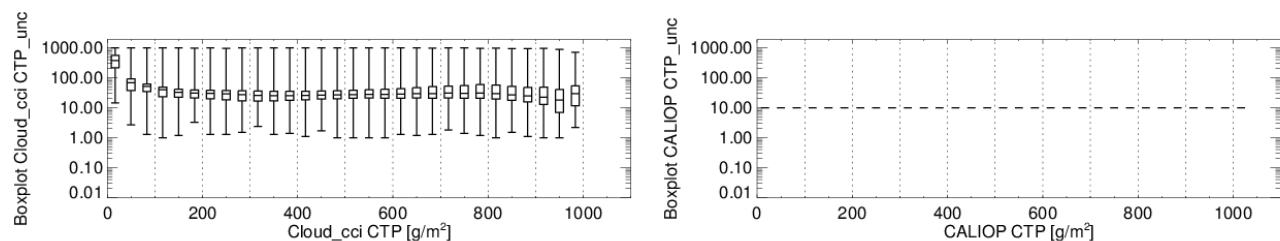



Figure 6-6 Boxplot of CTP uncertainties (CTP_{unc}) shown as function of CTP for *Cloud_cci* (left) and *CALIOP* (right) for ice clouds. Boxes reflect minimum and maximum as well as 25%, 50% and 75% percentiles of

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 121				

CTP_unc distribution in each CTP bin. As no CTP uncertainty is available for CALIOP, all values have been set to 10hPa.

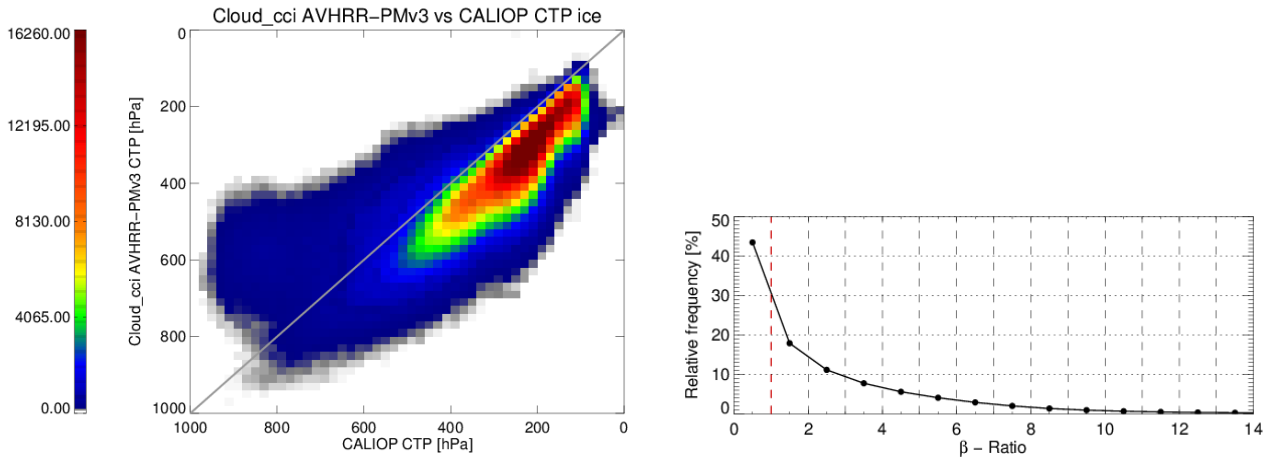


Figure 6-7 Left: 2-dimensional frequency histogram of colocated *Cloud_cci* and CALIOP CTP pairs for ice clouds. Right: Histogram of β -ratio (as defined in Equation 6-3) over all collocations. Vertical dashed lines reflect bin borders (bin width is 1).

In Figure 6-8, except from the very high ice clouds, for which the β_1 -portion is 50% or above, for most parts of the retrieved CTP spectrum the *CTP_unc* are clearly too low.

Looking at the results stratified for CALIOP CTP bins (Figure 6-9), the β_1 -portion is between 20% and 40% for almost all clouds, with higher values for high clouds and smaller values for low clouds. As for liquid cloud the β_1 -portion is large enough when the differences between *Cloud_cci* and CALIOP small (Figure 6-10).

Table 6-2 reports correlation coefficients of d with (a) *CTP_unc^{cci}*, (b) the combined uncertainty u and (c) β . As for liquid clouds, also for ice clouds the correlation coefficients for (a) and (b) are very close to zero. Correlation (c) is 0.72.

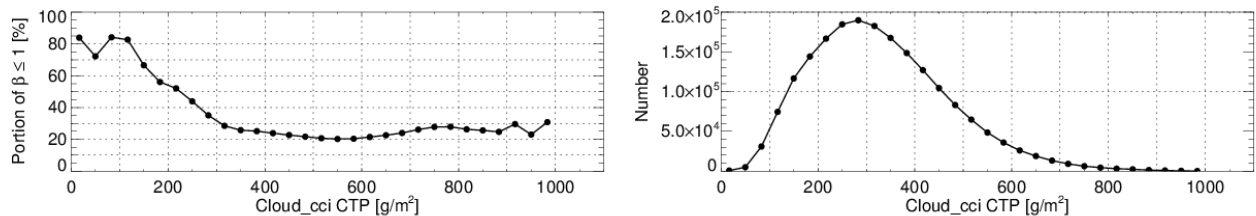


Figure 6-8 Left: Portion of all cases for which $\beta \leq 1$ (β_1 -portion) shown as function of *Cloud_cci* CTP for ice clouds. Right: Corresponding number of pixels per *Cloud_cci* CTP bin.

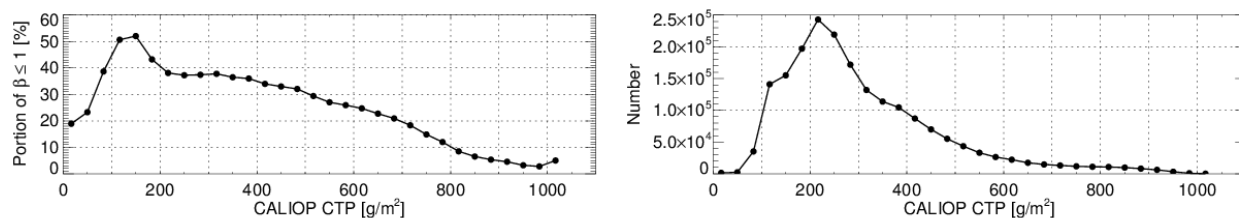


Figure 6-9 As Figure 6-8 but as function of CALIOP CTP.

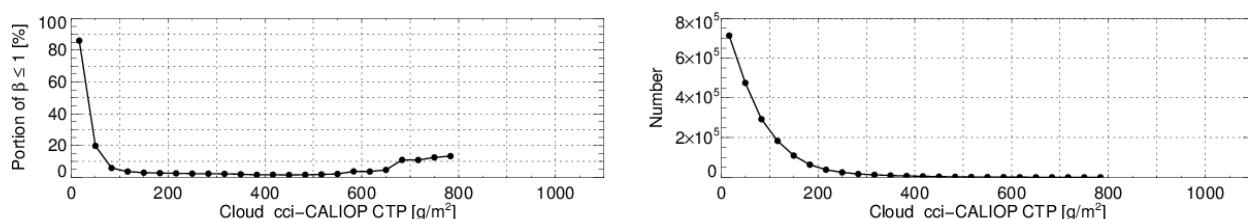


Figure 6-10 As Figure 6-8 but as function of the absolute difference Cloud-cci CTP - CALIOP CTP.

Table 6-2 Correlation coefficients for ice clouds. See text for details.

(a)	Correlation ($ d $, CTP_{unc}^{cci})	-0.1
(b)	Correlation ($ d $, u)	0.12
(c)	Correlation ($ d $, β)	0.72


6.2 Liquid water path

As mentioned in the beginning of Section 6, AMSR-E LWP data (Wentz and Meissner, 2004) is used as best estimate of the LWP truth. There is no pixel-based uncertainty reported for AMSR-E LWP. Instead we use 0.017mm, which is reported as the overall root mean square error in Table 9 of Wentz and Meissner (2002), as pixel-based uncertainty measure. In the discussion below we will tackle what the impact of this compared to no uncertainty or compared to an assumed relative uncertainty is.

Only a very small systematic bias is found between LWP^{cci} and LWP^{amsre} for liquid clouds. As the frequency distribution of d was already centred around 0 (not shown), no further bias correction was applied and all remaining deviations to be of random nature only.

Equations 6-4 to 6-6 list the calculations of d , u and β as defined above. For LWP uncertainty (LWP_{unc}) validation all cases for which either Cloud_cci or AMSR-E reported an ice cloud top were excluded.

$$d_i = LWP_i^{cci} - LWP_i^{amsre} \quad \text{Equation 6-4}$$

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
		Page 123		

$$u_i = \sqrt{(LWP_unc_i^{cci})^2 + (LWP_unc_i^{amsre})^2} \quad \text{Equation 6-5}$$

$$\beta_i = \frac{|d_i|}{u_i} \quad \text{Equation 6-6}$$

Figure 6-11 shows boxplots for LWP_unc^{cci} and LWP_unc^{amsre} as function of LWP. For LWP larger than 30g/m^2 the LWP_unc^{cci} grows exponentially with LWP^{cci} (near-linear relationship in logarithmic plot). As we have set a fixed uncertainty (see above) for LWP_unc^{amsre} , no sensitivity to the AMSR-E LWP exists (right hand side of Figure 6-11).

Figure 6-12 shows a 2-dimensional histogram over all LWP^{cci} , LWP^{amsre} pairs for all liquid cloud pixels available in the collocation set. In addition a histogram over β is given. As mentioned above, the β_1 -portion should ideally amount to 68%. Thus 68% should fall into the first bin. And indeed, the histogram reports about 68% with $\beta \leq 1$. In the second bin ($1 < \beta \leq 2$) about 22% are found. The number decrease further with increasing bin. This investigation suggests that appropriate values are provided by LWP_unc^{cci} given that the LWP_unc^{amsre} correctly presents the 68% confidence level itself. In the following it is investigated if the appropriateness of LWP_unc^{cci} is a function of LWP^{cci} itself. For this, the portion of cases with $\beta \leq 1$ with respect to all cases within certain LWP bins (Figure 6-13 to Figure 6-15).

Figure 6-13 shows that for larger LWP^{cci} the LWP_unc^{cci} seems to have the right amplitude as the β_1 -portion increases to values even above 80%. For LWP^{cci} values below 10g/m^2 (which are only a minority of all cases) the LWP_unc^{cci} seems to be too small. These results however become a bit inconclusive because Figure 6-14 indicates that for larger LWP^{amsre} (above 50g/m^2) the LWP_unc^{cci} seems too low, while also here the appropriateness of LWP_unc^{cci} increase from low to mid-range LWP. Figure 6-15 reveals that the β_1 -portion is large enough when the differences between Cloud_cci and AMSR-E are small (below 20g/m^2), but too low for larger LWP. With respect to the latter, the static LWP_unc^{amsre} might contribute to this finding in case of being too low for larger LWP.

Table 6-3 reports correlation coefficients of the LWP differences (Cloud_cci minus AMSR-E) with (a) the LWP_unc^{cci} , (b) with the combined uncertainty (Cloud_cci and AMSR-E) and (c) with β . The correlation coefficients for (a) and (b) are around 0.28. Correlation (c) is higher (0.84).

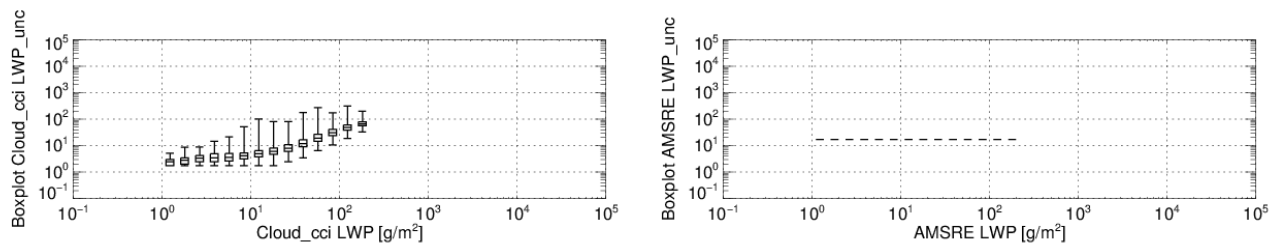


Figure 6-11 Boxplot of LWP uncertainties (LWP_unc) shown as function of LWP for Cloud_cci (left) and AMSR-E (right). Boxes reflect minimum and maximum as well as 25%, 50% and 75% percentiles of LWP_unc distribution in each LWP bin.

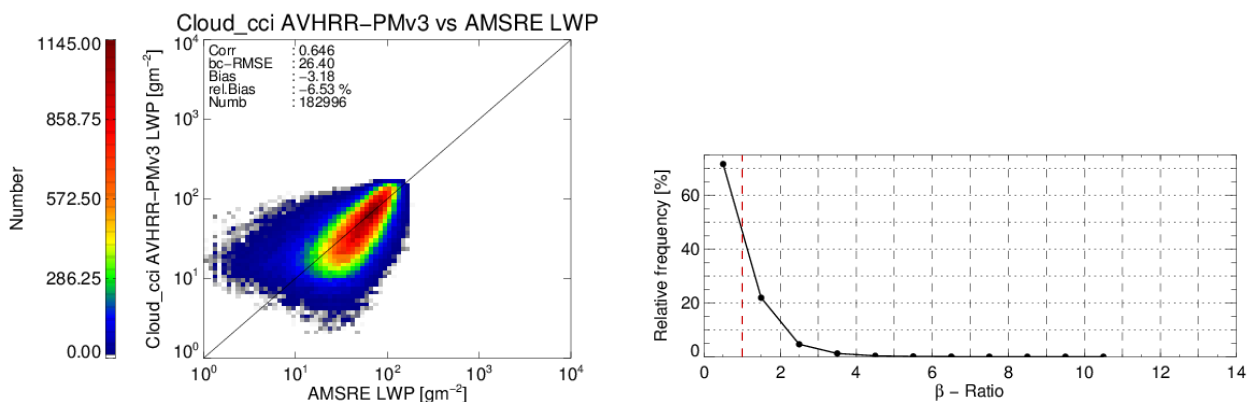


Figure 6-12 Left: 2-dimensional frequency histogram of LWP^{cci} , LWP^{amsre} pairs (about 1 million collocated pixels). Right: Histogram of β -ratio (as defined in Equation 6-6) over all collocations. Vertical dashed lines reflect bin borders (bin width is 1).

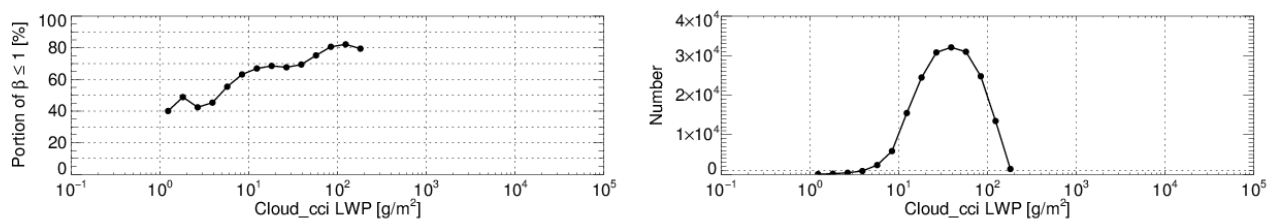


Figure 6-13 Left: Portion of all cases for which $\beta \leq 1$ (β_1 -portion) shown as function of $Cloud_cci$ LWP. Right: Corresponding number of pixels per $Cloud_cci$ LWP bin.

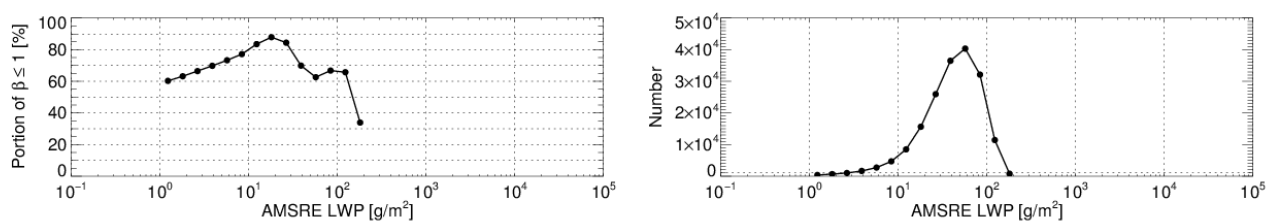


Figure 6-14 As Figure 6-18 but as function of AMSR-E LWP.

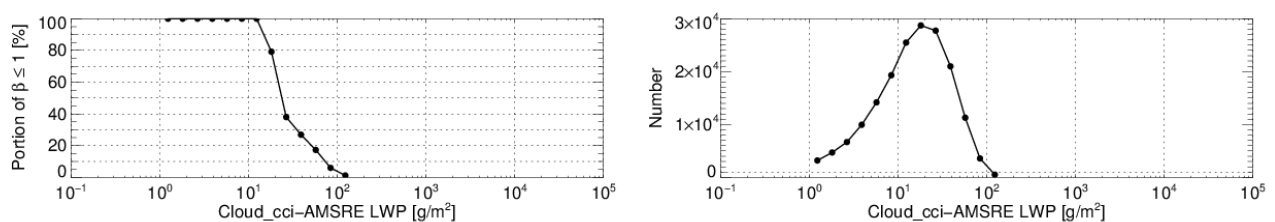


Figure 6-15 As Figure 6-18 but as function of the difference $Cloud_cci$ LWP - AMSR-E LWP.

Table 6-3 Correlation coefficients. See text for details.

(a)	Correlation ($ d $, IWP_{unc}^{cci})	0.28
(b)	Correlation ($ d $, u)	0.27
(c)	Correlation ($ d $, β)	0.84

6.3 Ice water path

As mentioned in the beginning of Section 6, DARDAR IWP data (IWP^{dardar}) is used as best estimate of the truth. IWP^{dardar} values are results of an variational scheme (Delanoë and Hogan, 2008,2010) for retrieving ice cloud properties from combined radar CloudSat, lidar (CALIOP), and infrared radiometer (e.g. MODIS) and come along with a pixel based uncertainty measure similar to Cloud_cci.

Only a very small systematic bias is found between IWP^{cci} and IWP^{dardar} . As the frequency distribution of d was already centred around 0 (not shown), no further bias correction was applied and all remaining deviations to be of random nature only.

Equations 6-7 to 6-9 list the calculations of d , u and β as defined above. For IWP uncertainty (IWP_{unc}) validation all cases for which either Cloud_cci or DARDAR reported a liquid cloud top were excluded.

$$d_i = IWP_i^{cci} - IWP_i^{dardar} \quad \text{Equation 6-7}$$

$$u_i = \sqrt{(IWP_{unc}^{cci})^2 + (IWP_{unc}^{dardar})^2} \quad \text{Equation 6-8}$$

$$\beta_i = \frac{|d_i|}{u_i} \quad \text{Equation 6-9}$$

Figure 6-16 shows boxplots for IWP_{unc}^{cci} and IWP_{unc}^{dardar} uncertainty as function of IWP. For IWP larger than 30g/m² the IWP_{unc}^{cci} grows exponentially with IWP^{cci} (near-linear relationship in logarithmic plot). Right hand side of Figure 6-16 shows the same for IWP_{unc}^{dardar} and IWP^{dardar} , indicating similarities to Cloud_cci, with the spread however being lower than for Cloud_cci. It is also important to note that IWP_{unc}^{dardar} are about one order of magnitude smaller than the retrieved IWP^{dardar} , whereas for Cloud_cci it seems that the IWP_{unc}^{cci} is nearly as large as the retrieved IWP^{cci} .

Figure 6-17 shows a 2-dimensional histogram over all IWP^{cci} , IWP^{dardar} pairs for all ice cloud pixels used. In addition a histogram over β is given. As mentioned above, the β_1 -portion should ideally amount to 68%. Thus 68% should fall into the first bin. However, the histogram only reports 31% with $\beta \leq 1$. In the second bin ($1 < \beta \leq 2$) about 23% are found. The number decrease further with increasing bin. This investigation suggests too small values for IWP_{unc}^{cci} given that the IWP_{unc}^{dardar} correctly presents the 68% confidence level itself. In the following it is investigated if the appropriateness of IWP_{unc}^{cci} is a function of IWP itself. For this, the portion of cases with $\beta \leq 1$ with respect to all cases within certain IWP bins (Figure 6-18 to Figure 6-20).


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 126

Figure 6-18 shows that for large IWP^{cci} values the IWP_{unc}^{cci} seems to have the right amplitude as the β_1 -portion increases to values above 60%. However, for IWP values below 1000g/m^3 the IWP_{unc}^{cci} seems to be too small. Figure 6-19 shows on the one hand a similar peak for very high IWP^{dardar} , on the other hand however another peak for β_1 -portion for IWP^{dardar} between 100 and 1000g/m^2 . Figure 6-20 reveals that the β_1 -portion is large enough when the differences between Cloud_cci and CALIOP are small. However, for larger IWP differences only a small portion of the data is correctly characterized by the IWP_{unc} (about 20%).

Table 6-4 reports correlation coefficients of the IWP differences ($IWP^{cci} - IWP^{dardar}$) with (a) the IWP_{unc}^{cci} (b) with the combined uncertainty (IWP_{unc}^{cci} and IWP_{unc}^{dardar}) and (c) with β . The correlation coefficients for (a) and (b) are 0.41 and 0.49. Correlation (c) is moderate (0.27).

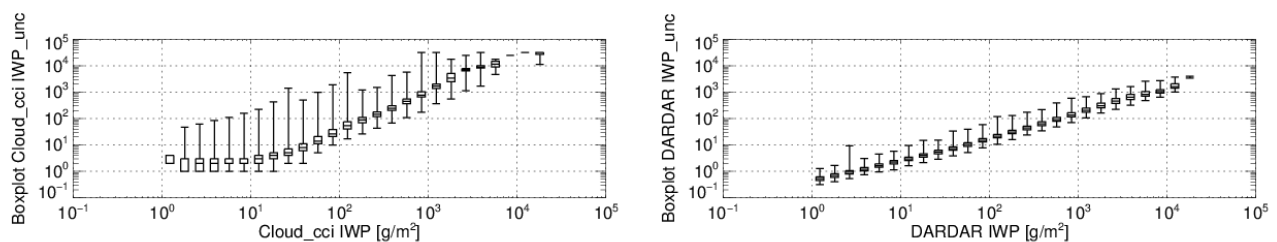


Figure 6-16 Boxplot of IWP uncertainties (IWP_{unc}) shown as function of IWP for Cloud_cci (left) and DARDAR (right). Boxes reflect minimum and maximum as well as 25%, 50% and 75% percentiles of IWP_{unc} distribution in each IWP bin.

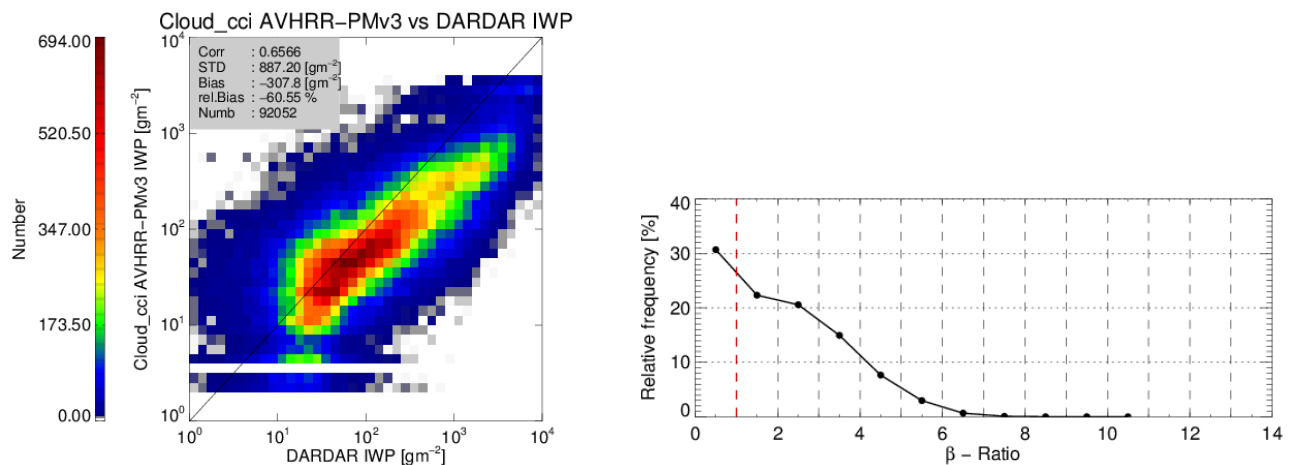


Figure 6-17 Left: 2-dimensional frequency histogram of Cloud_cci and DARDAR IWP pairs for about 14000 collocated pixels. Right: Histogram of β -ratio (as defined in Equation 6-9) over all collocations. Vertical dashed lines reflect bin borders (bin width is 1).

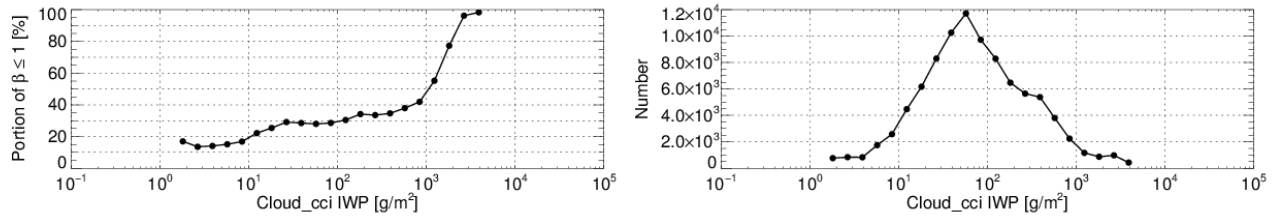


Figure 6-18 Left: Portion of all cases for which $\beta \leq 1$ (β_1 -portion) shown as function of Cloud_cci IWP. Right: Corresponding number of pixels per Cloud_cci IWP bin.

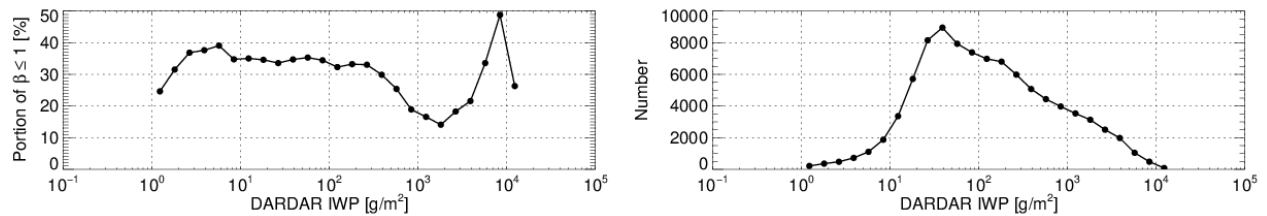


Figure 6-19 As Figure 6-18 but as function of DARDAR IWP.

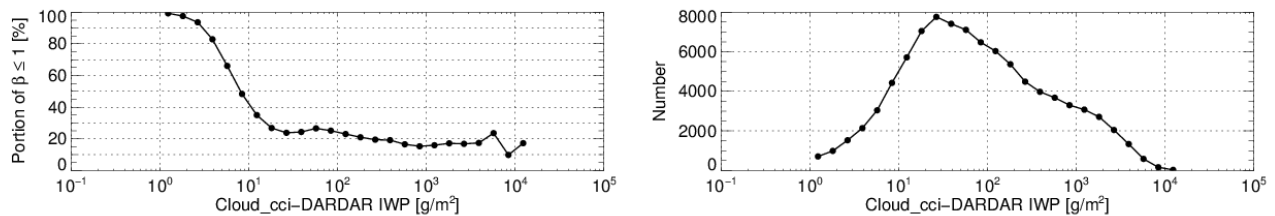



Figure 6-20 As Figure 6-18 but as function of the absolute difference Cloud-cci IWP - DARDAR IWP.

Table 6-4 Correlation coefficients. See text for details.

(a)	Correlation ($ d $, IWP_{unc}^{cci})	0.41
(b)	Correlation ($ d $, u)	0.49
(c)	Correlation ($ d $, β)	0.27

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 128


6.4 Summary of the uncertainty validation

In the previous three subsections, the Cloud_cci CTP, LWP and IWP uncertainties were validated using very accurate reference observations of CTP, LWP and IWP and, if existent, their uncertainties. As the OE uncertainty describes the interval around the retrieved value in which the truth can be found with a confidence of 68%, we determined the number of cases for which the uncertainty covered the difference to the reference data as portion from all cases.

For CTP and IWP the found portion was significantly too low (between 31% and 47%) revealing that the Cloud_cci uncertainties being generally too low for CTP and IWP. This probably means that not all uncertainty sources are (correctly) included/modelled yet. One example is the AVHRR measurement uncertainty, which is likely to be larger than currently specified. On the other hand, for LWP 68% is reported which is agreeing to the expectation.

Further studies revealed that for CTP nearly no correlation between the uncertainties and the actually difference to the estimated truth could be found, which indicates the CTP uncertainty being not a good measure of the actual uncertainty of the retrieved CTP value. In contrast, the results for LWP and IWP uncertainties indicate their utility, as the correlation between the combined LWP/IWP uncertainty (Cloud_cci + reference uncertainty) and the LWP/IWP difference (between Cloud_cci and the reference) were 0.27 (for LWP) and 0.49 (for IWP).

It needs to be noted that the uncertainties of the reference data also have an important impact on the results inferred. Having these specified too low (by us or the providers of the reference data), could also explain a significant portion of the results.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 129

7. Summary

In this document, the evaluation results for Cloud_cci v3 datasets (AVHRR-AMv3, AVHRR-PMv3 and ATSR2-AATSRv3) are presented. In Section 2 a general overview of Cloud_cci datasets was given and the derived cloud properties were introduced together with a discussion on the available processing levels.

In Section 3 the results of validating Cloud_cci v3 datasets against benchmark reference data from CALIOP (a space-based Lidar), CloudSat (space-based radar), AMSR-E (space-based microwave sensor), SYNOP (human ground cloud cover observations), BSRN (ground-based radiative flux measurements at the surface) and CERES (space-based TOA radiative flux measurements) were shown and discussed.

Cloud mask (CMA) validation against CALIOP proves the good cloud detection performance of CC4CL hitrates generally higher than 74%. For nearly all comparisons a slight underestimation of cloud occurrences is found in the Cloud_cci data compared to CALIOP, which is primarily due to a lack of sensitivity of passive imager data with respect to optically very thin clouds. Removing these optically thin clouds from the statistic, which is done by using CALIOP cloud optical thickness (COT), shows a clear improvement in the found detection scores (e.g. hitrates above 80% for all datasets).


Cloud phase (CPH) validation against CALIOP shows hitrates of 74% and higher when using the CALIOP phase of the uppermost detected cloud layer. Neglecting the very thin cloud layers (with COT lower than 0.15) improves the scores significantly: increasing hitrate scores to 77% and beyond. All datasets have a bias towards liquid clouds, when compared against to uppermost layer. The bias becomes negative for the scenario of neglecting the very thin cloud layer.

Cloud top height retrievals were also validated against CALIOP, stratified by cloud phase. For liquid clouds only very small biases (-0.1 to 0.12 km) and bc-RMSD (0.8 to 0.98 km) are found, which are only slightly sensitive to the chosen COT threshold when removing the uppermost clouds layers. For ice clouds, the strong underestimation of cloud top height is evident and a common feature for all three v3 datasets, and mainly caused by high-level, optically thin clouds. Biases are around -3.5 km and bc-RMSE around 2.3 km. Removing the optically very thin cloud layers at the top of the CALIOP profiles, improves the agreement between Cloud_cci and CALIOP significantly. When removing all cloud layers with COTs (integrated from layer bottom to top of all cloud layers) lower than 1.0, biases are reduced to approximately -1.5 km while bc-RMSE remain nearly unchanged.

Validating liquid water path over ocean against AMSR-E gives very convincing results for all three Cloud_cci v3 datasets, with bc-RMSE values of 25 to 28 g/m², only small biases and high correlations (above 0.62). Validating ice water path against the combined CALIPSO-CloudSat product DARDAR shows good agreement with correlations between 0.45 and 0.62. There is a general underestimation of IWP by Cloud_cci which in terms of relative bias partly exceeds 50%.

Cloud_cci monthly mean cloud fraction data (CFC) were validated against equivalent values from SYNOP observations. It could be shown, that all datasets reveal a high stability of the bias compared to SYNOP throughout the time period covered. An exception here is AVHRR-AM, which clearly reveals shortcomings for AVHRR on board NOAA12, the satellite that that was used for 1991 to 1999 for AVHRR-AM. As the SYNOP data was purposely not subsampled to the satellite overpass time, the comparisons include sampling errors. Comparing the biases it seems that the morning satellites do provide a better estimate of the climatological mean (24-hour mean) than the afternoon satellites.

Validating monthly mean radiative fluxes reveal an excellent agreement of Cloud_cci v3 data with CERES at top of atmosphere with standard deviations ranging from 11.6 W/m² to 15.6 W/m² and biases ranging from 3.6 W/m² to 7 W/m² for upwelling shortwave fluxes. For upwelling longwave fluxes at top of atmosphere biases are between -2 W/m² and 4 W/m² and standard deviations between 3.4 W/m² and 12.4 W/m². Generally, much higher correlations are found. Validation of BOA fluxes against BSRN stations confirms the TOA validation results by yielding standard deviations between 13 W/m² and 24 W/m² and biases ranging from 1.9 W/m² to 8.2 W/m². Correlations are again very high and above 0.96 for all three datasets. For BOA longwave fluxes, results give standard deviations of 11.4 W/m² to 14.6 W/m² and correlations from 0.97 to 0.99. BOA longwave biases are in the range of 7.3 W/m² to 11.9 W/m².

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 130

In Section 4 Cloud_cci L3C v3 cloud products were compared against their precursor versions and confronted with well-established, existing datasets ('reference datasets' hereafter) of the same kind (CLARA-A2, PATMOS-x, MODIS Collection 6.1, ISCCP). These comparisons are used for plausibility studies as none of the used external datasets can be considered as truth. The comparisons were separated into morning orbit satellites and afternoon orbit satellite, with morning and afternoon referring to the local observation time of the daytime equator-crossing node of the orbits.

The results show that Cloud_cci v3 datasets are comparable to reference datasets. They reflect the same global patterns when comparing multi-annual fields. In some comparison (e.g. COT in polar regions), the spread among all datasets (including the reference datasets) is large, which prevents any sound conclusion on the quality of the Cloud_cci datasets. However, most comparisons clearly indicate consensus between Cloud_cci v3 and the reference datasets highlighting the comparable quality of the Cloud_cci v3 datasets for most cloud properties. Larger improvements can be seen at the variables CER_{ice} and IWP, for which the Cloud_cci AVHRR v2 datasets revealed a strong overestimation, which was caused by erroneous 3.7mic LUTs for ice clouds).

The same approach was followed for the radiative flux properties, comparing Cloud_cci v3 dataset to CERES, HIRS OLR, CM-SAF and ERA-Interim. A lot of consensus was found in these comparisons demonstrating the high quality of the Cloud_cci radiative flux properties.

Time series analyses indicate that difficulties remain at the transitions from one satellite to another in each individual Cloud_cci dataset and that the drifting of the earlier NOAA satellite cause inhomogeneities in the Cloud_cci AVHRR datasets that need to be tackled by drift correction before the datasets are being used for e.g. trend studies.

The validation results were compared to GCOS requirements (see compliance matrix in Table 7-2). For nearly all validations for which a trustworthy reference data source is available, the compliance to GCOS requirements could be shown, e.g. cloud fraction stability, cloud-top height accuracy and stability (after removing optically thin cloud layers at the top), liquid water path accuracy and stability and ice water path stability. Cloud fraction accuracy is close to the GCOS target requirements for many situations. Generally, for effective radius and optical thickness no reliable reference data is available for accuracy compliance analysis. A general problem is the assessment of the stability. In this report most stability assessments are based on comparisons to MODIS, which in turn however, is not an entirely reliable source itself as it is sometime characterized by significant trends which may or may not be true. For this reason stability compliance against MODIS is often not met, however, it often is when considering a constant climate.

Furthermore, the uncertainty estimates provided along with the Cloud_cci products were evaluated - at least for those properties for which a reliable Level-2 reference exists: CTP, LWP, IWP. For CTP the uncertainty estimates do not provide a useful measure of the CTP retrieval error on pixel level. In contrast, for LWP and IWP some skill could be revealed for the corresponding pixel-level uncertainties, although the uncertainties still seem too low for the majority of all pixels. It is assumed that deficient uncertainties of the input data (Level-1 and auxiliary) are major constituents for the found shortcomings in the Level-2 uncertainties.

Based on the evaluation results documented in this report, the project team summarized recommendations on the usage of Level-2, Level-3U and Level-3C products of Cloud_cci datasets versions 3 (Table 7-1).


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 131

Table 7-1 Recommendation on the usage of Level-2, Level-3U and Level-3C products of Cloud_cci version 3 datasets.

Cloud mask (CMA) & Cloud fraction (CFC)

Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- Discrimination of heavy aerosol and cloudy is not optimal, thus aerosol is sometimes flagged as clouds in such conditions. It is advised to be careful in the interpretation cloudiness in periods with dust / volcanic ash outbreaks. Cloudiness is overestimated in these conditions.
- Cloud detection during polar night over snow and ice is generally difficult. Comparisons to other AVHRR datasets reveal that the CC4CL cloud detection has also some shortcoming during polar day although it performs better than during polar night.
- Cloud detection in twilight conditions is of rather poor quality due to the small number of channels used. Due to the orbit configurations, this issue is impacting most significantly the AVHRR-AM set before 1999.
- AVHRR-PM/AM time series of CFC should not be used for trend analysis without performing a proper satellite drift correction first.
- The ATSR2-AATSR cloud detection consistency between ATSR2 and AATSR is not optimal
- Due to a limitation of passive imagers, the cloud fraction is usually biased toward lower values, compared to example CALIOP.

Cloud phase & Liquid cloud fraction (CPH)

Recommendation for use:

Can be used (See limitations below)

Known limitations:

- AVHRR-PM/AM time series of CPH should not be used for trend analysis without performing a proper satellite drift correction first.
- ATSR-2/AATSR is biased towards ice phase and should be used carefully; and shows some instabilities in the beginning of the record.

Cloud top pressure (CTP)


Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- In semi-transparent (ice) cloud conditions, the cloud top will be assigned too low.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 132

- Multi-layer clouds are not modelled hence the CTH for cases of an upper layer of thin cirrus will effectively retrieve a radiative height (approx. 1 optical depth into the cloud).
- The supplemental products *corrected* cloud top pressure/height/temperature are not radiatively consistent. More information on these products can be found in PUGv5.1 and ATBDv6.2

Cloud optical thickness - liquid and ice clouds (COT_{liq} , COT_{ice})

Recommendation for use:

Can be used (See limitations below)

Known limitations:

- COT is a daytime product only
- In cases of wrong phase assigned, the optical thickness is likely to have significant errors.
- In the case of incorrectly assigned surface BRDF the optical depth is likely to be biased. Too high BRDF the COT will be biased low. Too low BRDF the COT will be biased high.
- In case of sub-pixel clouds or cloud borders the COT is likely to have significant errors.
- In case of optically thin clouds above (especially poorly) defined highly reflecting surface, the COT retrieval might be problematic
- For very optically thick clouds, the measurements go into saturation and thus the sensitivity of the measurement to the COT is small. Those values should be accompanied by large uncertainty values.

Cloud effective radius - liquid clouds (CER_{liq})

Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- CER is a daytime product only
- In cases of wrong phase assigned, the effective radius is likely to have significant errors.
- In case of sub-pixel clouds or cloud borders, the effective radius is likely to have significant errors.

Cloud effective radius - ice clouds (CER_{ice})


Recommendation for use:

Can be used with caution (See limitations below)

Known limitations:

All datasets:

- CER is a daytime product only
- In cases of wrong phase assigned, the effective radius is likely to have significant errors.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 133

- In case of sub-pixel clouds or cloud borders, the effective radius is likely to have significant errors.

Liquid water content (LWP)

Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- LWP is a daytime product only
- Since LWP is computed from retrieved COT and CER, same limitations as for COT and CER apply for LWP.
- The method used assumes vertically homogeneous clouds, which might deviate from truth. In case of vertically inhomogeneous cloud layers, e.g. multi-layered clouds, the LWP retrieval is likely to show large errors, since the CER is retrieved from the uppermost cloud layers and may not be representative for the entire vertical column.
- In cases of wrongly assigned cloud phase, i.e. ice cloud is treated as liquid cloud, the retrieved LWP will show large errors.

Ice water content (IWP)

Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- IWP is a daytime product only
- Similar limitations as mentioned for Cloud liquid water path

Top of atmosphere (TOA) shortwave and longwave fluxes


Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- Higher uncertainties in twilight conditions, especially in the shortwave fluxes, due to limitation in retrieving COT and CER (input to the radiation calculation) in these condition
- Partly sparse temporal/spatial sampling, partly compensated by introduced diurnal cycles correction
- Somewhat higher uncertainties expected for TOA shortwave fluxes for conditions with low clouds frequencies and elevated surface albedo uncertainties.
- Upwelling longwave fluxes seem biased low

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 134

Bottom of atmosphere (TOA) shortwave and longwave fluxes (including PAR)

Recommendation for use:

Can be used (See limitations below)

Known limitations:

All datasets:

- Higher uncertainties in twilight conditions, especially in the shortwave fluxes, due to limitation in retrieving COT and CER (input to the radiation calculation) in these condition
- Partly spare temporal/spatial sampling, partly compensated by introduced diurnal cycles correction
- Somewhat higher uncertainties expected for TOA shortwave fluxes for conditions with low clouds frequencies and elevated surface albedo uncertainties.
- Downwelling longwave fluxes seem biased high



	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 135

Table 7-2 Compliance matrix confronting validation results of Cloud_cci data with GCOS requirements. The latter are taken from Section 2.2.1 of URDv3 (2018). Green shaded cell indicate compliance with the requirements, yellow cells nearly compliance and red cell no compliances. However, only for validation scores reported in black bold font, the reference data is trusted to provide an accurate reference for compliance analysis. For some cloud properties and datasets no validation ('n/v') was possible. Last column provides tractability of the used validation scores reported earlier in this document.

		GCOS target requirements	Achieved Cloud_cci accuracy			Comment
			AVHRR-AMv3	AVHRR-PMv3	ATSR2-AATSRv3	
Cloud cover	accuracy	5%	-6.2 % ²	-5.4 % ²	-5.1 % ²	Level-2 validation against CALIOP
	stability ¹ (per decade)	3%	0.65%	0.07%	-0.52%	Values ⁴ taken from Table 4-2 and Table 4-13 (L3C comparisons to MODIS C6.1)
Cloud top height/ pressure	accuracy ⁵ (low/mid/high)	0.5/0.7/1.6km	-0.05km ⁶ -1.25km ⁷	-0.1km ⁶ -1.62km ⁷	0.12km ⁶ -1.76km ⁷	Level-2 validation against CALIOP
	stability (per decade)	15hPa	3.45hPa	11.25hPa	0.45hPa	Values ⁴ taken from Table 4-4 and Table 4-15 (L3C comparisons to MODIS C6.1)
Liquid cloud optical depth	accuracy	10%	n/v	n/v	n/v	No validation possible due to a lack of reliable reference data. through LWP and IWP validation
	stability (per decade)	2%	-0.05%	0.07%	-0.03%	Values ⁴ taken from Table 4-5 and Table 4-16 divided by mean MODIS C6.1 Terra COT _{liq} (13) (L3C comparisons to MODIS C6.1)
Ice cloud optical depth	accuracy	10%	n/v	n/v	n/v	No validation possible due to a lack of reliable reference data. through LWP and IWP validation
	stability (per decade)	2%	0.02%	-0.08%	-0.05%	Values ⁴ taken from Table 4-6 and Table 4-17 divided by mean MODIS C6.1 Terra COT _{ice} (10) (L3C comparisons to MODIS C6.1)
Liquid water path	accuracy	25%	4.9%	-8.5%	-2.4	Level-2 validation against AMSR-E (Figure 3-1)
	stability (per decade)	5%	-0.08%	0.09%	-0.06%	Values ⁴ taken from Table 4-9 and Table 4-20 divided by mean MODIS C6 LWP (123g/m ²) (L3C comparisons to MODIS C6)

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 136

Ice water path	accuracy	25%	-49%	-59%	-39.9%	Level-2 validation against DARDAR (Figure 3-2)
	stability (per decade)	5%	0.04%	0.03%	-0.04%	Values ⁴ taken from Table 4-10 and Table 4-21 divided by mean MODIS C6 IWP (208g/m ²) (L3C comparisons to MODIS C6)
Cloud liquid effective particle radius	accuracy	10%	n/v	n/v	n/v	No validation possible due to a lack of reliable reference data. through LWP and IWP validation
	stability (per decade)	1µm	-1.33µm	-0.1µm	-0.96µm	Values ⁴ taken from Table 4-7 and Table 4-18 (L3C comparisons to MODIS C6)
Cloud ice effective particle radius	accuracy	10%	n/v	n/v	n/v	No validation possible due to a lack of reliable reference data. through LWP and IWP validation
	stability (per decade)	1µm	0.81µm	-0.77 µm	-0.33µm	Values ⁴ taken from Table 4-8 and Table 4-19 (L3C comparisons to MODIS C6)

¹ The "stability" figures provided are a first tentative attempt to quantify stability via intercomparison of the long term linear trend against MODIS. Needs to be noted that instability of the MODIS reference, or other aspects such as non-linear instabilities in the CDR, are not accounted for by this method.

² after applying COD threshold


³ global ocean

⁴ derived with respect to MODIS C6 which itself reveals a significant (negative) trend for some cloud properties

⁵ As most of the validation results (discussed in the following) are given in terms of CTH, which is also the direct measurement of CALIOP data used as validation source, we defined CTH requirements by converting GCOS CTP accuracy requirements CTH via the US standard atmosphere (1976) profiles, obtained from <https://www.digitaldutch.com/atmoscalc/>, using the $\Delta H/\Delta P$ derivative at three representative levels for low, mid-level and high clouds. For mid-level clouds the middle pressure level between the 680hPa and 440hPa borders was chosen: 560hPa. For low and high clouds distinct peaks exist in frequency distributions of global CALIOP CTP measures around 850hPa and 200hPa. At these two levels the conversion from GCOS CTP requirements for low and high clouds to CTH requirements were used.


⁶ all liquid clouds after applying a COD threshold of 1.0

⁷ all ice clouds after applying a COD threshold of 1.0


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 137

8. Glossary


AATSR	Advanced Along Track Scanning Radiometer
ACDD	Attribute Convention for Dataset Discovery
AIRS	Atmospheric Infrared Sounder
ATBD	Algorithm Theoretical Baseline Document
ATSR2	Along-Track Scanning Radiometer 2
AVHRR	Advanced Very High Resolution Radiometer
BOA	Bottom of the Atmosphere
BRDF	Bidirectional Reflectance Distribution Function
Bc-RMSE	Bias-corrected (bc) Root Mean Square Error
Bc-RMSD	Bias-corrected (bc) Root Mean Square Difference
BSRN	Baseline Surface Radiation Network
C6	(MODIS) Collection 6
C6.1	(MODIS) Collection 6.1
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization
CC4CL	Community Cloud retrieval for Climate
CCI	Climate Change Initiative
CDR	Climate Data Record
CER	Cloud Effective Radius
CERES	Clouds and the Earth Radiant Energy System
CF	Climate and Forecast
CFC	Cloud Fractional Coverage
CFMIP	Cloud Feedback Model Intercomparison Project
CLA	Spectral Cloud Albedo
CLARA-A2	CM SAF cLOUDs, Albedo and Radiation dataset from AVHRR data - Edition 2

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 138


CMA	Cloud Mask
CM SAF	Satellite Application Facility on Climate Monitoring
COSP	CFMIP Observation Simulator Package
CPH	Cloud Phase
COSP	CFMIP Observation Simulator Package
COT	Cloud Optical Thickness
CTH	Cloud Top Height
CTP	Cloud Top Pressure
CTT	Cloud Top Temperature
CWP	Cloud Water Path
DCHP-A	Daytime Cloud Height Property AATSR
DCHP-M	Daytime Cloud Height Property MERIS
DOI	Digital Object Identifier
DSRDP	Data Standards Requirements for CCI Data
DWD	Deutscher Wetterdienst
EBAF	Energy Balanced and Filled
ECV	Essential Climate Variable
ENVISAT	Environmental Satellite
ERB	Earth Radiation Budget
ERS	European Remote Sensing
ESA	European Space Agency
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAME-C	FUB AATSR MERIS Cloud retrieval algorithm

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 139

FAR	False alarm rate
FUB	Freie Universität Berlin
FCDR	Fundamental Climate Data Record
GAC	Global Area Coverage - globally available AVHRR dataset with reduced resolution (4 km).
GEOS	Goddard Earth Observation System
GEWEX	Global Energy and Water Cycle Experiment
GMT	Greenwich Mean Time
HIRS	High Resolution Infrared Radiation Sounder
JCH	Joint Cloud property Histogram
ISCCP	International Satellite Cloud Climatology Project
IWP	Ice Water Path
ITCZ	Inter-Tropical Convergence Zone
LCF	Liquid Cloud Fraction
LUT	Look-up Table
LW	Longwave
LWP	Liquid Water Path
MERIS	Medium Resolution Imaging Spectrometer
Metop	Meteorological Operational Satellite
MODIS	Moderate Resolution Imaging Spectroradiometer
MW	Microwave
NCEI	National Centers for Environmental Information
NIR	Near-Infrared
NOAA	National Oceanic & Atmospheric Administration


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 140

OE	Optimal Estimation
OLR	Outgoing Longwave Radiation
PATMOS-x	AVHRR Pathfinder Atmospheres-Extended dataset
POD	Probability of detection
PUG	Product User Guide
PVIR	Product Validation and Intercomparison Report
RAL	Rutherford Appleton Laboratory
SDL	Surface Downwelling Longwave Radiation
SIS	Surface Incoming Shortwave Radiation
SOL	Surface Outgoing Longwave Radiation
SW	Shortwave
SYNOP	Surface synoptic observations
TCDR	Thematic Climate Data Record
TIROS	Television and Infra-Red Observation Satellite
TOA	Top of the Atmosphere
VIIRS	Visible Infrared Imaging Radiometer Suite
WCRP	World Climate Research Programme
WRMC	World Radiation Monitoring Center
UWISC	University of Wisconsin


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 141

9. References

- Aguilar, E., Auer, I., Brunet, M., Peterson, T.C., Wieringa, 2003. Guidance on metadata and homogenization, WMO TD N. 1186 (WCDMP N. 53), 51pp.
- Alexandersson, H., 1986. A homogeneity test applied to precipitation data. J. Climatol. 6, 661-675. doi:10.1002/joc.3370060607
- ATBDv5.1, Algorithm Theoretical Baseline Document (ATBD) - ESA Cloud_cci, Issue 5, Revision: 1, date of Issue: 12/09/2017, Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- ATBDv6.2, Algorithm Theoretical Baseline Document (ATBD) - ESA Cloud_cci, Issue 6, Revision: 2, date of issue: 14/10/2019, Available at: http://www.esa-cloud-cci.org/?q=documentation_v3
- ATBD-FAME-Cv5, Algorithm Theoretical Baseline Document (ATBD) FAME-C - ESA Cloud_cci, Issue 5, Revision: 0, date of Issue: 12/09/2017, Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- ATBD-CC4CLv5, Algorithm Theoretical Baseline Document (ATBD) CC4CL - ESA Cloud_cci, Issue 5, Revision: 0, date of Issue: 12/09/2017, Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- ATBD-CC4CLv6.2, Algorithm Theoretical Baseline Document (ATBD) CC4CL - ESA Cloud_cci, Issue 6, Revision: 1, date of Issue: 18/10/2019, Available at: http://www.esa-cloud-cci.org/?q=documentation_v3
- ATBD-CC4CL_TOA_FLUXv1.1, Algorithm Theoretical Basis Document (ATBD) of the Community Code for CLimate (CC4CL) Broadband Radiative Flux Retrieval (CC4CL-TOAFLUX) - ESA Cloud_cci, Issue 1, Revision: 1, date of Issue: 14/10/2019, Available at: http://www.esa-cloud-cci.org/?q=documentation_v3
- Baum, B. A., Menzel, W. P., Frey, R. A., Tobin, D. C., Holz, R. E., Ackerman, S. A., Heidinger, A. K., and Yang, P.: MODIS cloud-top property refinements for collection 6, Journal of applied meteorology and climatology, 51, 1145-1163, 2012.
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57, 289-300.
- Berrisford, P., Dee, D., Poli, P., Brugge, R., Fielding, K., Fuentes, M., Kallberg, P., Kobayashi, S., Uppala, S. and Simmons, A., 2011: The ERA-Interim archive version 2.0, ERA report series. 1. Technical Report: pp23. ECMWF.
- Bodas-Salcedo, A., Webb, M.J., Bony, S., Chepfer, H., Dufresne, J.L., Klein, S.A., Zhang, Y., Marchand, R., Haynes, J.M., Pincus, R. and John, V.O., 2011. COSP: Satellite simulation software for model assessment. Bulletin of the American Meteorological Society, 92(8), p.1023.
- Carbajal Henken, C.K., Lindstrot, R., Preusker, R. and Fischer, J.: FAME-C: cloud property retrieval using synergistic AATSR and MERIS observations. Atmos. Meas. Tech., 7, 3873-3890, doi:10.5194/amt-7-3873-2014, 2014
- CARv3, Climate Assessment Report (CAR) - ESA Cloud_cci, Issue 3, Revision: 0, planned date of Issue: 06/2017, Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- CECRv3, Comprehensive Error Characterization Report (CECR) - ESA Cloud_cci, Issue 3, Revision: 1, Date of Issue: 07/03/2017 Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. & Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteorol. Soc., 137, 553-597, 2011.
- 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.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 142

- Delanoë, J., and R. J. Hogan, 2010:** Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *J. Geophys. Res.*, 115, D00H29, doi:10.1029/2009JD012346.
- Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., Wielicki, B. A., Young, D. F. and M. Sun, 2013:** Geostationary enhanced temporal interpolation for CERES flux products. *J. Atmos. Oceanic Technol.*, 30: 1072-1090.
- Driemel, A., Augustine, J., Behrens, K., Colle, S., Cox, C., Cuevas-Agulló, E., Denn, F. M., Duprat, T., Fukuda, M., Grobe, H., Haefelin, M., Hodges, G., Hyett, N., Ijima, O., Kallis, A., Knap, W., Kustov, V., Long, C. N., Longenecker, D., Lupi, A., Maturilli, M., Mimouni, M., Ntsangwane, L., Ogihara, H., Olano, X., Olefs, M., Omori, M., Passamani, L., Pereira, E. B., Schmithüsen, H., Schumacher, S., Sieger, R., Tamlyn, J., Vogt, R., Vuilleumier, L., Xia, X., Ohmura, A., and König-Langlo, G, 2018:** Baseline Surface Radiation Network (BSRN): structure and data description (1992-2017), *Earth Syst. Sci. Data*, 10: 1491-1501, doi:10.5194/essd-10-1491-2018
- González, A., 2009:** Measurement of Areas on a Sphere Using Fibonacci and Latitude--Longitude Lattices. *Mathematical Geosciences*. 42 (1), 49-64. doi:10.1007/s11004-009-9257-x.
- Feofilov, A. G., C. J. Stubenrauch, S. Protopapadaki, and R. Armante, 2017:** Diurnal variation of high-level clouds from a synergy of the space-borne infrared sounders AIRS and IASI: detection and radiative effects. In preparation for submission to ACPD.
- Häder, Donat-P., Michael Lebert, Roberto Marangoni, and Giuliano Colombetti. "ELDONET—European light dosimeter network hardware and software." *Journal of Photochemistry and Photobiology B: Biology* 52, no. 1-3 (1999): 51-58.**
- Häder, Donat-Peter, Michael Lebert, Martin Schuster, Lineu del Ciampo, E. Walter Helbling, and Richard McKenzie. "ELDONET—a decade of monitoring solar radiation on five continents." *Photochemistry and photobiology* 83, no. 6 (2007): 1348-1357.**
- Hai-Tien, L. & NOAA CDR Program, 2018:** NOAA Climate Data Record (CDR) of Monthly Outgoing Longwave Radiation (OLR), Version 2.7. NOAA National Centers for Environmental Information. <https://doi.org/10.7289/V5W37TKD> [12/2018].
- Heidinger, A. K., M.D. Goldberg, A. Jelenak and M.J. Pavolonis, 2005:** A new AVHRR cloud climatology, *Proc. SPIE* 5658, 197, doi: 10.1117/12.579047.
- Heidinger, A. K., M.J. Pavolonis, 2009:** Gazing at Cirrus Clouds for 25 Years through a Split Window. Part I: Methodology. *J. Appl. Meteor. Climatol.*, 48, 1100-1116. doi: 10.1175/2008JAMC1882.1
- Heidinger, A. K., Straka, W. C., Molling, C. C., Sullivan, J. T. and X. Q. Wu (2010):** Deriving an inter-sensor consistent calibration for the AVHRR solar reflectance data record. *Int. J. Rem. Sens.*, 31: 6493-6517.
- Heidinger, A. K., Evan, A. T., Foster, M., and Walther, A., 2012:** A Naïve Bayesian cloud detection scheme derived from CALIPSO and applied within PATMOS-x. *Journal of Applied Meteorology and Climatology*, 51, 1129-1144.
- Heidinger, A. K., Foster, M. J., Walther, A., and Zhao, X.: The pathfinder atmospheres-extended AVHRR climate dataset, *Bulletin of the American Meteorological Society*, 95, 909-922, 2014.**
- Hollmann, R., Merchant, C.J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R. and Holzer-Popp, T., 2013.** The ESA climate change initiative: Satellite data records for essential climate variables. *Bulletin of the American Meteorological Society*, 94(10), pp.1541-1552.
- Karlsson, K. -G. & E. Johansson (2013):** On the optimal method for evaluating cloud products from passive satellite imagery using CALIPSO-CALIOP data: example investigating the CM SAF CLARA-A1 dataset. *Atm. Meas. Tech.*, 6, 1271-1286, <http://dx.doi.org/10.5194/amt-6-1271-2013>
- Karlsson, K.-G., Anttila, K., Trentmann, J., Stengel, M., Meirink, J. F., Devasthale, A., Hanschmann, T., Kothe, S., Jääskeläinen, E., Sedlar, J., Benas, N., van Zadelhoff, G.-J., Schlundt, C., Stein, D.,**


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 143

Finkensieper, S., Håkansson, N., Hollmann, R., Fuchs, P. and M. Werscheck (2017): CLARA-A2: CM SAF cCloud, Albedo and surface Radiation dataset from AVHRR data - Edition 2, Satellite Application Facility on Climate Monitoring, DOI: https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V002.


- Karlsson, K.-G., Anttila, K., Trentmann, J., Stengel, M., Meirink, J. F., Devasthale, A., Hanschmann, T., Kothe, S., Jääskeläinen, E., Sedlar, J., Benas, N., van Zadelhoff, G.-J., Schlundt, C., Stein, D., Finkensieper, S., Håkansson, N., and Hollmann, R.:** CLARA-A2: The second edition of the CM SAF cloud and radiation data record from 34 years of global AVHRR data, *Atmospheric Chemistry and Physics Discussions*, **2016**, 1-41, doi:10.5194/acp-2016-935, <http://www.atmos-chem-phys-discuss.net/acp-2016-935/>, 2016.
- Kato, S., Rose, F. G., Rutan, D. A., Thorsen, T. E., Loeb, N. G. Doelling, D. R., Huang, X., Smith, W. L., Su, W. and S.-H. Ham,** **2018:** Surface irradiances of Edition 4.0 Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) data product, *J. Climate*, **31**: 4501-4527.
- Kendall, M.G., 1938.** A New Measure of Rank Correlation. *Biometrika* **30**, 81-93.
- Khaliq, M.N., Ouarda, T.B.M.J., 2007.** On the critical values of the standard normal homogeneity test (SNHT). *Int. J. Climatol.* **27**, 681-687. doi:10.1002/joc.1438
- Lebert, Michael, Martin Schuster, and Donat-Peter Häder.** "The European light dosimeter network: four years of measurements." *Journal of Photochemistry and Photobiology B: Biology* **66**, no. 1 (2002): 81-87.
- Loeb, N. G., Priestley, K. J., Kratz, D. P., Geier, E. B., Green, R. N., Wielicki, B. A., Hinton, P. O. R. and S. K. Nolan, 2001:** Determination of unfiltered radiances from the Clouds and the Earth's Radiant Energy System (CERES) instrument. *J. Appl. Meteor.*, **40**: 822-835.
- Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., Manalo-Smith, N. and T. Wong, 2009:** Toward optimal closure of the Earth's top-of-atmosphere radiation budget, *J.Climate*, **22**: 748-756. Doi: <https://doi.org/10.1175/2008JCLI2637.1>.
- Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., Liang, L., Mitrescu, C., Rose, F. G., and S. Kato, 2018:** Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Top-of-Atmosphere (TOA) Edition-4.0 Data Product. *J. Climate*, **31**: 895-918. Doi: https://doi.org/10.5067/Terra+Aqua/CERES/EBAF-TOA_L3B004.0.
- Mann, H.B., 1945.** Nonparametric Tests Against Trend. *Econometrica* **13**, 245-259.
- Marangoni, Roberto, Domenico Gioffré, Giuliano Colombetti, Michael Lebert, and Donat-P. Häder.** "ELDONET-European Light Dosimeter Network: Structure and functions of the ELDONET server." *Journal of Photochemistry and Photobiology B: Biology* **58**, no. 2-3 (2000): 178-184.
- Marchant, B., Platnick, S., Meyer, K., Arnold, G. T., and Riedi, J.:** MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP, *Atmospheric Measurement Techniques*, **9**, 1587-1599, doi:10.5194/amt-9-1587-2016, <http://www.atmos-meas-tech.net/9/1587/2016/>, **2016**.
- McGarragh, G. R., Poulsen, C. A., Thomas, G. E., Povey, A. C., Sus, O., Stapelberg, S., Schlundt, C., Proud, S., Christensen, M. W., Stengel, M., Hollmann, R., and Grainger, R. G.:** The Community Cloud retrieval for CLimate (CC4CL) - Part 2: The optimal estimation approach, *Atmos. Meas. Tech.*, **11**, 3397-3431, <https://doi.org/10.5194/amt-11-3397-2018>, **2018**.
- O'Dell, C.W., Wentz, F.J. and Bennartz, R., 2008.** Cloud liquid water path from satellite-based passive microwave observations: A new climatology over the global oceans. *Journal of Climate*, **21**(8), pp.1721-1739.
- Pavolonis, M.J., A.K. Heidinger, T.Uttal, 2005:** Daytime Global Cloud Typing from AVHRR and VIIRS: Algorithm Description, Validation, and Comparisons. *J. Appl. Meteor.*, **44**, 804-826. doi: 10.1175/JAM2236.1
- Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007.** Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.* **11**, 1633-1644. doi:10.5194/hess-11-1633-2007

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:	03 February 2020		
	Issue:	6	Revision:	1
Page 144				

- Platnick, S., King, M. and Hubanks, P., **2017a**. MODIS Atmosphere L3 Monthly Product. NASA MODIS Adaptive Processing System, Goddard Space Flight Center. doi:10.5067/MODIS/MOD08_M3.061; doi:10.5067/MODIS/MYD08_M3.061.
- Platnick, S., King, M. D., Meyer, K. G., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Ridgway, B., et al.: MODIS cloud optical properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 product and associated Level-3 datasets, Version 0.9 (beta), 17, **2014**.
- Platnick, S., Meyer, K.G., King, M.D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G.T., Zhang, Z., Hubanks, P.A., Holz, R.E. and Yang, P., **2017b**. The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua. *IEEE Transactions on Geoscience and Remote Sensing*, 55(1), pp.502-525.
- PUGv3.1, Product User Guide (PUG) - ESA Cloud_cci, Issue 3, Revision: 1, Date of Issue: 18/04/2017 Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- PUGv5.1, Product User Guide (PUG) - ESA Cloud_cci, Issue 5, Revision: 1, Date of Issue: 19/12/2019 Available at: http://www.esa-cloud-cci.org/?q=documentation_v3
- Rossow, W.B. and Schiffer, R.A., **1991**. ISCCP cloud data products. Bulletin of the American Meteorological Society, 72(1), pp.2-20.
- Rossow, W.B., Walker, A.W., Beuschel, D.E. and Roiter, M.D., **1996**. International Satellite Cloud Climatology Project (ISCCP) documentation of new cloud datasets. WMO/TD, 737, p.115.
- Rossow, William B., and Robert A. Schiffer. "Advances in understanding clouds from ISCCP." Bulletin of the American Meteorological Society 80, no. 11 (**1999**): 2261.
- RAFCDRv1.0, Technical Report on AVHRR GAC FCDR generation - ESA Cloud_cci, Issue 1, Revision: 0, planned date of Issue: 06/2017. Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- RODCv1.0, Report on Orbital Drift Correction for AVHRR - ESA Cloud_cci, Issue 1, Revision: 0, planned date of Issue: 06/2017. Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- Stengel, M., Stapelberg, S., Sus, O., Schlundt, C., Poulsen, C., Thomas, G., Christensen, M., Carbajal Henken, C., Preusker, R., Fischer, J., Devasthale, A., Willén, U., Karlsson, K.-G., McGarragh, G. R., Proud, S., Povey, A. C., Grainger, R. G., Meirink, J. F., Feofilov, A., Bennartz, R., Bojanowski, J. S., and Hollmann, R.: Cloud property datasets retrieved from AVHRR, MODIS, AATSR and MERIS in the framework of the Cloud_cci project, Earth Syst. Sci. Data, 9, 881-904, <https://doi.org/10.5194/essd-9-881-2017>, **2017**.
- Stengel, M., Schlundt, C., Stapelberg, S., Sus, O., Eliasson, S., Willén, U., and Meirink, J. F.: Comparing ERA-Interim clouds with satellite observations using a simplified satellite simulator, Atmos. Chem. Phys., 18, 17601-17614, <https://doi.org/10.5194/acp-18-17601-2018>, **2018**.
- Stengel, M., Stapelberg, S., Sus, O., Finkensieper, S., Würzler, B., Philipp, D., Hollmann, R., Poulsen, C., Christensen, M., and McGarragh, G.: Cloud_cci Advanced Very High Resolution Radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties, Earth Syst. Sci. Data, 12, 41-60, <https://doi.org/10.5194/essd-12-41-2020>, **2020**.
- Stubenrauch, C.J., W.B. Rossow, S. Kinne, S. Ackerman, Cesana G., H. Chepfer, et al. Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel, Bulletin of the American Meteorological Society (**2013**) <http://dx.doi.org/10.1175/BAMS-D-12-00117>
- Stubenrauch, C. J., A. G. Feofilov, S. E. Protopapadaki, R. Armante, **2017**: Cloud climatologies from the InfraRed Sounders AIRS and IASI: Strengths, Weaknesses and Applications In preparation for submission to ACPD
- Su, W., Corbett, J., Eitzen, Z. and L. Liang, **2015**: Next-generation angular distribution models for top-of-atmosphere radiative flux calculation from CERES instruments: methodology. Atmos. Meas. Tech., 8 (2): 611-632. Doi: <http://dx.doi.org/10.5194/amt-8-611-2015>.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 145

- Sus, O., Stengel, M., Stapelberg, S., McGarragh, G., Poulsen, C., Povey, A. C., Schlundt, C., Thomas, G., Christensen, M., Proud, S., Jerg, M., Grainger, R., and Hollmann, R.: The Community Cloud retrieval for CLimate (CC4CL) - Part 1: A framework applied to multiple satellite imaging sensors, *Atmos. Meas. Tech.*, 11, 3373-3396, <https://doi.org/10.5194/amt-11-3373-2018>, **2018**.
- Theil, H., **1950**. A rank-invariant method of linear and polynomial regression analysis. Parts 1-3. *Ned Akad Wetensch* 53, 386-392, 521-525, 1397-1412.
- Thomas, S.M., A.K. Heidinger, and M.J. Pavolonis, **2004**: Comparison of NOAA's Operational AVHRR-Derived Cloud Amount to Other Satellite-Derived Cloud Climatologies. *J. Climate*, 17, 4805-4822. doi: 10.1175/JCLI-3242.1
- Toreti, A., Kuglitsch, F.G., Xoplaki, E., Della-Marta, P.M., Aguilar, E., Prohom, M., Luterbacher, J., **2011**. A note on the use of the standard normal homogeneity test to detect inhomogeneities in climatic time series. *Int. J. Climatol.* 31, 630-632. doi:10.1002/joc.2088
- URDv3, User Requirements Document (URD) - ESA Cloud_cci, Issue 3, Revision: 0, Date of Issue: 12/02/2018 Available at: <http://www.esa-cloud-cci.org/?q=documentation>
- Vaughan, M., Powell, K., Kuehn, R., Young, S., Winker, D., Hostetler, C., Hunt, W., Liu, Z., McGill, M., and Getzewich, B., **2009**: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements, *J. Atmos. Oceanic Technol.*, 26, 2034-2050, doi: 10.1175/2009JTECHA1228.1.
- Wang, T., Fetzer, E.J., Wong, S., Kahn, B.H. and Yue, Q., **2016**. Validation of MODIS cloud mask and multilayer flag using CloudSat-CALIPSO cloud profiles and a cross-reference of their cloud classifications. *Journal of Geophysical Research: Atmospheres*, 121(19).
- Wentz, F. and T. Meissner, **2000**. AMSR Ocean Algorithm. Algorithm Theoretical Basis Document, Version 2. Santa Rosa, California USA: Remote Sensing Systems. Available at: <https://eosps.gsfc.nasa.gov/sites/default/files/atbd/atbd-amr-ocean.pdf>
- Wentz, F. J. and T. Meissner. **2004**. *AMSR-E/Aqua L2B Global Swath Ocean Products derived from Wentz Algorithm, Version 2*. [Indicate subset used]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center. doi: http://dx.doi.org/10.5067/AMSR-E/AE_OCEAN.002. [Date Accessed: 07 Sept. 2017].
- Wielicki, B. A., Barkstrom, B. R., Harrison, E. F., Lee III, R. B., Smith, G. L. and J. E. Cooper, **1996**: Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment, *Bull. Amer. Meteor. Soc.*, 77: 853-868.
- Winker, D. M., Vaughan, M.A., Omar, A., Hu, Y., Powell, K.A., Liu, Z., Hunt, W.H., and Young, S.A., **2009**: Overview of the CALIPSO mission and CALIOP data processing algorithms, *J. Atmos. Oceanic. Technol.*, 26, 2310-2323, doi:10.1175/2009JTECHA1281.1.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 146

Annex A - Datasets for comparisons with Cloud_cci products

A.1 CALIPSO-CALIOP

Measurements from space-born active instruments (radar + lidar) provide probably the most accurate information we can get about cloud presence in the atmosphere. The reason is the fact that the measured reflected radiation comes almost exclusively from cloud and precipitation particles and is therefore not “contaminated” by radiation from other surfaces or atmospheric constituents as is the case for measurements from most passive radiometers. In this validation study we have decided to utilise measurements from the CALIOP lidar instrument carried by the CALIPSO satellite (included in the A-Train series of satellites - Figure A-1).

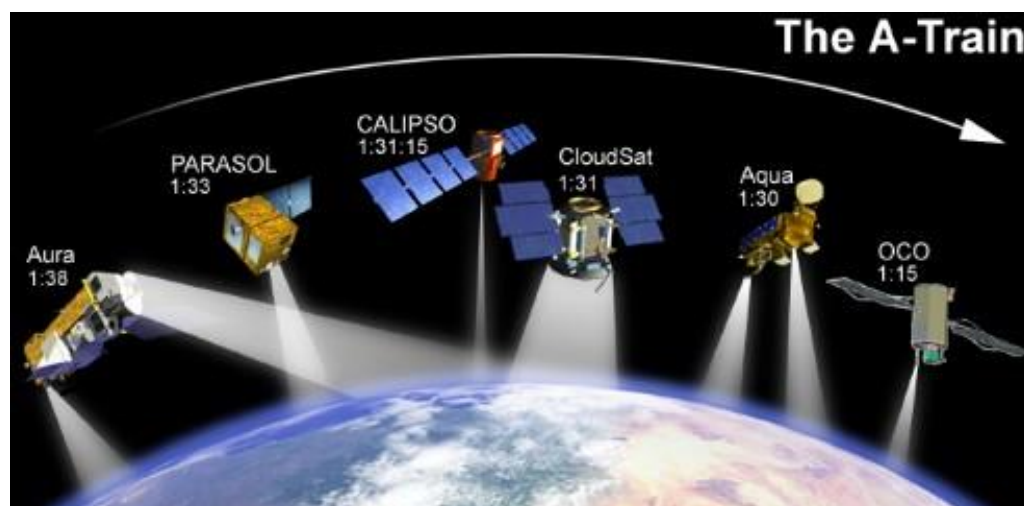



Figure A-1 The Aqua-Train satellites. (Image credit: NASA)

The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite was launched in April 2006 together with CloudSat. The satellite carries the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) and the first data became available in August 2006 ([Winker et al., 2009](#)). CALIOP provides detailed profile information about cloud and aerosol particles and corresponding physical parameters ([Vaughan et al., 2009](#)).

CALIOP measures the backscatter intensity at 1064 nm while two other channels measure the orthogonally polarized components of the backscattered signal at 532 nm. The CALIOP cloud product we have used report observed cloud layers i.e., all layers observed until signal becomes too attenuated. In practice the instrument can only probe the full geometrical depth of a cloud if the total optical thickness is not larger than a certain threshold (somewhere in the range 3-5). For optically thicker clouds only the upper portion of the cloud will be sensed. The horizontal resolution of each single FOV is 333 m and the vertical resolution is 30-60 m.

The CALIOP products are available in five different versions with respect to the along-track resolution ranging from 333 m (individual footprint resolution), 1 km, 5 km, 20 km and 80 km. The four latter resolutions are consequently constructed from several original footprints/FOVs. This allows a higher confidence in the correct detection and identification of cloud and aerosol layers compared to when using the original high resolution profiles. For example, the identification of very thin Cirrus clouds is more reliable in the 5 km data record than in the 1 km data record since signal-to-noise levels can be raised by using a combined data record of several original profiles.

We used the CALIOP level-2 1 km and 5 km cloud layer data record versions 3-01, 3-02 and 3-30 (CALIPSO Science Team, 2015) for the validation purpose. The 5 km resolution data record is closest to the nominal AVHRR GAC resolution but according to [Karlsson and Johansson \(2013\)](#) there are some inconsistencies

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 147

between results for the two resolutions which means that the total cloud amounts from the 5 km is often slightly underestimated. It means that some of the thick (opaque) boundary layer clouds that are reported in fine resolution (333 m and 1 km) data records are not reported in the higher resolution (5 km or higher) data records. This has to do with the methodology to do averaging at the longer scales (5 km or higher) where contributions from strongly reflecting clouds are removed from the original signal to facilitate detection of very thin cloud layers and aerosols. Thus, we use here the method proposed by [Karlsson and Johansson \(2013\)](#) combining the two CALIPSO data records (i.e., adding missed clouds at 5 km resolution which are detected at 1 km resolution). This normally gives almost 5 % higher global cloud amounts compared to if just relying on 5 km data.

The CALIOP cloud layer product reports up to 10 cloud layers per column and provides information about cloud phase and cloud type of each layer as well as the pressure, height and temperature at each layer's top.


The CALIOP data record classifies cloud layers into cloud types according to Table A-1. To be noticed here is that the ISCCP cloud type method has been used in the sense that the vertical separation of Low (categories 0-3), Medium (categories 4-5) and High (categories 6-7) clouds is defined by use of vertical pressure levels of 680 hPa and 440 hPa. However, the separation of thin and thick clouds is made using the information on whether the surface or lower layers below the current layer can be seen by CALIOP.

Table A-1 *Cloud type categories according to the CALIOP Vertical Feature Mask product*

Category 0	Low, overcast, thin (transparent St, StCu, and fog)
Category 1	Low, overcast, thick (opaque St, StCu, and fog)
Category 2	Transition stratocumulus
Category 3	Low, broken (trade Cu and shallow Cu)
Category 4	Alto cumulus (transparent)
Category 5	Altostratus (opaque, As, Ns, Ac)
Category 6	Cirrus (transparent)
Category 7	Deep convective (opaque As, Cb, Ns)

We only give a quite general description of the CALIPSO data records in this section. The details concerning the actual use of the data records are elaborated further in the following sections 6.1.1.2 and 6.1.2.1.

It should be emphasized that the CALIOP measurement is probing the atmosphere very efficiently in the along-track direction since it is a nadir pointing instrument. Here, cloud dimensions down to the original FOV resolution (333 m) will be detected. However, it should be made clear that the across-track extension of the observation is still limited to 333 m. Thus, to compare CALIOP-derived results with the results of 4 km GAC AVHRR pixel data is not entirely consistent (i.e., CALIOP is only capable of covering the GAC pixel properly in one direction and not in the perpendicular direction). However, we believe that this deficiency is of marginal importance. Most cloud systems on the GAC scale will be detected, e.g., it is very unlikely to imagine elongated clouds with size and shapes below 0.3x4 km that might risk remaining undetected within a GAC pixel that coincides with a CALIOP measurement. Most clouds will have aspect ratios for the two horizontal directions that guarantee detection by CALIOP. However, it is also clear that in situations with scattered (sub-pixel) cloudiness within the GAC FOV some optically thick clouds may be detected by AVHRR cloud


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 148

schemes while not being covered at all by CALIOP FOVs. Thus, some small bias between AVHRR and CALIOP observations due to this effect appears unavoidable.

It is important to consider that the CALIOP lidar instrument is much more sensitive to cloud particles than the measurement from a passively imaging instrument. It means that a significant fraction of all CALIOP-detected clouds will not be detected from imagers. This sensitivity difference also propagates into CPH and CTH, which will typically be sensed at a lower cloud layer by passive instruments compared to CALIOP (see e.g., [Hamann et al., 2014](#)). Thus, to get reasonable and justified results one should theoretically consider filtering out the contributions from the very thinnest clouds. We have applied this approach in this validation study, both in the study of cloud amounts (CFC) and cloud top heights (CTO).

The cloud detection efficiency with CALIOP is slightly different day and night because of the additional noise from reflected solar radiation at daytime that can contaminate lidar backscatter measurements. However, Chepfer et al. (2010) reports that this can introduce an artificial difference of not more than 1 % when comparing night time and daytime data.

In conclusion: Despite the fact that the CALIPSO cloud observations most likely are the best available cloud reference data record being released so far, we might still see a negative bias of a few percentage points in cloud cover when using exclusively the 5 km data record. However, in this validation effort we have tried to compensate for this effect by combining the 1 km and 5 km data records following [Karlsson and Johansson \(2013\)](#). Other errors, e.g. due to mis-interpretation of heavy aerosol loads as clouds, are in this respect of minor importance when judging the effect on full global orbits.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 149

A.2 AMSR-E LWP

Passive microwave imagers, such as the Advanced Microwave Scanning Radiometer - EOS (AMSR-E), can be used to retrieve column-integrated liquid water along with water vapour and surface wind speed.

AMSR-E is a dual-polarization conical-scanning passive microwave radiometer with 12 channels ranging from 6.9 to 89 GHz. This instrument was designed to measure cloud properties, sea surface temperature and surface water, ice and snow. Because the microwave (MW) channels usually fully penetrate clouds, they provide a direct measurement of the total liquid (but not solid) cloud condensate amount. For precipitating clouds an estimate of the rain water path has to be made and subtracted from the total liquid water path to retrieve the cloud liquid water path. AMSR-E data are in this document used for Level-2 validation of LWP.

For the validation results presented in this document the Version 06 Level 2B Ocean product suite (see Wentz and Meissner (2000) for more information on the algorithm) was used. The data was collected from the EOS Data Pool at NSIDC (<http://nsidc.org/>). AMSR-E LWP products are derived primarily from the liquid sensitive 37 GHz channel measurement, which has a spatial footprint of 10×14 km. A list of all parameters included in this product can be found at http://nsidc.org/data/docs/daac/ae_ocean_products.gd.html. This page also indicates a root mean square error of 17 g/m^2 as preliminary error estimate for LWP retrievals.


Setup of collocation

Spatiotemporally collocated AMSR-E and AVHRR/AATSR pixels were collected over ocean and during daylight conditions. The LWP of all AVHRR/AATSR pixels that fall into one AMSR-E footprint were averaged, this included clear-sky pixels that assigned to have a LWP of 0 g/m^2 . The latter is done to produce a clear-sky LWP, thus aligning the Cloud_cci data to what is measured by AMSR-E.

A.3 DARDAR IWP

DARDAR is a combined product based on measurement by CALIOP lidar and CPR onboard CloudSat. CPR is a nadir-looking cloud profiling radar sensing the atmosphere from above at 94 GHz. The instrument's sensitivity is defined by a minimum detectable reflectivity factor of -30 dBZ and calibration accuracy of 1.5 dB. The minimum detectable reflectivity factor requirement was reduced to -26 dBZ when the mission was changed to put CloudSat into a higher orbit for formation flying in A-train. The DARDAR dataset (Delanoë and Hogan, 2008) provides the result from a synergistic retrieval method combining the measurements from the CALIOP lidar, the CLOUDSAT radar and the MODIS imager, all three elements of the A-Train satellite constellation. By combining these different measurements, consistent profiles of microphysical properties are retrieved based on the specific particle size (instrumental) wavelength sensitivities. The lidar signals for instance are sensitive to the particle surfaces in the line of sight ($\sim \text{radius}^2$), which is dominated by the smaller particles in a particle size distribution (PSD) whereas the radar signals are sensitive to the square of the particle volume which is dominated by the larger particles in the PSD. When both signals are available the combined PSD sensitivities provide the best guess of extinction, effective particle radius and IWC. When only one of the signals is available, i.e. when the lidar is fully attenuated or when the particles are too small to be detected by radar, the DARDAR retrievals are based on the single instrument parameterizations. The optimal estimation framework used for this retrieval ensures a smooth transition from these different regimes. The DARDAR product has the vertical resolution of CALIOP (30/60 m) and a horizontal resolution given by the radar footprint (700m). This is in contrast to the comparison to the CALIOP data (Section 5.2) which has been averaged to 5 km wide layers before being compared to the CLAS-2 data records.

The DARDAR data for the current evaluation has been downloaded from the ICARE site: http://www.icare.univ-lille1.fr/projects/dardar/overview_dardar_cloud. DARDAR data is used for level-2 evaluation of IWP. Important to note, DARDAR is much more sensitive to thin ice cloud than passive imagers. In the lower part of the atmosphere, the reflectance of the surface affects the backscattered radar signal, so clouds may not be properly detected below 1 km distance to the surface.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 150

A.4 SYNOP: manual cloud observations from surface stations

Observations of total cloud cover made at meteorological surface stations (i.e. synoptic observations - hereafter called SYNOP) constitute one of the data records used to evaluate the cloud fractional coverage estimates. The SYNOP data used is from the local DWD archive of collected global SYNOP reports following the guidance of the *Guide to Meteorological Instruments and Methods of Observations* (WMO, 2008)

At manned stations the total cloud cover is visually estimated by human observers, at automated stations in contrast ceilometers are used for that purpose. For data quality reasons, only those SYNOP reports provided by manned airport stations were taken into account (~1800 stations globally).

SYNOP total cloud cover observations are used for the evaluation of level-3 cloud cover estimates.


Manual cloud observations are affected by many sources of error. We list some of the most important in the following:

- The observation is subjective in nature, i.e., despite clear instructions on how to make an observation, differences will appear because of different interpretations from person to person. This introduces a random noise in global cloud amount observations but may also lead to geographical biases (reflecting some systematic behaviour related to the way people have been educated/trained).
- The human eye has a detection limit for when a cloud can be clearly discernible against a cloud-free sky. This limit is somewhere in the cloud optical thickness range of 0.5-1.0 (with some dependence on solar zenith angle and on which viewing angles clouds are observed and the degree of aerosol load or haze in the troposphere). Thus, many satellite sensors have a higher sensitivity to e.g. cirrus detection than SYNOP observations.
- At night, the random error in the observations increases, naturally since the observer does not have a clear sky background against which a cloud can be observed (i.e., clouds are as dark as the cloud-free sky). However, accuracies improve in the presence of moonlight. Nevertheless, the overall effect is normally a negative bias (underestimated cloud amounts) since the observer is tempted to report cloud free conditions as soon as stars become visible, thus neglecting that large fractions of thin cirrus and other cloud types may still be present.
- A well-known deficiency of SYNOP observations is the scenery effect, i.e. overestimation of convective cloud towers at a slanted view (Karlsson, 2003). This effect is thus most pronounced in the summer season and for low to moderate cloud amounts when the overestimation easily can reach values of 20-30 % (1-2 octas).
- It is important to consider that most SYNOP stations are located at land stations and with higher density in developed countries. Thus, global averages tend to be biased towards land conditions in densely populated countries.

Since no rigorous study has been able to cover all those aspects in a quantitative manner (mainly because of lack of an absolute truth as reference) we can only make a very general qualitative statement about the overall quality. We would suggest that the accuracy of SYNOP observations vary between approximately +10 % (some overestimation) at daytime conditions changing to -10 % or worse (some underestimation) at night time. However, the variability (precision) probably reaches higher absolute values and it is largest during night conditions. This may lead to a strong seasonal variation with the worst accuracy and precision features during the winter season (at least at middle and high latitudes including the polar regions).

It is worth noting that the increasing trend to replace manual cloud observations with automatic observations from ceilometers will change the accuracy and precision of cloud observations in several ways.

Despite their subjective character and varying quality, SYNOP observations still provide a useful reference data set suitable for monitoring and validating space-based estimations of cloud coverage, especially due to their long-term availability.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 151


A.5 Baseline Surface Radiation Network

The international Baseline Surface Radiation Network (BSRN) provides ground-level radiation measurements with high temporal resolution for the validation of satellite data, the validation and improvement of radiative transfer calculations in climate models as well as the detection and monitoring of long-term changes in surface radiation fluxes (Driemel et al., 2018). The BSRN started working in 1992 with 9 stations and currently comprises a network of 59 stations (status 2018-01-01), which are distributed over all continents and oceanic environments. The stations deliver data to the archive called the World Radiation Monitoring Center (WRMC), which is hosted at the Alfred Wegener Institute (AWI) in Bremerhaven, Germany. The data archive offers about 10300 months of data from the years 1992 to 2017 and is freely available at <https://doi.org/10.1594/PANGAEA.880000>.

BSRN stations measure direct, diffuse and global downwelling shortwave and longwave fluxes in 1 min temporal resolution. The manned stations are located at positions, which are representative of a relatively large surrounding area for the use in satellite and climate model validation. The quality controlled datasets are available for the years 1992 to 2017 in ASCII file format. In this work we used specially calculated monthly means of daily mean products:

- Global shortwave radiation (SWD)
- Long-wave downward radiation (LWD)
- Shortwave upward radiation (SWU)
- Long-wave upward radiation (LWU)



	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 152

A.6 CERES EBAF Ed4.0 TOA & Surface fluxes

The Cloud and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) top-of-atmosphere (TOA), Edition 4.0 (Ed.4.0), data product provides long-term shortwave (SW) and longwave (LW) TOA fluxes for all- and clear-sky conditions (Loeb et al., 2018). The CERES instruments fly on the *Terra* and *Aqua* satellites and cover a period from March 2000 to June 2002 for *Terra* only, and cover combined *Terra* and *Aqua* observations from July 2002 to January 2017. Each instrument measures filtered radiances in the shortwave wavelengths between 0.3 and 5 μm , total wavelengths between 0.3 and 200 μm and window wavelengths between 8 and 12 μm . The filtered radiances are converted to unfiltered SW, LW and window radiances following Loeb et al. (2001), which in turn are converted to instantaneous TOA radiative fluxes using empirical angular distribution models (Su et al., 2015). The CERES instruments provide global coverage daily, and monthly mean regional fluxes are based upon daily samples over the entire globe.

However, the standard CERES data products are produced using coincident imager data from the Moderate Resolution Imaging Spectrometer (MODIS) and the Visible Infrared Imaging Radiometer Suite (VIIRS). To provide a diurnally complete representation of Earth's radiation budget, all available geostationary imager data between 60°S and 60°N are also used to account for cloud radiation changes between CERES observation times. A specialty about the CERES EBAF product is the adjustment of occurring radiation imbalances within the CERES-based Earth Radiation Budget (ERB). Therefore an objective algorithm is applied to adjust SW and LW TOA fluxes within their ranges of uncertainty to remove the inconsistency between average global net TOA flux and heat storage in the earth-atmosphere system (Loeb et al. 2009).

For intercomparison we used the CERES EBAF TOA longwave and shortwave monthly mean fluxes for all-sky conditions. In CERES EBAF Ed4.0, the monthly mean (Level 3) fluxes are determined by spatially averaging the instantaneous TOA flux values on a 1°x1° grid, temporally interpolating between observed values at 1-h increments for each GMT hour of every month and then averaging all hour boxes in month (Doelling et al. 2013). The used products are summarized as follows:


- Top of the Atmosphere Longwave Flux, Monthly Means, All-Sky conditions
- Top of the Atmosphere Shortwave Flux, Monthly Means, All-Sky conditions

In addition to the TOA fluxes, we also used the CERES EBAF Ed4.0 Surface Fluxes (Kato et al., 2018). Surface fluxes in EBAF-Surface are derived from the CERES data products: (i) CERES SYN1deg-Month Ed4.0 provides computed surface fluxes to be adjusted, (ii) CERES EBAF-TOA Ed4.0 provides the CERES-derived TOA flux constraints by observations and (iii) SYN1deg-Hour provides weights to compute monthly mean computed clear-sky TOA fluxes. The monthly mean all-sky surface shortwave and longwave fluxes for 1°x1° grids are computed by averaging hourly CERES SYN1deg Ed4.0 all-sky fluxes (Wielicki et al., 1996).

In order to minimize the error in surface fluxes due to uncertainties in the input data sources, the EBAF-Surface data product introduces several additional constraints based upon information from other independent data sources, such as CERES TOA fluxes, Atmospheric Infrared Sounder (AIRS) derived temperature and humidity profiles as well as CALIPSO/CloudSat-derived vertical profiles of clouds. With the help of Goddard Earth Observing System (GEOS) atmospheric profiles and CERES-derived surface albedo radiative transfer model calculations can be applied. From CERES EBAF Ed4.0 Surface fluxes we used the following products:

- Surface Shortwave Flux Down, Monthly Mean, All-Sky conditions
- Surface Longwave Flux Down, Monthly Mean, All-Sky conditions

The CERES EBAF TOA Ed4.0 dataset for TERRA and AQUA is freely available at https://doi.org/10.5067/Terra+Aqua/CERES/EBAF-TOA_L3B004.0 as well as the CERES EBAF SURFACE Edition 4.0 dataset at https://doi.org/10.5067/Terra+Aqua/CERES/EBAF-Surface_L3B004.0.

	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 153

A.7 ELDONET

The European light dosimeter network (ELDONET) contains in total more than 40 stations spread over 24 countries and 5 continents. The ELDONET data provides ground-based reference data for photosynthetic active radiation among others. Figure A.7-1 provides a map of the station locations. More information on this network and the data can be found in Häder et al. (1999), Maragoni et al. (2000), Lebert et al. (2002) and Häder et al. (2007). A subset of the data was acquired from Häder (personal communication). After applying rigorous quality control, the data of only 7 stations were used. The temporal coverage and density of the data remain very divers among these stations.

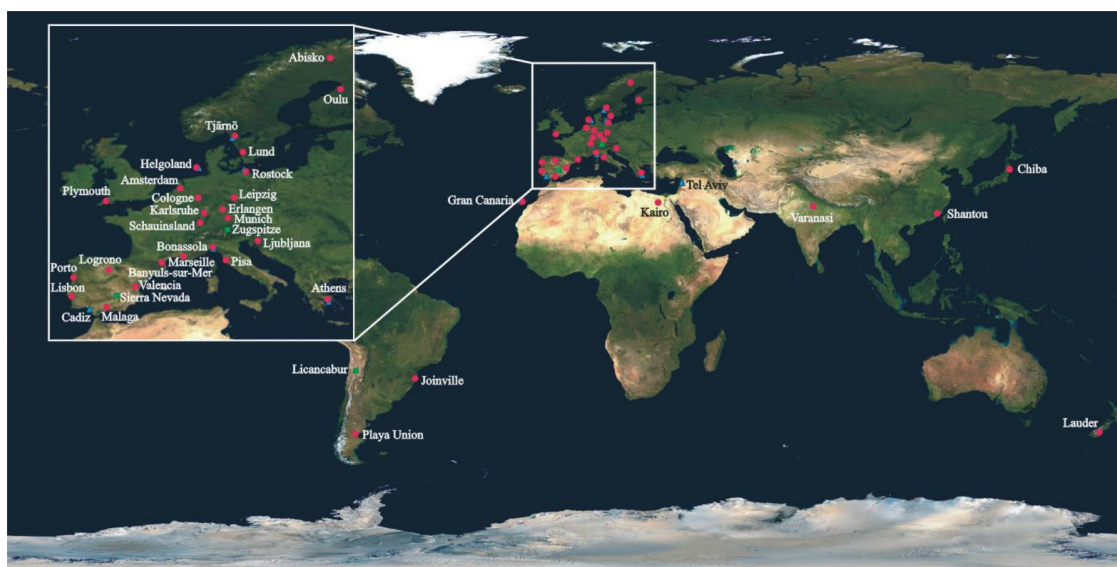



Figure A.7-1 Map of the terrestrial ELDONET instruments in the world and in Europe (inset). Figure and Caption taken from Häder et al. (2007).

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 154

A.8 The CLARA-A2 dataset


The CM SAF cloud, albedo and radiation (CLARA-A2, [Karlsson et al., 2017](#)) dataset provides a long term data record from January 1982 to December 2015 based on homogenized AVHRR (Advanced Very High Resolution Radiometer) measurements. The AVHRR Instrument is stationed on board the polar orbiting NOAA satellites and the EUMETSAT METOP satellites measuring in five spectral bands in the beginning of the dataset (1982 AVHRR/2) and was extended to a sixth channel in 1998 (AVHRR/3), although only accessible if switched with the previous third channel at 3.7 micron. Original visible radiances were inter-calibrated and homogenized, using MODIS data as a reference, before applying the various parameter retrievals. The inter-calibration is based on a method introduced by [Heidinger et al. \(2010\)](#) and updated to MODIS Collection 6. The horizontal resolution of the dataset is close to 1 km at nadir but only reduced resolutions of approximately 4 km are permanently archived and available with global coverage. The CLARA-A2 dataset includes measurements from NOAA-7 up to 19 and the METOP-A and METOP-B missions.

The cloud products, including macro physical and microphysical cloud properties, are provided on a L2 and L3 (daily and monthly mean) basis and are summarized as follows:

- Fractional Cloud Cover - CFC
- Cloud Top Pressure - CTP
- Cloud Top Temperature - CTT
- Cloud Top Height - CTH
- Cloud Optical Thickness - COT
- Cloud Phase - CPH
- Liquid Water Path - LWP
- Ice Water Path - IWP
- Cloud effective radius - REF
- Joint Cloud property Histogram - JCH

The surface radiation data records derived from AVHRR contain information on the shortwave and longwave radiation. The shortwave incoming surface radiation data (SIS) is based on the retrieval of the surface irradiance using information from the Nowcasting SAF cloud detection algorithm PPSv2014 and the satellite-derived radiances in the visible and near-infrared satellite channels. The longwave surface radiation data records for surface outgoing longwave radiation (SOL) and the surface downwelling longwave radiation rely on information obtained from the ERA-Interim reanalysis ([Berrisford et al., 2011](#)) and the monthly averaged cloud fraction of CLARA-A2. All products are globally available as monthly averages (SIS is also available as daily averages) between 1982 and 2015 on a 0.25°-regular longitude-latitude grid. The radiation products are provided on a L2 and L3 basis (see https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V002) and are summarized as follows:

- Surface Incoming Shortwave Radiation - SIS
- Surface Outgoing Longwave Radiation - SOL
- Surface Downwelling Longwave Radiation - SDL

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 155


A.9 PATMOS-x

The most appropriate satellite-derived climatology to compare AVHRR-based Cloud_cci datasets with is (besides CLARA-A2) the PATMOS-x data record. The acronym stands for “AVHRR Pathfinder Atmospheres - Extended” and the corresponding cloud products have been derived using the CLAVR-x method (Clouds from AVHRR - Extended, see [Heidinger et al, 2005](#), [Pavolonis et al., 2005](#), [Thomas et al., 2004](#) and [Heidinger and Pavolonis, 2009](#)). AVHRR radiances in all available spectral channels have been used to derive global cloud and radiation products over the entire lifetime of the AVHRR sensor. Some basic information about the used methodology for the derivation of various parameters is given in Table A.. To notice is that the cloud screening methodology of CLAVR-x has undergone a substantial revision lately compared to the method described by the cited references. The previous multispectral threshold approach has been replaced by a probabilistic methodology (naïve Bayesian classifier - see [Heidinger et al., 2012](#)). We have compared Cloud_cci results against the results produced by this new method. This means we have compared to PATMOS-x version v05r03. The most up-to-date publication describing the PATMOS-x data record is provided by [Heidinger et al. \(2014\)](#).

Table A.1 *Some basic characteristics of the PATMOS-x retrieval methods*

Product	Methodology
Cloud amount	Computed from results of a statistical naïve Bayesian cloud mask trained from CALIPSO-CALIOP cloud information
Cloud top level	Optimum Estimation (OE) retrieval
Cloud phase	Multi-channel test scheme
Cloud optical thickness	OE retrieval (with look-up tables as CM SAF but with different radiative transfer models and ice particle definitions)
Cloud effective radius	OE retrieval (with look-up tables as CM SAF but with different radiative transfer models and ice particle definitions)
Cloud liquid water path	Calculated from optical thickness and effective radius (Stephens' parameterization – same as CM SAF)
Cloud ice water path	Calculated from optical thickness and effective radius (Stephens' parameterisation – same as CM SAF)

The PATMOS-x data record is prepared exclusively as so-called level-2b products. This means that, for each satellite, data from all orbits during one day have been sub-sampled to produce only two global products per day valid for the nominal local solar time for both the descending (southbound) and ascending (northbound) observation nodes.


	Doc:	Cloud_cci_D4.1_PVIR_v6.1.docx			
	Date:	03 February 2020			
	Issue:	6	Revision:	1	Page 156

A.10 MODIS Collection 6.1

MODIS (or Moderate Resolution Imaging Spectroradiometer) is an advanced imaging instrument onboard the Terra (EOS AM) and the Aqua (EOS PM) polar satellites (see <https://modis-atmosphere.gsfc.nasa.gov/products/monthly>). Both Terra and Aqua orbits around the Earth are sun synchronous. Terra passes from north to south across the equator in the morning (local solar time 10:30), while Aqua passes south to north over the equator in the afternoon (local solar time 13:30). MODIS Terra and MODIS Aqua are viewing the entire Earth's surface every 1 to 2 days, acquiring data in 36 spectral bands or groups of wavelengths.

Since the Terra and Aqua satellites passes in very similar orbits (at least the afternoon orbit of Aqua) as the NOAA and Metop-A satellites and since MODIS observes with as much as 36 spectral channels (including all the AVHRR-like channels), corresponding cloud products from MODIS should serve as a top quality reference for corresponding cloud products retrieved from AVHRR data. MODIS uncertainties are indeed expected to be somewhat smaller than what can be obtained with AVHRR retrievals. For example: multiple CO₂ channels allow a more accurate cloud-top height determination, additional shortwave channels allow better discrimination of (thin) cirrus and a more reliable retrieval of cloud optical properties over very bright surfaces. Otherwise, uncertainties should lie in the same ballpark as for Cloud_cci, CLARA-A2 and PATMOS-x. The main limitation of MODIS is the relatively short duration of the observation period, starting in 2000. We have used the level-3 MODIS gridded atmosphere monthly global products - MOD08_M3 (Terra) and MYD08_M3 (Aqua). They contain monthly 1° × 1° degree grid average values of atmospheric parameters related to atmospheric aerosol particle properties, total ozone burden, atmospheric water vapour, cloud optical and physical properties, and atmospheric stability indices. Statistics are sorted into 1° × 1° degree cells on an equal-angle grid that spans a (calendar) monthly interval and then summarized over the globe.


For this particular study we have used the data from MODIS Collection 6.1 Terra (dx.doi.org/10.5067/MODIS/MOD08_M3.061) as well as the data from MODIS Collection 6.1 Aqua (dx.doi.org/10.5067/MODIS/MYD08_M3.061) (Platnick et al., 2017a; Baum et al., 2012; Platnick et al., 2017b). Validation results for MODIS Collection 6.1 cloud properties are not yet available, but for MODIS Collection 6.0 cloud properties validation results can for example be found in Baum et al., (2012), Marchant et al. (2016) and Wang et al., (2016).

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 157

A.11 ISCCP

The International Satellite Cloud Climatology Project (ISCCP) provides cloud properties over a period of more than 35 years (Rossow and Schiffer, 1991; Rossow et al., 1996; Rossow and Schiffer, 1999). This project was established in 1982 as part of WCRP to collect weather satellite radiance measurements (from geostationary and polar orbiting satellites) and to analyze them to infer the global distribution of clouds, their properties, and their diurnal, seasonal and inter-annual variations. The resulting data records and analysis products are being used to study the role of clouds in climate, both their effects on radiative energy exchanges and their role in the global water cycle. This project and its results are considered to be the state of the art today on what can be derived from routine weather satellite data. ISCCP is the only other existing TCDR for cloud physical property products (here we mean products CPH, LWP and IWP). However, it has the disadvantage that it is based on different satellite types - polar and geostationary - of which most of the latter do not contain the necessary narrow-band channels for accurate retrieval of LWP and IWP.

The production of ISCCP has recently been transferred to the National Centers for Environmental Information (NCEI) and a new high-resolution version of the data record (to be denoted ISCCP-H) is under production (see <https://www.ncdc.noaa.gov/isccp>). Unfortunately, the new data record is still not released which means that we have compared to the previous ISCCP-D2 version covering the period 1983-2008 as prepared in the Global Energy and Water Cycle Experiment (GEWEX) database (Stubenrauch et al, 2013). I.e. we have used the 0300AMP data products which contain daytime and night-time data valid at 03:00 AM and 03:00PM local time, which is close to the observation times of the prime AVHRR afternoon satellites and Aqua.


	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 158

A.12 NOAA (HIRS) Climate Data Record OLR

The High Resolution Infrared Radiation Sounder (HIRS) Outgoing Longwave Radiation (OLR) Climate Data Record (CDR) at the top of the atmosphere is obtained from algorithm estimates that use narrowband radiance observations for OLR under all-sky conditions from the HIRS instruments onboard NOAA TIROS-N series and EUMETSAT MetOp-A (and B) operational polar-orbiting satellites. Thereby, a continuous monthly mean OLR time series in $2.5^{\circ} \times 2.5^{\circ}$ equal-angle gridded global maps spanning from January 1979 to the present is generated (Hai-Tien Lee & NOAA CDR Program, 2018).

HIRS consists of nineteen infrared channels and one visible channel. The HIRS OLR CDR production starts with the derivation of the radiance from the HIRS Level-1B data by the radiance calibration module. The radiance calibration is performed consistently throughout the whole time series. Within the OLR retrieval a multi-spectral OLR estimation technique is applied. The OLR is retrieved at each HIRS field of view with the given radiances and local zenith angles. For each satellite, one month worth of field of view retrievals are compiled and gridded into equal-angle maps, consisting of observations from the ascending and descending orbits, respectively.

The OLR monthly mean is derived from the gridded OLR monthly orbital maps of all the contributing satellites. The monthly orbital maps can be considered as the monthly mean OLR at particular local times. Then, the OLR inter-satellite bias adjustments are first applied to the OLR monthly orbital maps before the temporal integration. The temporal integral is then independently performed for each grid. The diurnal variation is constrained by the observed OLR retrievals obtained at different observing local times. The OLR monthly mean values is then determined as the integral of the constrained envelop curve over the 24 hours period (Hai-Tien Lee & NOAA CDR Program, 2018). In short, by an application of a combination of statistical techniques, includes OLR, instrument ambient temperature prediction coefficients, and inter-satellite bias corrections produce the final records. All records are freely available under <https://doi.org/10.7289/V5W37TKD>.

	Doc:		Cloud_cci_D4.1_PVIR_v6.1.docx		
	Date:		03 February 2020		
	Issue:	6	Revision:	1	Page 159

A.13 ERA-Interim reanalysis

The ERA-Interim dataset is a reanalysis dataset provided by the ECMWF. The offered fluxes are either downward or net. For the validation we used the surface downward fluxes and the top of the atmosphere upward fluxes, calculated with the help of the net fluxes. In this work we used the following products:

- Top net solar radiation – TSR
- Top net thermal radiation – TTR
- TOA incident solar radiation – TISR
- Surface solar radiation downward – SSRD
- Surface thermal radiation downward – STRD

All the data starts in 1979 and provides until present meteorological parameters in near real-time. The ERA-Interim atmospheric model has a spatial resolution of 0.75 degree and it is based on 60 atmospheric layers. The reanalysis dataset is produced with a sequential data assimilation scheme, advancing forward in time using 12-hourly analysis cycles. In each cycle, available observations are combined with prior information from a forecast model to estimate the evolving state of the global atmosphere and its underlying surface (Berrisford et al., 2011; Dee et al., 2011). For this study we obtained monthly mean values of the different products from <https://www.ecmwf.int/node/8173>.