ESA Cloud_cci

Climate Assessment Report

(Applicable to Cloud_cci version 2.0 products)

Issue 3 Revision 1

18 September 2017

Deliverable No.:	D-5.1
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

Document Change Record

Document, Version	Date	Changes	Originator
Version 3.0, submitted version	13/05/17	Final editing	R. Hollmann
Version 3.1_trackchange resubmitted version	28/08/17	Revision after ESA review with partly substantial changes, which include e.g. - Updating Executive Summary - New Section 6, giving results of Option 1. - New Annex A, comparing SIMFERA and full Cloud_cci simulator based on results achived when applied to the same data.	J. Bojanowski, E. van Meijgaard, J.F. Meirink, M. Keller, S. Eliasson, U. Willén, C. Stubenrauch, R. Hollmann, M. Stengel

Purpose

This Climate Assessment Report (CAR) complements the detailed Level-2/Level-3C validation and comparisons described in the ESA Cloud_cci Product Validation and Intercomparison Report D-4.1 by:
1) Studying the differences between Cloud_cci datasets and datasets of the GEWEX Cloud Assessment (CA) database,
2) Comparing the Cloud_cci datasets to modelled clouds of reanalyses as well as of regional and global climate models.

LIST OF FIGURES

LIST OF TABLES

Executive Summary

This Climate Assessment Report (CAR) complements the more detailed Level-2/Level-3C validation and comparisons described in the ESA Cloud_cci Product Validation and Intercomparison Report D-4.1 by (1) studying the differences between Cloud_cci datasets and datasets of the GEWEX Cloud Assessment (CA) database and (2) comparing the Cloud_cci datasets to clouds from meteorological reanalyses as well as from regional and global climate model simulations. The document describes analyses and results using the gridded and time/space averaged and sampled (histograms) Level-3C products of ESA Cloud_cci. In addition, (3) summarizes a first steps in using the provided Cloud_cci uncertainties in model evaluation.

(1) *ESA* Cloud_cci cloud retrievals are based on an Optimal Estimation (OE) technique and use similar instruments (AVHRR, MODIS and AATSR), with spectral information ranging from visible to infrared during daytime and from near-infrared to infrared during night-time.
As the Cloud_cci *datasets* use different spectral information during daytime and night-time, they were separately evaluated using appropriate data of the GEWEX Cloud Assessment data base as well as data from IR sounders (AIRS, IASI). The latter use the same IR spectral information during daytime and night-time and also provide, due to their good spectral resolution, a very reliable cirrus identification (down to an IR optical depth of 0.1). Global total cloud amount compares well to the reference datasets from other passive remote sensing cloud datasets (0.68±0.03), with a similar performance of Cloud_cci AVHRR and MODIS. Cloud amount of AATSR is slightly lower (0.66), and the one of MERIS-AATSR is underestimated (0.60). The data show a good coherence of latitudinal variation and the seasonal cycle, except for MERIS-AATSR.
One finding of this assessment is that the performance of the Cloud_cci datasets in identification of high-level clouds is worse during daytime than during night-time. This might be explained by the fact that VIS information in combination with spectral IR information in passive imager methods may lead to a misidentification in the case of thin cirrus overlying low-level clouds (e.g.(Stubenrauch et al., 2013)). During daytime the relative amount of high-level clouds (globally averaged to 28%), in particular over land, is underestimated. This is lower than the 30% quoted by ISCCP and 40% from AIRS/IASI. The latter identify cirrus under all conditions.
During night-time, when only spectral IR information is available, Cloud_cci relative high-level cloud amounts increase to 35-39%, in agreement with AIRS/IASI. However, total cloud amount seems to be slightly underestimated, mostly by missing low-level clouds.
The vertical distribution of cloud pressure which issupposed to be bimodal in the tropics with peaks around 250 and 950 hPa, reveal that the height of low-level clouds from the Cloud_CCI data sets seems to be slightly overestimated. The peak for high-level clouds, which is located at the right height, is smaller than for other other datasets during daytime / night-time. To overcome this situation and to improve the performance for high-level clouds, one path might be to adapt the OE method using daytime spectral information to the one using night-time spectral information (which excludes VIS information). By applying both methods during daytime, keeping the VIS information for the cloud mask, one might even get information of multi-layer cloud situations from a comparison of VIS optical depth to IR emissivity.

(2) Comparisons with cloud properties simulated by meteorological reanalyses, regional and global climate models show that the current Cloud_cci products are very useful. Based on the comparisons with the global climate model, polar regions as well as high altitude snow covered regions became visible where passive instrument satellites have in general problems detecting clouds. The assessment shows that an evaluation of high-level clouds from climate models with the Cloud_cci datasets should be currently done with care, for example by excluding clouds with optical depth less than 1 – 2 or by considering night-time observation.
The use and application of the developed and providedCloud-CCI simulator is essential in setting up a direct comparison between model predicted cloud parameters and Cloud_cci based retrievals. An application of the Cloud-CCI simulator to 25-km resolved multi-level cloud fields obtained with a one month integration in hindcast mode of the regional climate model RACMO2 shows that the

simulated fields are consistently (much) closer to the Cloud_cci AVHRR cloud property retrievals than the original model fields, the exception being the ice water path field. First results indicates that the model produces higher and colder cloud tops and a higher ice water path than the retrieval indicates. Nevertheless, the difference between the model and observations remains large. It remains very difficult to attribute the differences to one or the other, thus more in depth studies are needed to understand the origins of these differences.
An application of a simplified cloud simulator to COSMO simulations with 12 and 2 km resolution during a period with strong convective activity shows with the help of COT/CTP histograms a significant reduction in the positive high-level cloud bias from 12 to 2km compared to Cloud_cci MODIS data.

(3) Cloud_cci provides in addition to their gridded and averaged data sets information about associated uncertainty. As Cloud_cci AVHRR total cloud cover compares well with other existing long term AVHRR cloud datasets, a major advantage of the Cloud_cci data are the included pixel based uncertainties. They show the user which areas should be carefully treated. To optimize the use of the Cloud_cci uncertainties it is recommended to provide more guidance from the data providers of how average uncertainties should be calculated as well as to have a closer insight in the derivation of L3 uncertainties from L2 data beyond the existing ATBDs.

To improve the usability of Cloud_cci CFC in climate studies, a simple statistical method is proposed for correcting CFC by debiasing the AVHRR-PM CFC data using synoptic observations. The method is based on geostatistical interpolation of these observations using satellite data as an explanatory variable. The method relies on the strong assumption that synoptic observations are accurate and homogenous. The corrected (debiased) dataset significantly outperforms the original one in terms of accuracy and precision, as well as reveals decreased performance differences among NOAA satellites. Therefore, debiasing can implicitly remove the inhomogeneity in CFC time series due to changing overpass times and unresolved diurnal cycle. The correction decreased magnitude of trends but keeps their signs unchanged (positive trends are less positive, negative trends are less negative).

1. Introduction

1.1 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.

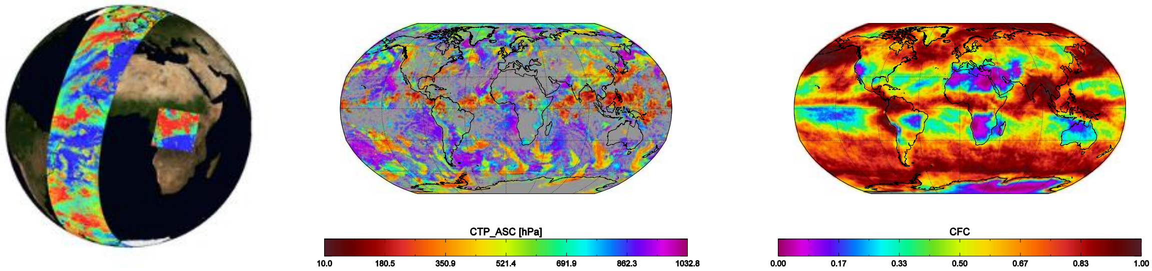


Figure 1-1 *Examples of Cloud_cci cloud products. Left: Pixel-based (Level 2), middle: daily composite on a global grid (Level 3U), right: monthly averaged on a global grid (Level 3C)*

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., 2017; McGarragh et al., 2017) 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.

All parts of the datasets generation effort were properly documented with the major components being the Algorithm Theoretical Baseline Documents (ATBD; ATBDv5, ATBD-FAME-Cv5, ATBD-CC4CLv5), the Product Validation and Intercomparisons Report (PVIR; PVIRv4.1) and the Product User Guide (PUGv3.1).

Furthermore, to facilitate the utilization for evaluation of regional and global atmospheric models, the development of a satellite simulator package for Cloud_cci datasets were fostered, which is planned to be part of one of the upcoming releases of the CFMIP Observation Simulator Package (COSP, Bodas-Salcedo et al. 2011).

1.2 The Cloud_cci datasets

In Cloud_cci two families of global cloud property datasets have been generated. The first family comprises datasets for individual sensor groups such as AVHRR, MODIS, ATSR2/AATSR, for which the AVHRR-heritage channels (0.6, 0.8, 1.6/3.7, 10.8, 12.0 μm) were utilized to retrieve cloud properties using the CC4CL algorithm. The second family comprises a dataset of cloud properties retrieved from simultaneous usage of AATSR and MERIS sensors (both mounted on ENVISAT) by applying the FAME-C algorithm. Since MODIS and AVHRR sensors are separated into morning and afternoon orbits, 6 distinct Cloud_cci datasets exist, which can be seen in Figure 1-2.

Table 1-1 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 1-2), 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 1-3 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 (1x1km^2 for ATSR2, AATSR, MERIS, MODIS; 5x1km^2 for AVHRR). This, together with the sensor swath width, lead to very different observation frequency and spatial coverage. While MODIS and AVHRR 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.

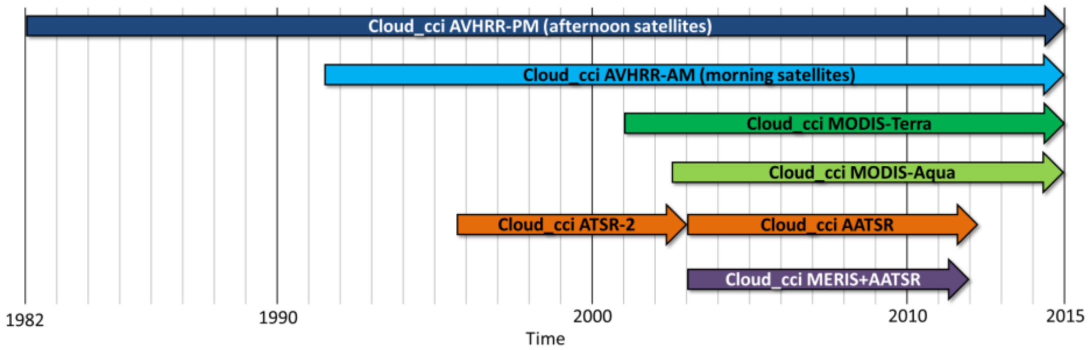


Figure 1-2 *Overview of Cloud_cci datasets and the time periods they cover.*

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). The data are 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° 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 (except CPH) 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 1.4 for more information.

In addition to the passive imager based datasets mentioned so far, in Cloud_cci an IASI-based demonstrator dataset has been created, with more details to be found in Feofilov et al. (2017) and Stubenrauch et al. (2017).

Key strengths of Cloud_cci datasets:

- The Cloud_cci datasets are based on two newly-developed, state-of-the art retrieval systems named CC4CL and FAME-C that use the optimal estimation (OE) technique and are applied to passive imager sensors of current and past European and non-European satellite missions.
- The measurement records of the utilized sensors have been revisited, re-characterized and, in case of AVHRR, re-calibrated.
- Two special features of CC4CL and FAME-C are, among others, their applicability to multiple sensors: ATSR2, AATSR, MODIS, AVHRR (CC4CL) and the simultaneous utilization of AATSR and MERIS measurements (FAME-C, i.e. utilizing the O2-A band of MERIS) 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 1-1 *Cloud_cci 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-PM DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-PM/V002	AVHRR-2/-3	NOAA-7,-9,-11,-14,-16,-18,-19	1982-2014	CC4CL
Cloud_cci AVHRR-AM DOI:10.5676/DWD/ESA_Cloud_cci/AVHRR-AM/V002	AVHRR-2/-3	NOAA-12,-15,-17, Metop-A	1991-2014	CC4CL
Cloud_cci MODIS-Terra DOI:10.5676/DWD/ESA_Cloud_cci/MODIS-Terra/V002	MODIS	Terra	2000-2014	CC4CL
Cloud_cci MODIS-Aqua DOI:10.5676/DWD/ESA_Cloud_cci/MODIS-Aqua/V002	MODIS	Aqua	2002-2014	CC4CL
Cloud_cci ATSR2-AATSR DOI:10.5676/DWD/ESA_Cloud_cci/ATSR2-AATSR/V002	ATSR2, AATSR	ERS2, ENVISAT	1995-2012	CC4CL
Cloud_cci MERIS+AATSR DOI:10.5676/DWD/ESA_Cloud_cci/MERIS+AATSR/V002	MERIS, AATSR	ENVISAT	2003-2011	FAME-C

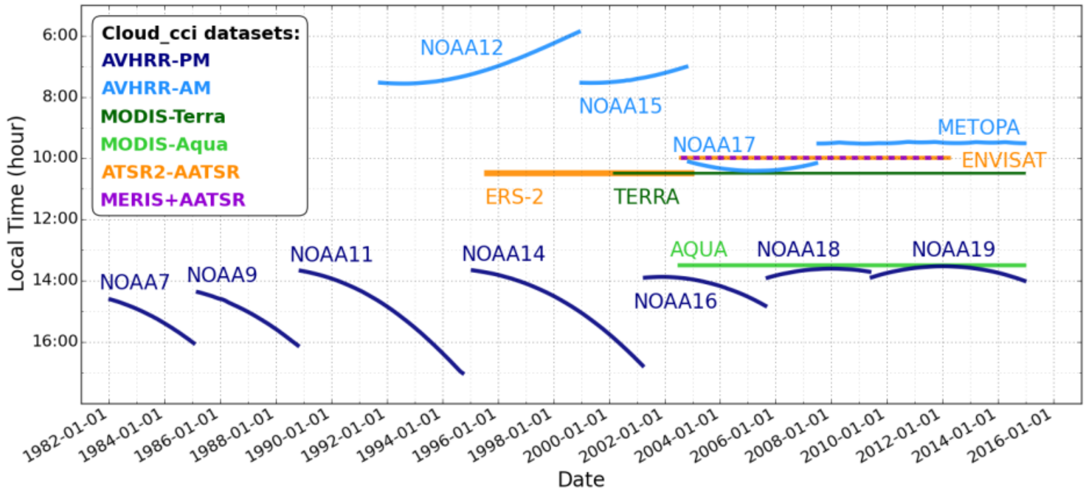


Figure 1-3 *Time periods and local observation times (equator crossing times) of each satellite sensor considered in Cloud_cci. Figure is taken from Stengel et al. (2017).*

1.3 Cloud_cci cloud products

The cloud properties derived on pixel level of each utilized sensor are listed in Table 1-2. It is important to note that the 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. Based on these pixel level retrievals the data is further processed into different processing levels as summarized in Table 1-3. Level-3U denotes a composite on a global Latitude-Longitude grid (of 0.05° resolution) onto which the Level-2 data is sampled (see ATBDv5 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 1-4). Level-3S products are generated merging the Level-3C of all individual sensors. Using Level-3S products requires careful consideration of the partly large and time-varying discrepancies between the used sensors. Please contact the Cloud_cci team for more information (http://www.esa-cloud-cci.org/?q=support).

Table 1-2 *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.*

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:

Variable	Abbrev.	Definition
		liquid or ice; in L2, L3U) and the therefrom derived monthly liquid cloud fraction (L3C, L3S).
Cloud optical thickness	COT	The line integral of the absorption coefficient and the scattering coefficient (at 0.55µm 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 µm) and 2 (0.87 µm), respectively (experimental product)

Table 1-3 *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)	MODIS: 1km AATSR: 1km AVHRR: 5 km MERIS+ AATSR: 1km	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. (MODIS-Europe: 0.02°)	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.
Level-3S (L3S)	Latitude-Longitude grid at 0.5° res.	Cloud properties of Level-2 orbits of all available single sensors combined (averaged / sampled for histograms) on a global space grid. Temporal coverage of this product is 1 month.

Table 1-4 *Cloud_cci product features incl. day and night separation, liquid water and ice as well as histogram representation. Level-3U refers to the non averaged, pixel-based cloud retrievals sampled onto a global Latitude-Longitude (lat/lon) grid. [1]CMA 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

1.4 Uncertainties

The retrieved cloud properties CMA, CTP, CTT, CTH, COT, CER, LWP and IWP (for CC4CL also CLA) are accompanied by pixel-based (Level-2) uncertainties, which are output of the OE technique and represent a rigorous propagation of the uncertainties in the input data, e.g. a-priori information, measurements, radiative transfer. These uncertainty values represent the 68% confidence interval of the true value being within the retrieved value ± uncertainty. These Level-2 uncertainties are also given in Level3U and further propagated into Level-3C. For this a sound mathematical framework has been developed and implemented taking into account the retrieval uncertainties but also the uncertainty correlations. The framework allows an estimation of both the real variability of the observed property and the uncertainty of the calculated mean. Determine and utilizing the uncertainty correlation is a particular key point for an appropriate propagation of Level-2 uncertainties into higher-level products (e.g. Level-3C). Please see the Comprehensive Error Characterization Report (CECRv3) and Stengel et al. (2017) for further details on the uncertainty measures provided.

2. Assessment of Cloud_cci Level-3 data using the GEWEX Cloud Assessment database

2.1 Cloud products participating in the assessment

2.1.1 Data from the GEWEX Cloud Assessment data base

The GEWEX Cloud Assessment cloud products (Stubenrauch et al., 2013) were retrieved from the following satellite instruments and missions: ISCCP (International Satellite Cloud Climatology Project), AVHRR (Advanced Very High Resolution Radiometer) multi-spectral imager aboard NOAA, ATSR (Along-Track Scanning Radiometer) aboard the European Space Agency (ESA) platform ERS-2, HIRS (High resolution Infrared Radiation Sounder) multi-channel radiometer aboard NOAA, AIRS (Atmospheric Infrared Sounder) aboard Aqua, TOVS (TIROS Operational Vertical Sounder) aboard NOAA, MODIS (MODerate resolution Imaging Spectroradiometer) aboard Aqua and Terra, POLDER (POLarization and Directionality of the Earth's Reflectances) multi-angle multi-spectral imager aboard PARASOL, MISR (Multiangle Imaging SpectroRadiometer) aboard Terra, and CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations). A detailed description of the GEWEX Cloud Assessment products can be found in (Stubenrauch et al., 2012). See also the appendix for a short description of the different instrument types. Here a brief overview of the datasets participating in the *ESA Cloud_cci* intercomparison is provided:

- **ISCCP** (Rossow and Schiffer, 1999) emphasizes temporal resolution (eight observations per day) over spectral resolution. For a better comparison with the other datasets in the assessment, the eight-times-daily ISCCP results have been averaged to four specific local observation times: 3:00 AM, 9:00 AM, 3:00 PM and 9:00 PM. Cloud top temperature (CT) is retrieved from the IR radiances. Cloud optical depth (COD) is obtained from the VIS radiances assuming effective particle radii for liquid and ice clouds. ISCCP data are available in the GEWEX Cloud Assessment data base only until 2007.
- **PATMOS-x** (Pathfinder Atmospheres Extended) developed by NOAA utilizes the data from the 5-channel AVHRR imager. Cloud detection is based on Bayesian classifiers derived from CALIPSO (Heidinger et al., 2012), and the retrieval is based on an optimal estimation approach. First cloud pressure (CP) and cloud emissivity (CEM) are retrieved using two IR channels at all times of day. Then COD and CRE are obtained from solar channels during daytime (Walther et al., 2012).
- The **ATSR-GRAPE** cloud products (CP, COD) are retrieved only during day, using an optimal estimation approach on the five available VIS / NIR / IR channels (Sayer et al., 2011).
- The **AIRS-LMD** (Stubenrauch et al., 2010, Guignard et al., 2012) cloud pressure and emissivity (CP, CEM) were retrieved by applying a weighted $\chi 2$ method using CO2 absorbing channels sounding throughout the whole atmosphere. Cloud temperature (CT) and height (CZ) are obtained from CP using retrieved temperature profiles.
- The **MODIS-ST** and **MODIS-CE** cloud products are retrieved the MODIS Science Team and the MODIS CERES science team, respectively. The former uses spectral testing to determine cloud amount (Frey et al., 2008), the "CO2 slicing" to determine CP and CEM (Menzel et al., 2008), and a LUT approach using solar reflectance channels to retrieve COD (Platnick et al., 2003). MODIS-CE determines CT and CEM from IR radiances. For the daytime observations, they retrieve COD using a reflectance-based LUT approach (Minnis et al., 2008, 2011).
- The active lidar measurements of the CALIPSO mission are analyzed by two teams. The CALIPSO Science Team (**CALIPSO-ST**) determines cloud top height from VIS backscatter and identifies cloud ice from depolarization (Winker et al., 2009). Noise is reduced by horizontal averaging. The GCM-Oriented CALIPSO Cloud Products (**CALIPSO-GOCCP**) reduce noise by vertical averaging (Chepfer et al., 2010). Both datasets are available in the GEWEX Cloud Assessment data base for 2007 and 2008.

Monthly L3 data of the GEWEX Cloud Assessment data base have been produced in the following way: in a first step, averages were made per observation over 1° x 1° grid boxes, and in a second step these cloud properties were averaged over the month. The assessed *ESA Cloud_cci* data have been prepared in the same manner. Furthermore, for the assessment of the *ESA Cloud_cci* datasets, ISCCP, AIRS-LMD, MODIS-CE, MODIS-ST, and PATMOSx have been grouped to what is called the "GEWEX Cloud Assessment reference product". These datasets have very similar observation times and have similar sampling and detection sensitivities (with AIRS-LMD the most sensitive to thin cirrus (Stubenrauch et al., 2010, 2013). For each cloud parameter under consideration an average and a root-mean-square (rms) is provided of the variation of this parameter within this reference dataset. Since ISCCP data are only available until 2007, it was decided to assess the *ESA Cloud_cci* data for 2007. This assessment approach was already developed during CCI phase 1, described in the CAR corresponding to D4.2 of Cloud_cci phase 1 (CAR-2013). To show the improvement of the versions for *Cloud_cci* AVHRR, Cloud_cci AATSR and Cloud_cci MERIS-AATSR, it is referred to the CAR report of Cloud_cci phase 1.

2.1.2 Data from IASI

The retrieval scheme for cloud properties from IR sounder observations developed at LMD (Stubenrauch et al., 1999) has been recently updated and adapted so that it can be applied to any IR sounder (CIRS) (Feofilov and Stubenrauch 2017). At LMD it has been applied to AIRS (2003-2015) and IASI (2008-2015) (Stubenrauch et al. 2017). The CIRS method has been evaluated by making use of the A-Train synergy: cloud detection and cloud height from AIRS have been assessed with synchronous data from active lidar and radar from the CALIPSO and CloudSat missions. Cloud detection agrees with CALIPSO-CloudSat about 84%-85% over ocean, 79-82% over land and 70-73% over ice / snow, depending on atmospheric ancillary data. Cloud height corresponds to the height at which the cloud reaches an optical depth of about 0.5 (Stubenrauch et al. 2017). Within a cooperation with the EUMETSAT Climate Monitoring Satellite Application Facility (CM SAF), the retrieval has been adapted to HIRS (Hanschmann et al. 2017). The L3 data are in the same format as other data of the GEWEX Cloud Assessment data base. Within the framework of *ESA Cloud_cci*, one year of IASI cloud data (2009) is used to evaluate the *ESA Cloud_cci* datasets. IASI, developed by CNES in collaboration with EUMETSAT, is a Fourier Transform Spectrometer based on a Michelson interferometer covering the IR spectral domain from 3.62 to 15.5 μm (Hilton et al., 2012). Two instruments were launched so far onboard the European Platforms Metop-A and Metop-B (in 2006 and 2012, respectively), with measurements of at 9:30 / 21:30 LT and 10:30 / 22:30 LT (local equator crossing time). IASI-CIRS cloud data is used for 2009 and AIRS-CIRS cloud data for 2007 (both based on ERA-Interim ancillary data) for an additional assessment of the Cloud_cci datasets.

2.2 Analyses and Results

2.1.2 Total and height-stratified cloud amount

The first characteristic of the dataset is the total cloud amount (CA), often referred to as cloud cover or cloud fraction. CA is defined as the ratio of the number of samples that contain clouds and the total number of measurement samples. When addressing the global CA (or any other global characteristic), the averaging is area-weighted so that a 1° × 1° grid box on the pole has a smaller contribution to the total value than a 1° × 1° grid box on the equator. By using in addition cloud pressure (CP), relative contributions of high-level, mid-level and low-level clouds have been determined, by dividing these cloud amounts (CAH, CAM, CAL) by CA. The sum of the relative contributions, CAHR, CAMR and CALR, should be 1.Pressure limits for high-level/mid-level and mid-level/low-level cloud separation are 440 hPa and 680 hPa, corresponding to altitudes of about 6 km and 3 km, respectively. Relative cloud amount values give an indication of how the detected clouds are vertically distributed in the atmosphere. Compared to the absolute values, they are less influenced by differences in cloud detection sensitivity and should be more useful for comparison with climate models.

Figure 2.1 presents global averages of these four variables, separately for observations mostly during day (left), corresponding to 1:30 PM (10:30 AM for MERIS, AATSR), and mostly during night (right), corresponding to 1:30 AM (10:30 PM for AATSR). In addition, results from IASI-CIRS are shown for 2009 (at 9:30 AM and 9:30 PM, respectively). Total cloud amount from the GEWEX Cloud Assessment data base is about 0.68±0.03, only CALIPSO-ST providing a cloud amount of 0.72, because it includes subvisible cirrus. Cloud_cci MODIS provides a similar global cloud amount and Cloud_cci AVHRR is about 0.02 lower, followed by Cloud_cci AATSR. *Cloud_cci MERIS-AATSR*, only available during daytime, is lower than 0.6 and therefore *underestimates CA*.

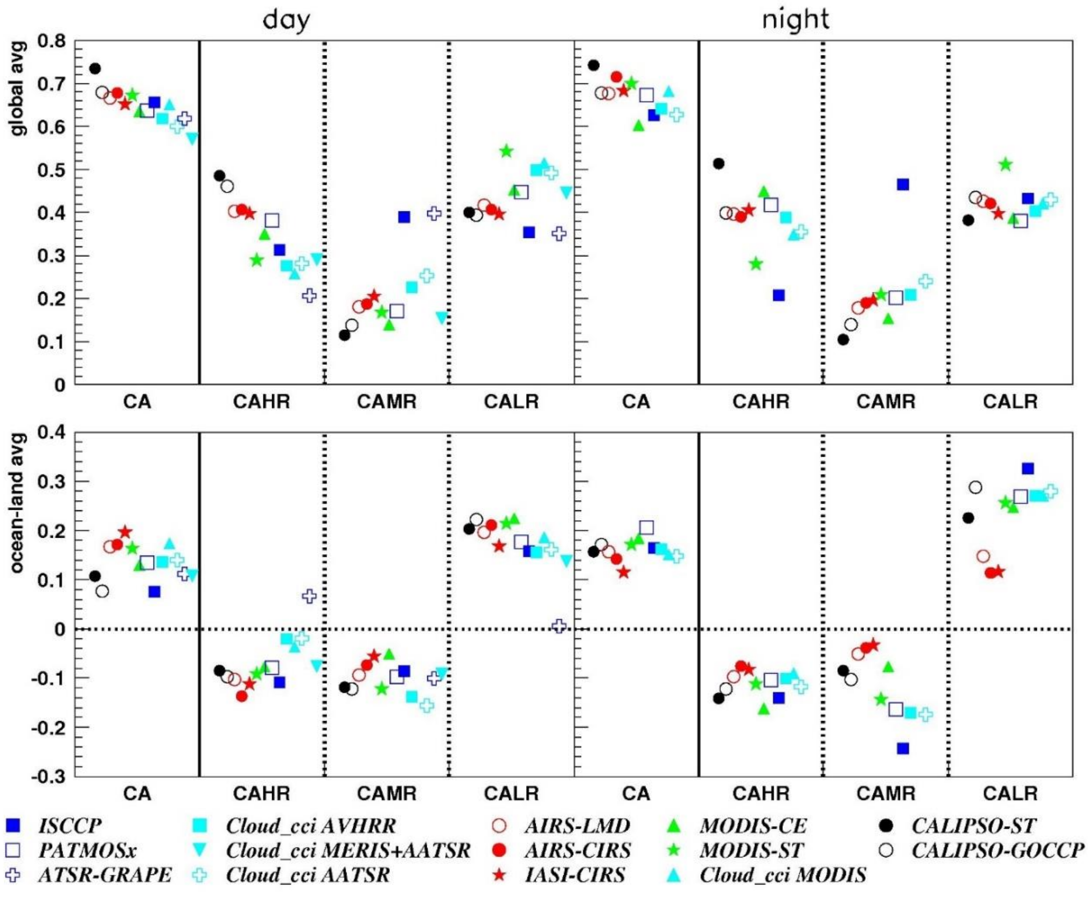


Figure 2-1 *Top: Global averages of total cloud amount (CA), as well as of fraction of high-level, mid-level and low-level cloud amount relative to total cloud amount (CAHR + CAMR + CALR = 1). Comparisons of Cloud_cci Data for 2007 with L3 data from the GEWEX Cloud Assessment data base, separately for observations mostly during day (left), corresponding to 1:30 PM (10:30 AM for MERIS, AATSR), and mostly during night (right), corresponding to 1:30 AM (10:30 PM for AATSR). In addition, results from IASI-CIRS are shown for 2009 (at 9:30 AM and 9:30 PM, respectively). Bottom: Averages of ocean-land differences for the same parameters.*

During day the CAHR of the reference dataset varies from 0.29 for MODIS-ST to 0.4 of AIRS-LMD, AIRS-CIRS and IASI-CIRS, while *all ESA Cloud_cci datasets lay at the lower limit of 0.28-0.29*. During night the Cloud_cci datasets perform better, with CAHR between 0.35 and 0.4, while MODIS-ST and ISCCP perform worse. The mid-level clouds of the reference dataset demonstrate a larger spread with CAMR varying from 0.12 for MODIS-CE to 0.4/0.5 for ISCCP (day/night). The ratio of low-level clouds has a spread between 0.35 (ISCCP) and 0.53 (MODIS-ST). These spreads can be explained by a difference in sensitivity to thin cirrus, linked to instrument and retrieval performance. Cirrus are well-identified by IR sounders (AIRS-LMD, AIRS-CIRS, IASI-CIRS) and active lidar (CALIPSO), also in the case of lower clouds underneath. In the latter case ISCCP confounds them with mid-level clouds (also during night) and MODIS-ST misidentifies these as low-level clouds (Stubenrauch et al., 2012; Stubenrauch et al., 2013, and references within).

In the current version, Cloud_cci AATSR shows a CAHR (0.28), which is comparable with Cloud_cci AVHRR, Cloud_cci MERIS-AATSR and ISCCP. The retrieval methodology of Cloud_cci (optimal estimation) using more spectral information than ISCCP did not improve the value of CAHR during day, while the results for CAHR are better during night-time. This can be explained by the fact that when using spectral information only in the IR, cirrus is in general well detected, while including visible (VIS) information brings more weight to low-level clouds which then introduces biases, especially in the case of thin cirrus overlying low-level clouds. In that case during daytime, IR sounders as well as MODIS-CE still provide information on the uppermost cloud (cirrus), while for all Cloud_cci datasets the uppermost cloud (thin cirrus) is not identified, but the the low-level clouds. The interpretation is then more complicated: instead of always identifying the uppermost cloud in the case of multilayer clouds, the Cloud_cci data sets indeed capture the uppermost cloud layer during night, while during daytime there is a chance that the uppermost cloud is not detected and instead the low-level cloud is detected (like for MODIS-ST).

Land-ocean differences, also shown in Figure 2-1, first confirm that CA is about 0.15 larger over ocean than over land, essentially due to low-level clouds, whereas there are about 0.1 more high-level and mid-level clouds over land than over ocean. Differences between the datasets tell us about their sensitivity to underlying surface parameters: *During daytime* (when including VIS) Cloud_cci AVHRR, Cloud_cci MODIS and Cloud_cci AATSR have a near-zero ocean-land difference for CAHR, which lets assume that these Cloud_cci datasets mostly miss thin cirrus over land. During night-time (when only using spectral IR) the Cloud_cci datasets agree with the other datasets on the ocean-land difference in CAHR, while CAMR and CALR behave more closely to ISCCP (which uses only one IR atmospheric window channel).

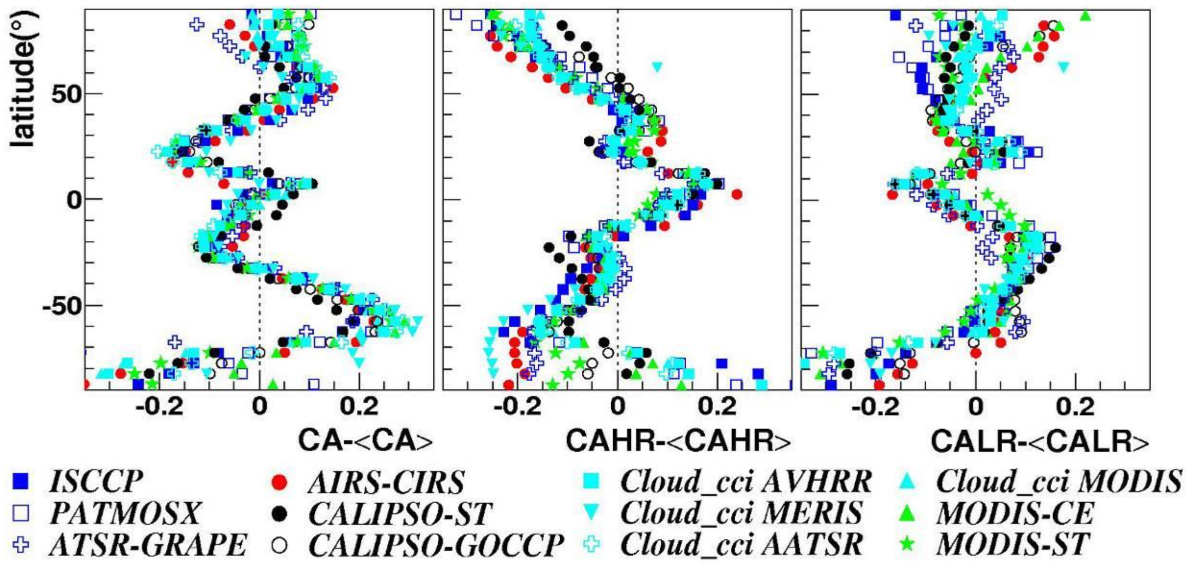


Figure 2-2 *Latitudinal anomalies of total cloud amount (left), relative high cloud amount (middle) and relative low cloud amount (right). Comparisons of Cloud_cci data are shown for 2007 with L3 data from the GEWEX Cloud Assessment data base.*

Figure 2-2 gives further insight, considering the latitudinal variability of CA, CAHR and CALR, which is presented here as a latitudinal anomaly. In general, the latitudinal variation in all three variables agrees quite well for all datasets within 60°N and 60°S (except ATSR-GRAPE). This shows an improvement of Cloud_cci AATSR compared to ATSR-GRAPE. At higher latitudes, Cloud_cci MERIS-AATSR is less coherent, but this is possible, since data are only available during daylight. At these latitudes snow and ice during winter make passive remote sensing also less reliable.

A more detailed way of estimating the reliability of cloud amount reproduced by a given dataset was described in CAR-2013. For each latitudinal/longitudinal grid box the mean GEWEX cloud assessment reference value is estimated as well as the rms of its variability among its datasets. For each of the tested Cloud_cci datasets, a deviation from the mean is calculated and expressed it in terms of rms values (sigmas).

The geographical maps of these values, separately for daylight conditions and night-time conditions, are shown in Figure 2-3 for Cloud_cci AVHRR, in Figure 2-4 for Cloud_cci MODIS, in Figure 2-5 for Cloud_cci AATSR and in Figure 2-6 for Cloud_cci MERIS-AATSR (only daytime).

When comparing daytime and night-time observations, the sensitivity to high-level clouds, in particular over land, is strongly reduced for all Cloud_cci datasets when using VIS information in addition to spectral IR information in the Optimal Estimation retrieval method by up to 6 sigmas from the reference dataset (as already deduced from Figure 2-1), whereas low-level clouds demonstrate a better match during daytime than during night-time.

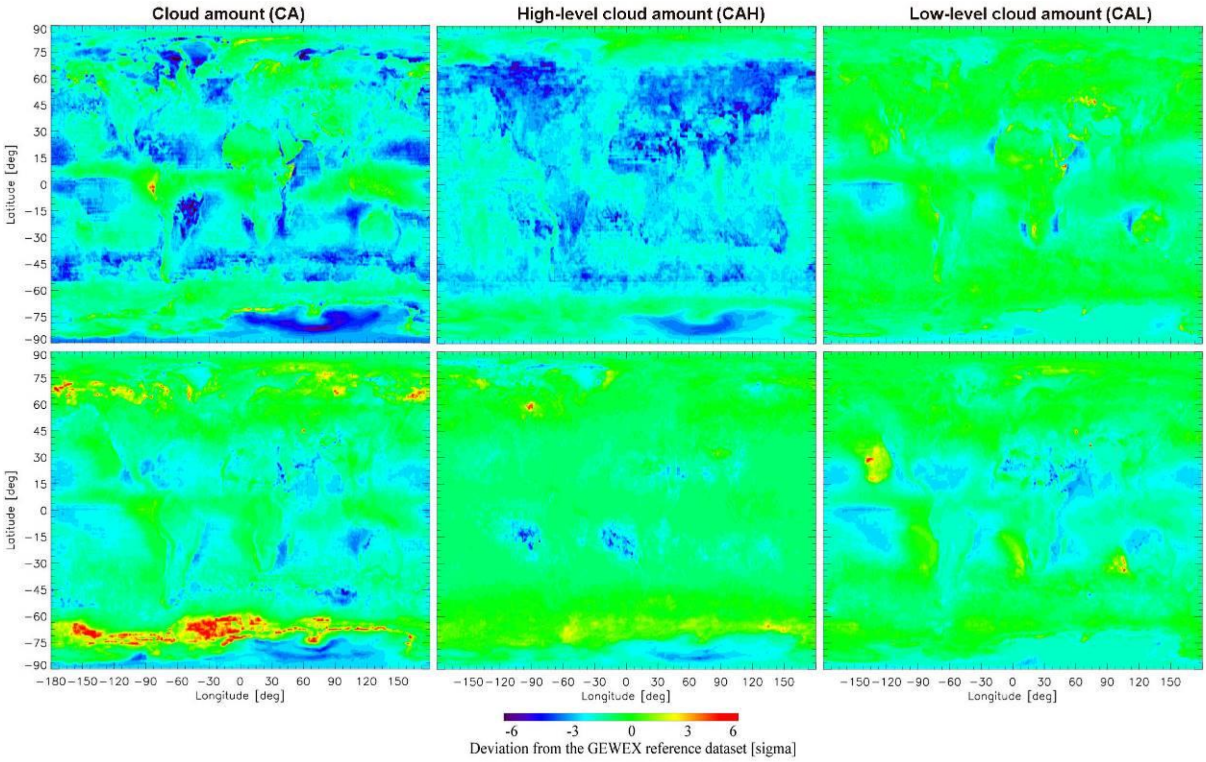


Figure 2-3 *Differences between annual averages of CA (left), high-level cloud amount (CAH, middle) and low-level cloud amount (CAL, right) from ESA Cloud_cci AVHRR and from the GEWEX Cloud Assessment reference dataset in units of standard deviation (σ), separately for mostly daytime conditions (1:30 PM, top) and mostly night-time conditions (1:30 AM, bottom).*

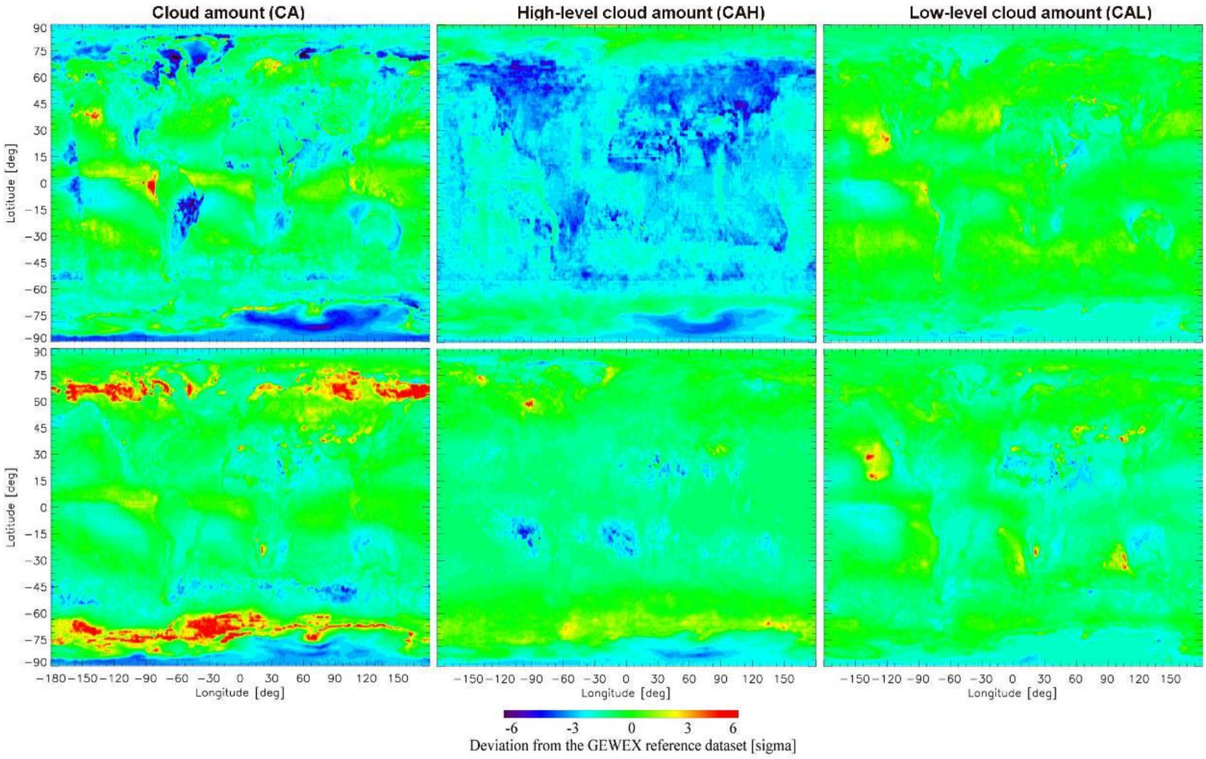


Figure 2-4 *Same as Figure 2-3, but for Cloud_cci MODIS.*

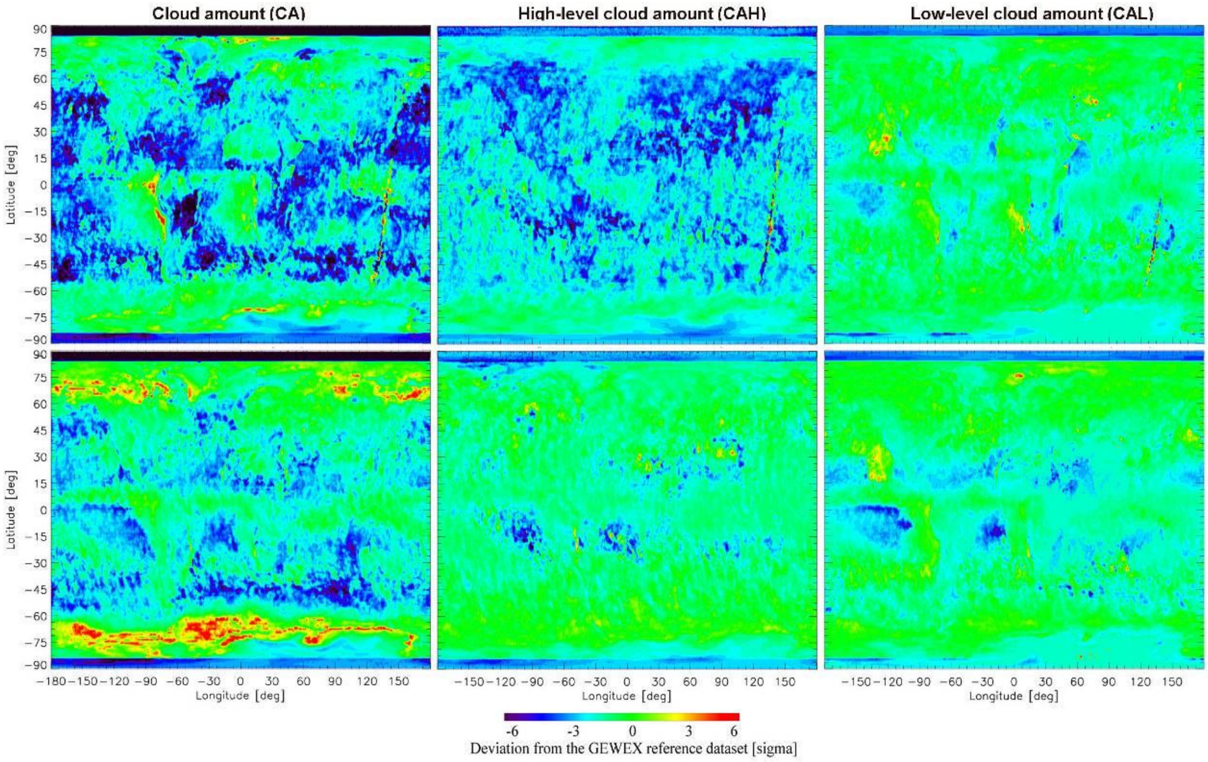


Figure 2-5 *Same as Figure 2-3, but for Cloud_cci AATSR.*

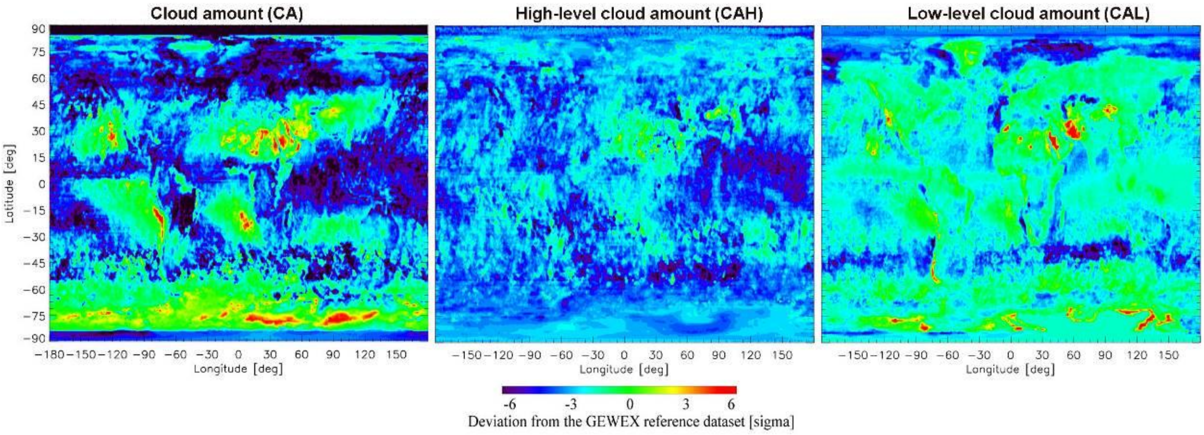


Figure 2-6 *Same as Figure 2-3, but for Cloud_cci MERIS-AATSR, only daytime data available.*

In general results for Cloud_cci AHVRR and Cloud_cci MODIS are similar, with a slightly better agreement to the reference dataset for MODIS during daytime. The performance for Cloud_cci AATSR is worse, followed by Cloud_cci MERIS-AATSR.

Figure 2-3 to Figure 2-6 reveal clearly that the Cloud_cci cloud retrievals, based on the optimal estimation technique, using spectral information from the VIS to the IR, have great difficulties to identify cirrus, especially over land, a fact which was already revealed during the GEWEX Cloud Assessment. The method performs better when excluding the VIS information, though at the cost of performance for low-level clouds.

2.1.3 Seasonal variations of clouds

Here, the seasonal variation of cloud amount and cloud temperature (CT) and its correlation between the Cloud_cci and reference datasets is studied. A seasonal variation is defined as a difference between the monthly averaged parameter in a given latitude/longitude grid box and an annual average of the same parameter. Figure 2.7 shows that the seasonal behaviour of CA, CAHR, CALR, and CT for the Cloud_cci datasets qualitatively agrees with that of the reference datasets. However, Cloud_cci MERIS-AATSR is off in many areas, especially in CT anomaly, which is coherent with the differences shown in the other figures and which is associated with cloud pressure mis-assignment and cloud detection sensitivity of this dataset.

As in CAR-2013, to investigate the seasonal coherence of the retrieved parameters, Pearson product-moment correlation coefficients (r) are estimated, built for each latitude/longitude grid box, between the 12 monthly means of the Cloud_cci datasets and of the reference dataset. A high correlation coefficient means similar behaviour and coherence in seasonal variability of a given parameter even though the absolute values might differ. Large values here mean that the combination of measurement, radiance model, and retrieval algorithm produce the picture which varies in accordance with reality (represented by the reference dataset and its uncertainties). One has to note that a high correlation coefficient does not guarantee the matching of absolute values of, let say, cloud radiative properties, but it tells that the retrieval algorithm in combination with its ancillary data properly capture all physical variations of atmosphere and corresponding measured radiance.

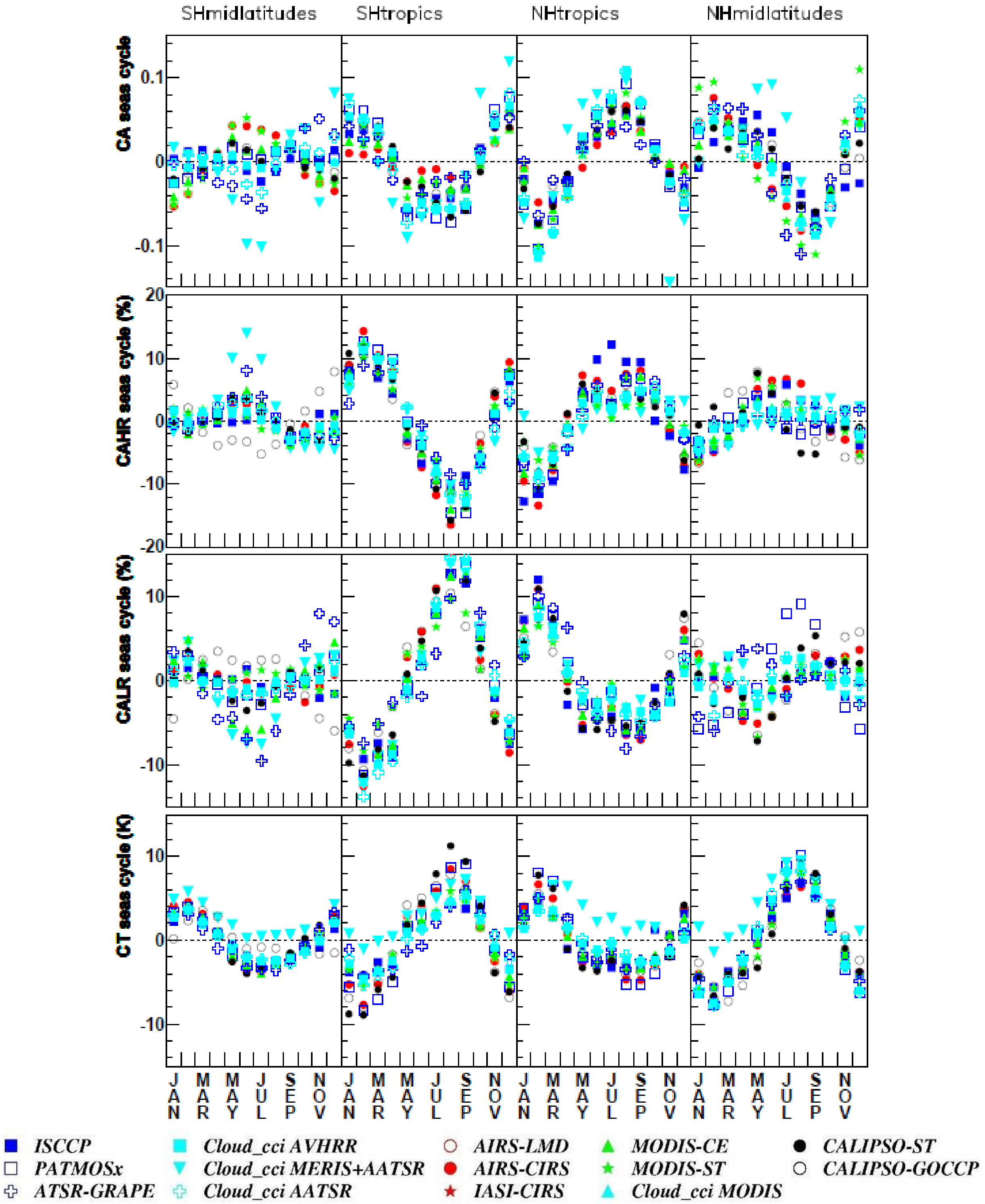


Figure 2-7 *Seasonal cycle (relative) of CA, CAHR, CALR and CT for NH midlatitudes (30°N-60°N), NH tropics (0°-30°N), SH tropics (0°-30°S), and SH midlatitudes (30°S-60°S). Comparisons of Cloud_cci Data for 2007, daytime observations, with L3 data from GEWEX Cloud Assessment data base.*

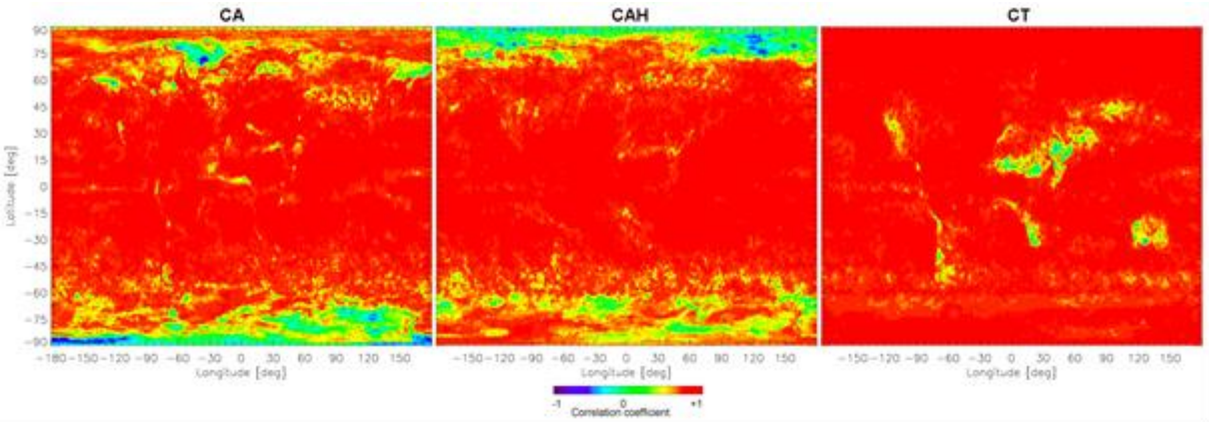


Figure 2-8 *Geographical maps of Pearson product-moment correlation coefficients calculated for CA (left), CAH (middle), and CT (right) temporal variation of Cloud_cci AVHRR versus reference.*

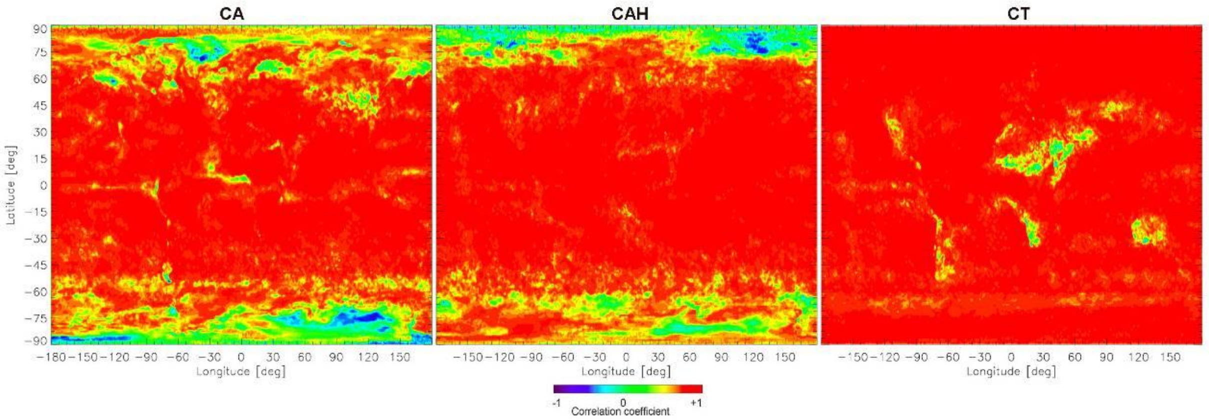


Figure 2-9 *Same as Figure 2-8, but for Cloud_cci MODIS dataset.*

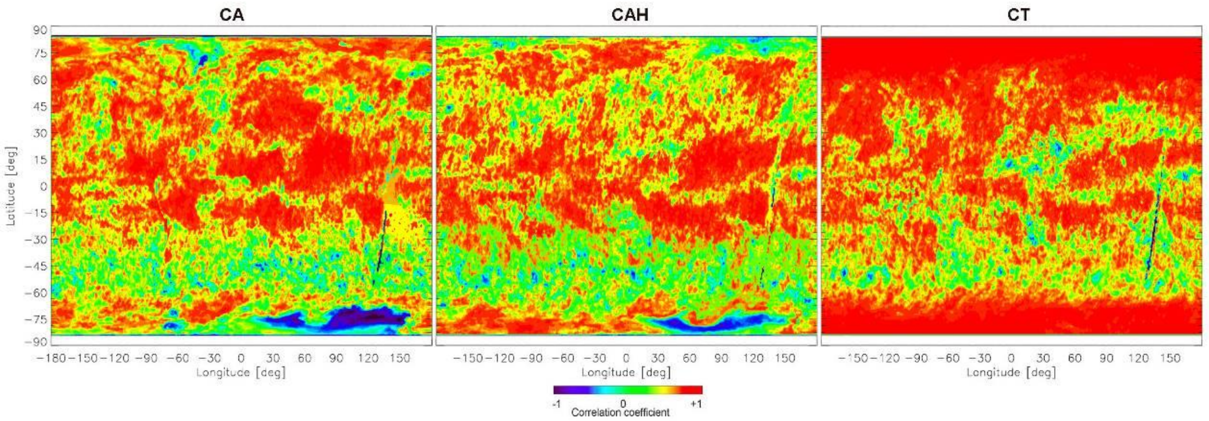


Figure 2-10 *Same as Figure 2-8, but for Cloud_cci AATSR dataset.*

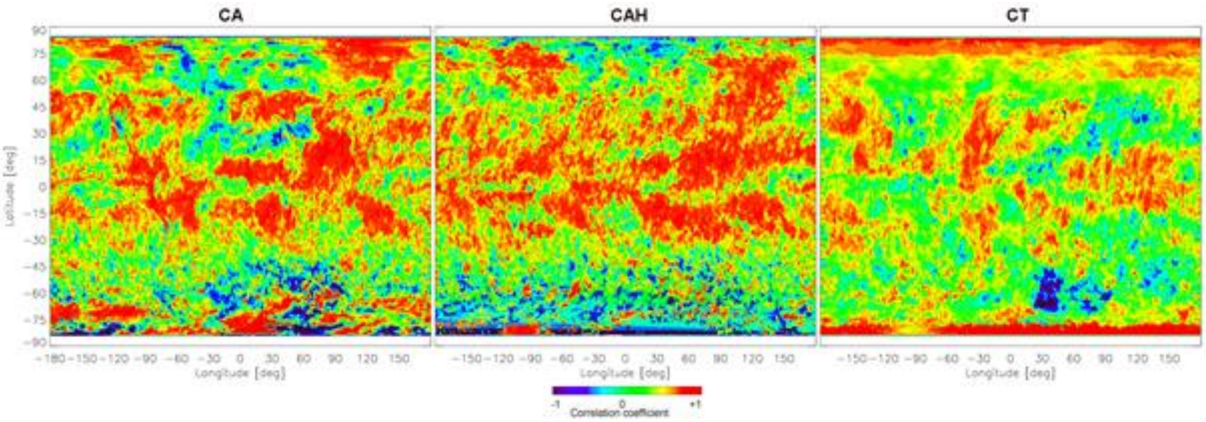


Figure 2-11 *Same as Figure 2-8, but for Cloud_cci MERIS-AATSR dataset.*

Figure 2-8 to Figure 2-11 present geographical maps of these coefficients built for CA, CAH, and CT for Cloud_cci AVHRR, Cloud_cci MODIS, Cloud_cci AATSR and Cloud_cci MERIS-AATSR, respectively. At present, the correlations between Cloud_cci AVHRR / MODIS and GEWEX are good everywhere except for the poles and Greenland, snow and ice covered areas which are known for difficulties in the retrievals associated with strong and varying background radiance. In these regions only active instruments, like CALIPSO and CloudSat, provide the most reliable information.

Considering Cloud_cci AATSR (Figure 2-10), even though its average values (Figure 2-1, Figure 2-2, Figure 2-7) are in relatively good agreement with the reference dataset, their seasonal behaviours are often uncorrelated, especially over the ocean in the Southern hemisphere.

As for Cloud_cci AATSR, the seasonal variation of high-level clouds of Cloud_cci MERIS-AATSR is more coherent with the one of the reference dataset in the tropics and the patterns are similar to that of Cloud_cci AATSR.

Summarizing this section, it is shown that Cloud_cci AVHRR and Cloud_cci MODIS datasets capture the seasonal variations well, while Cloud_cci AATSR and Cloud_cci MERIS-AATSR datasets are often uncorrelated with the reference dataset. Correspondingly, one should expect that the instantaneous radiation fields calculated with these datasets will be sometimes out of phase despite the fact that their averages can be coherent with reality.

2.1.4 Cloud-top pressure and temperature

All passive sensors selected for the study provide a '"radiometric height" of the cloud, which roughly corresponds to half-way between cloud top and "apparent" cloud base. The latter corresponds to the real cloud base when optical depth is smaller than 3. The "radiometric height" may lie as much as a few kilometres below the "physical height" of the cloud top, depending on the cloud extinction profile and vertical extent. High-level clouds in the tropics generally have "diffusive" cloud tops (meaning that the optical depth increases only slowly from cloud top downwards), for which retrieved cloud temperature may be as much as 10 K larger than the cloud top temperature corresponding to a lidar height.

Figure 2-12 and Figure 2-13 present probability density functions (PDFs) of CP, while Figure 2-14 presents PDFs of CT. These can be interpreted as the vertical distribution of the higher-most detected clouds in the atmosphere, since passive remote sensing only has access to the highest cloud layer in the case of multi-layer cloud situations. The distributions have been built using the histograms which are a part of the L3 products. The normalized PDFs have been obtained by dividing the histograms by the number of cloudy samples. It is stressed that it is essential to use histograms instead of the averaged values of CP and CT since the same averaged value can be obtained both from a bimodal distribution (as expected in the tropics) or from a single-mode distribution centred in the middle.

The strong bimodality in the tropics with strong peaks at 950 hPa and between 250 and 150 hPa means that the tropics have few mid-level clouds, in agreement with local observations using ground-based radar and with CALIPSO-CloudSat observations. The decrease of bimodality and range in CP and CT from tropics towards poles is essentially linked to the decrease of the tropopause height and a change in the type of atmospheric storm from convective to baroclinic cyclone.

Again, the assessment is done separately for observations mostly during daytime and mostly during night-time. For 2007, observations are compared around 1:30 PM (10:30 AM for AATSR and MERIS) and 1:30 AM (10:30 PM for AATSR), respectively. From Figure 2-12 it can be deduced that the PDFs of all four Cloud_cci datasets do not reproduce the bimodality in the tropical and midlatitudes during daytime and that the peak of low-level clouds is about 100 hPa lower (overestimation of low-level cloud height) than for the GEWEX reference dataset. The low-level cloud peak is another 100 hPa lower for Cloud_cci MERIS-AATSR, except in polar regions where it is higher, here in agreement with AIRS-CIRS. During night-time a hint of bimodality appears, but the overestimation of low-level cloud height stays the same. The net effect of these differences compared to the reference dataset is as follows: if these clouds are used in the radiative transfer calculations, this will lead to an overestimate of atmospheric cooling in the long wave.

In the CT histograms (Figure 2-14) the bimodal distributions are not visible, whereas CALIPSO, the GEWEX Cloud Assessment reference and AIRS-CIRS show bimodal distributions in the tropics and midlatitudes. This means that the pressure to temperature and L3 conversion in the Cloud_cci datasets seems to amplify the problems in the datasets. The difference between CALIPSO and AIRS-CIRS can be partly explained by the fact that CALIPSO determines the cloud top and the AIRS-CIRS cloud height corresponds to a height at which cloud optical depth of about 0.5 is reached and also on the difference in sampling, as IR sounders large footprints might include different cloud layers (Stubenrauch et al., 2017). The peak corresponding to low-level clouds is more comparable to that of the reference datasets, though still slightly colder especially during night-time. The polar distributions of Cloud_cci MERIS-AATSR show colder and lower clouds than retrieved by the other datasets and CALIPSO.

In addition, an assessment is presented of Cloud_CCI AVHRR, Cloud_cci MODIS and Cloud_cci AATSR at 7:30 AM / 7:30 PM and 10:30 AM and 10:30 PM, respectively, in comparison with IASI-CIRS, PATMOSx and MODIS-ST at similar observation times, for the year 2009. While the peak of high-level clouds of PATMOSx is much overestimated, IASI and MODIS-ST show similar peaks. Only MODIS-ST shows too many low-level clouds, because collection 5 of MODIS-ST misidentifies thin cirrus as low-level clouds (Stubenrauch et al., 2013). From Figure 2-13 it can deduced again that the PDFs of all Cloud_cci datasets do not reproduce the bimodality in the tropical and midlatitudes during daytime and that the peak of low-level clouds is about 100 hPa lower (overestimation of low-level cloud height) than for the GEWEX reference. While during night-time bimodality appears for Cloud_cci AVHRR and Cloud_cci AATSR, the performance of Cloud_cci MODIS on board Terra did not improve very much, because it looks like high-level clouds partly appear as mid-level clouds. Again, the overestimation of low-level cloud height stays the same.

Two-dimensional histograms of COD and CP distributions per 1° latitude x 1° longitude grid, available in the GEWEX Cloud Assessment data base, provide a basis for a more detailed comparison and are largely used for model evaluation. The two-dimensional normalized frequency distributions presented in Figure 2-15 show how the clouds of different optical depth and height are distributed in the atmosphere. This type of histogram is valuable for climate model evaluation and for estimating atmospheric, surface, and top of atmosphere radiative fluxes.

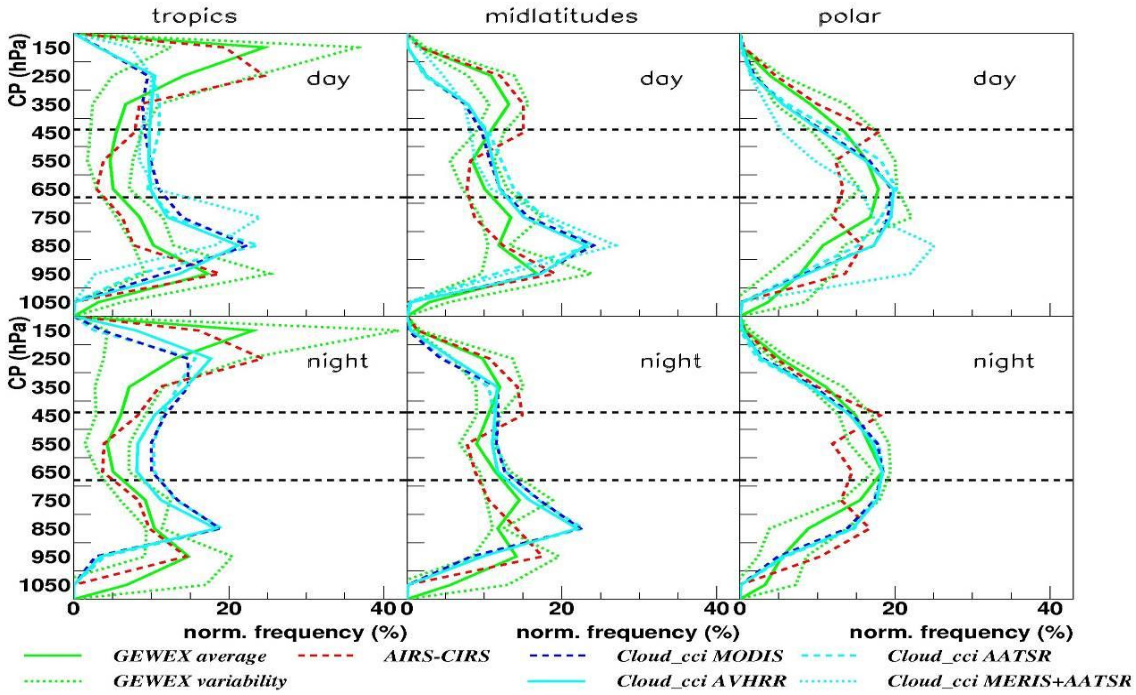


Figure 2-12 *Cloud pressure histograms, separated into tropical (30°N-30°S), midlatitude (30°-60°) and polar (60°-90°) latitude bands. Statistics are shown for 2007, with observation times of 1:30 PM (10:30 AM for AATSR and MERIS), corresponding mostly to day time (above), and of 1:30 AM (10:30 PM for AATSR), corresponding mostly to night time (below).*

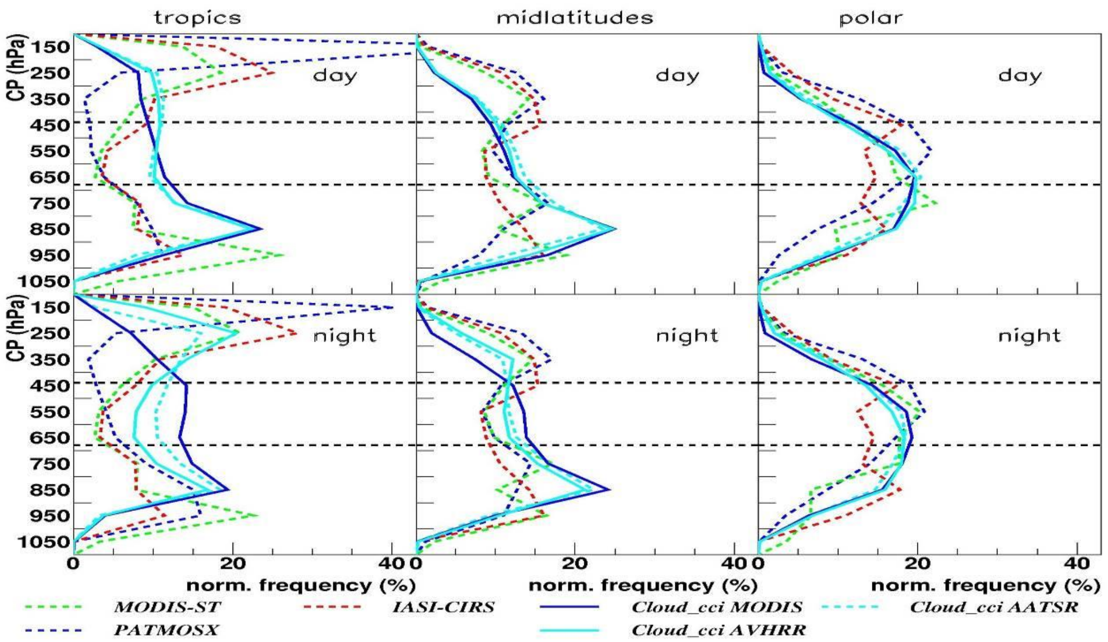


Figure 2-13 *Same as Figure 2-12. Statistics are shown for 2009, with observation times of 7:30 AM (PATMOSX, AVHRR), 9:30 AM (IASI-CIRS) and 10:30 AM (MODIS, AATSR), corresponding mostly to day time (above), and of 7:30 PM, 9:30 PM and 10:30 PM, corresponding mostly to night time (below).*

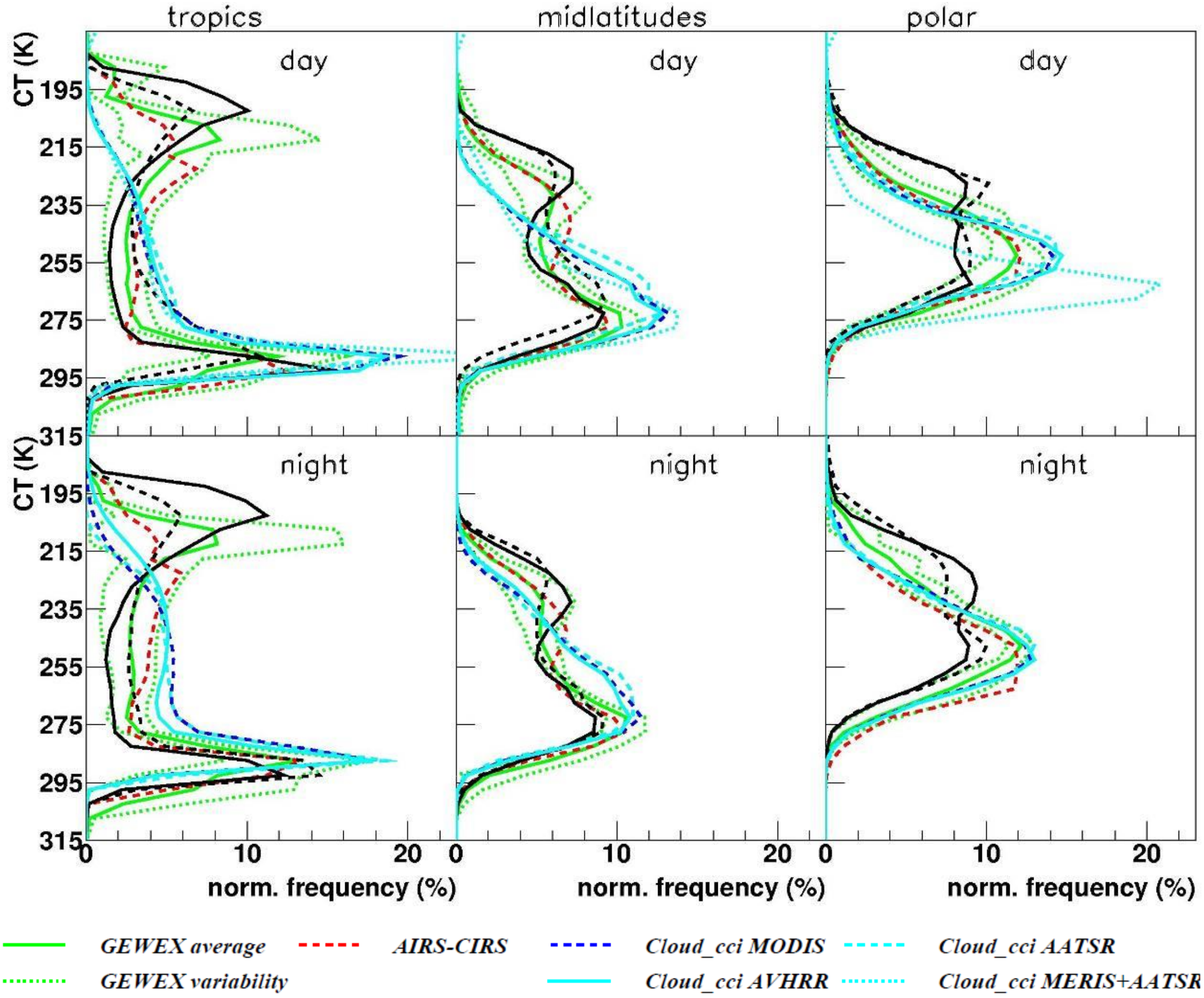


Figure 2-14 *Cloud temperature histograms are shown, separated into tropical (30°N-30°S), midlatitude (30°-60°) and polar (60°-90°) latitude bands. The statistics is presented for 2007, with observation times of 1:30 PM (10:30 AM for AATSR and MERIS), corresponding mostly to day time (above), and of 1:30 AM (10:30 PM for AATSR), corresponding mostly to night time (below).*

As range of reference two datasets of the GEWEX Cloud Assessment data base are used: AIRS-LMD, sensitive to cirrus, and ISCCP, the GEWEX dataset. In the case of AIRS-LMD, retrieved cloud emissivity was transformed to COD, and since cloud emissivity saturates at 1, the last COD interval has no values. For ISCCP clouds around 100 hPa with very small COD are suspicious. These two datasets, presented in the first two rows of Figure 2-15, show tropical distributions which are characterized by bimodal pressure distributions discussed above. This structure is also traceable in midlatitudes. Current versions of Cloud_cci AVHRR, Cloud_cci MODIS and Cloud_cci AATSR demonstrate a bimodal structure, which is consistent with the reference data, but the upper peaks are much less pronounced, that corresponds to less frequent (less effective) detection of high-level clouds. The Cloud_cci MERIS-AATSR dataset presents a distribution continuing at all pressures towards very high and thin clouds.

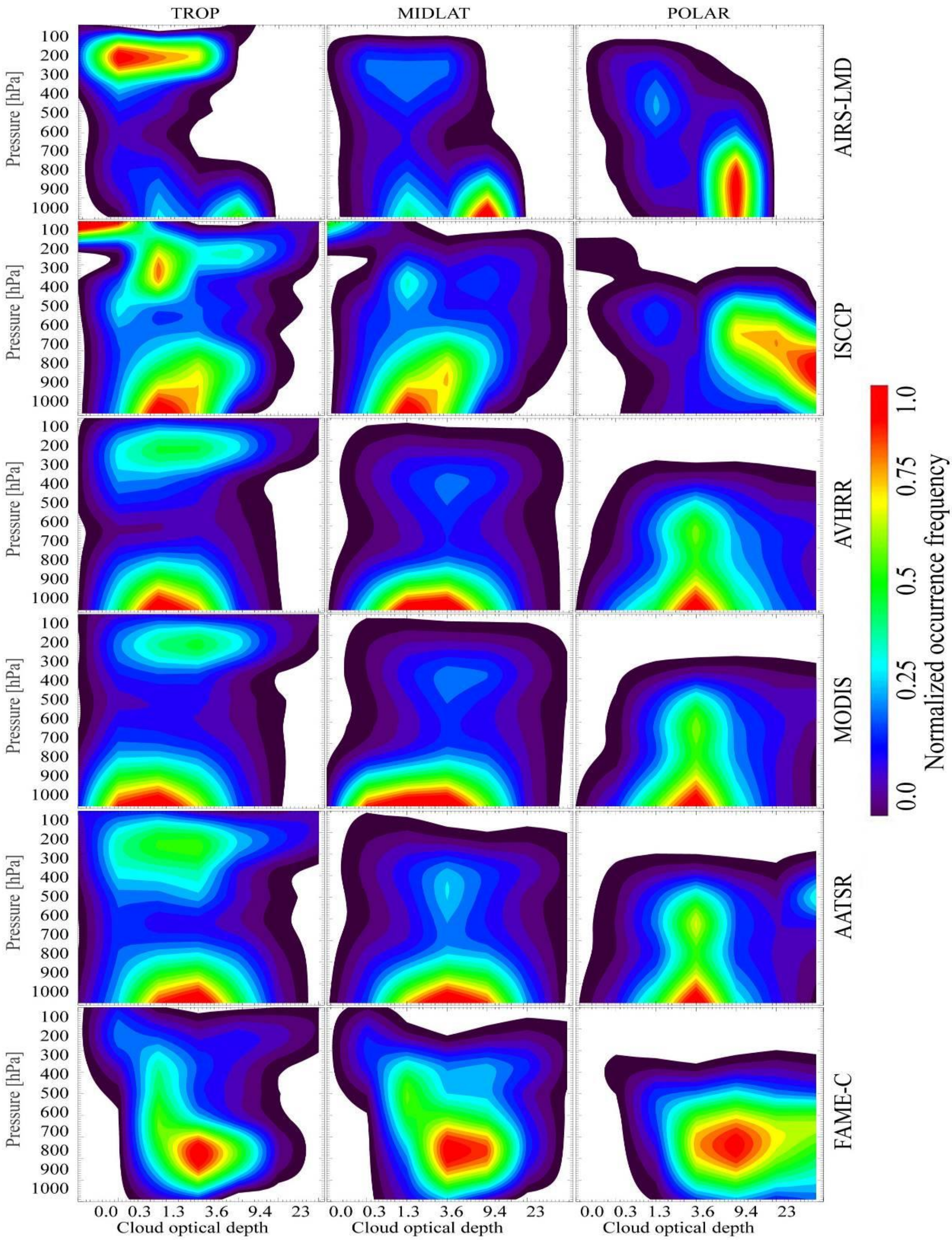


Figure 2-15 *Two dimensional histograms of annual COD versus CP for three latitude bands: tropical (15S-15N), midlatitude (15S/N-60S/N), and polar (60S/N-90S/N).*

2.3 Conclusions of Level-3 assessment with GEWEX data base

Based on the analysis of total and relative cloud amounts, their regional and seasonal behaviour, as well as of their distributions in cloud pressure and temperature the following conclusions is drawn regarding the *ESA Cloud_cci* datasets of AVHRR, MODIS, AATSR and MERIS-AATSR:

While global total cloud amount compares well to reference datasets from other passive remote sensing cloud datasets, cloud amount is underestimated over large parts of the ocean, especially for AATSR and MERIS-AATSR.

Global relative cloud amounts compare well to the reference cloud datasets during night-time, when only IR information is available. However, the amount of high-level clouds, in particular over land, is strongly underestimated for all *Cloud_cci* datasets when using VIS information in combination with spectral IR information in the Optimal Estimation retrieval method. This was already revealed during the GEWEX Cloud Assessment, with the participation of ATSR-GRAPE. Thin cirrus over low-level clouds is then misidentified as mid-level and low-level cloud.

Low-level clouds demonstrate a better match during daytime than during night-time, since these clouds are easier identified using VIS information. One should nevertheless keep in mind that passive remote sensing detects low-level clouds only when there is no higher cloud above them.

Distributions of cloud pressure reveal that the height of low-level clouds seems to be overestimated and that even during night the Cloud_cci datasets detect less high-level clouds than IR sounders. The latter are most abundant in the tropics.

The performance of Cloud_cci AVHRR and MODIS are comparable, with a good coherence of the seasonal cycle, whereas AATSR has a slightly lower performance and MERIS-AATSR has the worst performance.

3. Bias removal of CFC L3 using SYNOP

To be suitable for climate analysis, satellite-derived cloud datasets have to meet the challenging requirements including those for accuracy, precision and decadal stability (URDv2). The longest CFC dataset of Cloud_cci project (1982-2014) derived from the afternoon NOAA satellites has been thoroughly evaluated in the Product Validation and Intercomparison Report (PVIRv4).

The PVIR reveals some inhomogeneities in CFC between sensors (e.g. NOAA-7 and NOAA-9) as well as variation of performance within the lifetime of individual sensors. These inaccuracies are related to degradation of satellite sensors, radiometric and geometric calibration problems, and satellite orbital drift. Some of these aspects have been resolved within the scope of Cloud_cci retrieval algorithms (Poulsen et al., 2011; Stengel et al., 2015). However, orbital drift and varying equatorial crossing times of consecutive satellites have not been accounted for. This may inhibit the use of the original dataset for investigating the cloud variability and change over the last 3 decades.

Several methods exist for correcting satellite-derived cloud cover time series. Foster and Heidinger (2013) derived corrected CFC daily means from PATMOS-x dataset by fitting a mean diurnal cycle to a sinusoidal function derived over all NOAA's. An average corrected daily value of cloudiness is then interpolated from the fit function using all available ascending, descending, morning, and afternoon satellite overpasses. The application of this correction method is inhibited for the Cloud_cci CC4CL AVHRR-PM dataset since it only utilizes one satellite at a time. Devasthale et al. (2012) succeeded to delineate the signal of orbital drift in the AVHRR-based CFC time series by means of rotated empirical orthogonal function (REOF). However, the method is shown to be sensitive to decisions which EOF modes do reflect unnatural CFC variability. Therefore, it is necessary to rigorously test that the large scale geostatistical features are preserved in the corrected data set. Similar risk of removing from original dataset correct spatio-temporal patterns has been pointed out by Norris and Evan (2015) while detecting (by least squares best-fit) and then removing the CFC variability linked to known artefacts (e.g. orbital drift).

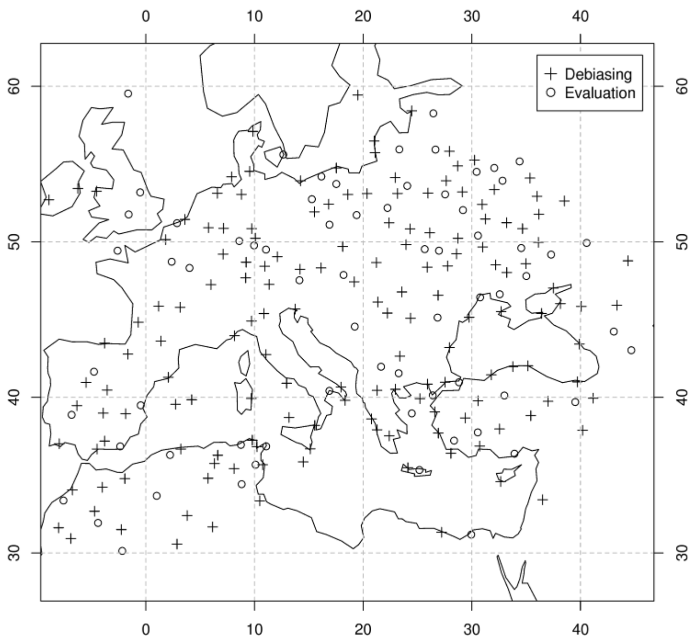


Figure 3-1 *SYNOP sites used for debiasing and evaluation of CC4CL AVHRR-PM CFC monthly means.*

In this section, a simple statistical method is proposed for correcting CFC by debiasing the AVHRR-PM CFC data using synoptic observations. The method is based on interpolation of these observations using satellite data as an explanatory variable. In this context, the proposed method relies on the strong assumption that synoptic observations are accurate and homogenous. Moreover, it is assumed that CFC errors caused by the artefacts such as orbital drift and transition between satellites are related to the occurrence and intensity of cloudiness diurnal cycle. Hence, it is assumed that these errors are spatially correlated (i.e. can be interpolated). Interpolation can be applied to monthly means, i.e. not explicitly resolving the diurnal cycle of cloudiness.

3.1 Method

Synoptic observations from the ECMWF archive were used for debiasing and evaluation. The archive initially contains data for over 6000 globally distributed sites. From these sites have been selected for a geographic range of 30N to 60N and 10W to 45E. In order to ensure the collection of long-term homogenous data series, only stations are selected where observations were continuously performed in 1982–2014, at least every 6 hours with a maximum break of 30 days. For each site cloud amount is used observed with the highest temporal frequency (up to 1 hour) that was reported for the whole 33-year period. Thus the frequency of observations could vary between sites, but remained stable in time for each site. Instantaneous observations were transformed from the okta scale to cloud fractional cover, and aggregated to monthly means. Further, sites have been excluded for which the Standard Normal Homogeneity Test (SNHT, Alexandersson, 1986, Khaliq and Ouarda, 2007) detected any inhomogeneity in a time series of cloud amount monthly anomalies.

Since SYNOP stations are unevenly distributed in geographic space, a regular grid was created. For each 2x2 degree geographic grid cell the SYNOP sites with the most valid observations were used. In case two or more sites were present within one grid cell, one site was used as training and one as validation site. This selection procedure yielded a total of 158 SYNOP sites used for training, and 68 sites used for validation (Figure 3-1).

Satellite data consisted of 33 years of CFC monthly means derived from NOAA-AVHRR afternoon satellites (CC4CL AVHRR L3C). These included data from consecutive afternoon NOAA satellites: NOAA-7, NOAA-9, NOAA-11, NOAA-12 (morning satellite due to a gap in data from afternoon satellites), NOAA-14, NOAA-16, NOAA-18 and NOAA-19.

Debiasing of satellite-derived CFC employed interpolation of synoptic observations using satellite data as an explanatory variable. For each month (i.e. 1982-01, 1982-02, etc.) kriging with external drift was applied to SYNOP-based monthly means (at 158 sites). The variogram was fit automatically to regression residuals and kriging performed using the R package 'automap' (Hiemstra et al., 2009). Based on explanatory analysis, the spherical variogram model was chosen and imposed for each fitting. The native spatial resolution of 0.05 degree was preserved during interpolation.

To assess the accuracy of debiased satellite CFC, the mean bias error is used, defined as mean difference between debiased CFC and reference SYNOP data. To express the precision, the bias-corrected root mean squared error (bcRMSE) was used. Finally, for a trend analysis linear trends are used derived using Theil-Sen estimates (Theil, 1950) and their significance was estimated with the Mann-Kendall test (Kendall, 1938; Mann, 1945) and adjusted using Benjamini-Hochberg (Benjamini and Hochberg, 1995) method.

Table 3-1 *Mean bias error (MBE) and bias-corrected root mean square error (bcRMSE) of AVHRR-PM uncorrected and corrected (debiased) mean monthly cloud fraction as compared to synoptic observations aggregated for NOAA missions.*

	Corrected		Uncorrected	
NOAA	MBE	bcRMSE	MBE	bcRMSE
AVHRR-PM	-0.68	6.42	4.05	14.90
7	-0.70	6.57	5.28	14.66
9	-0.73	6.55	4.25	14.90
11	-0.91	6.99	3.04	15.17
12	-0.24	6.97	9.16	15.41
14	-0.79	6.12	5.77	14.65
16	-0.80	6.03	3.68	14.08
18	-0.21	6.27	2.82	15.25
19	-0.53	6.32	3.25	14.52

3.2 Results of bias removal

Averaged over 68 evaluation sites (Table 3-1), the corrected (debiased) dataset (MBE=-0.68%, bcRMSE=6.42%) significantly outperforms the original one (MBE=4.05%, bcRMSE=14.90%).

Figure 3-2 reveals that the performance differs among sites. For the vast majority of them (11 out of 68), the original dataset overestimates the reference SYNOP CFC, while after correction 28 sites overestimate and 40 underestimate the reference. The are 16 sites for which the correction method increased the absolute bias, most of them located at the edges of the interpolation area. Concurrently, the bcRMSE decreased for all sites during correction.

The performance of the debiased dataset (Table 3-1) is stable among NOAA missions (absolute bias < 1%, and bcRMSE within 6-7%). Compared to the uncorrected data, the correction also decreases the performance differences among NOAA satellites. Figure 3-3 displays concurrently lower mean errors and lower variability. The errors are also more stable in time. These facts reveal that debiasing can implicitly remove the inhomogeneity in CFC time series due to changing overpass times.

Trends in CFC monthly anomalies are comparable in original and corrected dataset (Figure 3-4).The magnitude of trends is mostly smaller, but the correction keeps their signs unchanged (positive trends are less positive, negative trends are less negative). More prominent differences can be observed over water. It has to be noted however that CFC there was interpolated from inland sites. Therefore, the sharp edges between trends on land and water are obscured in debiased data. Nevertheless, the similarity of trends over land before and after the correction encourages to conclude that observed signal is related to natural changes in the last 30 years. Performance of both CFC datasets over water should be further investigated prior any conclusions about potentially observed trends.

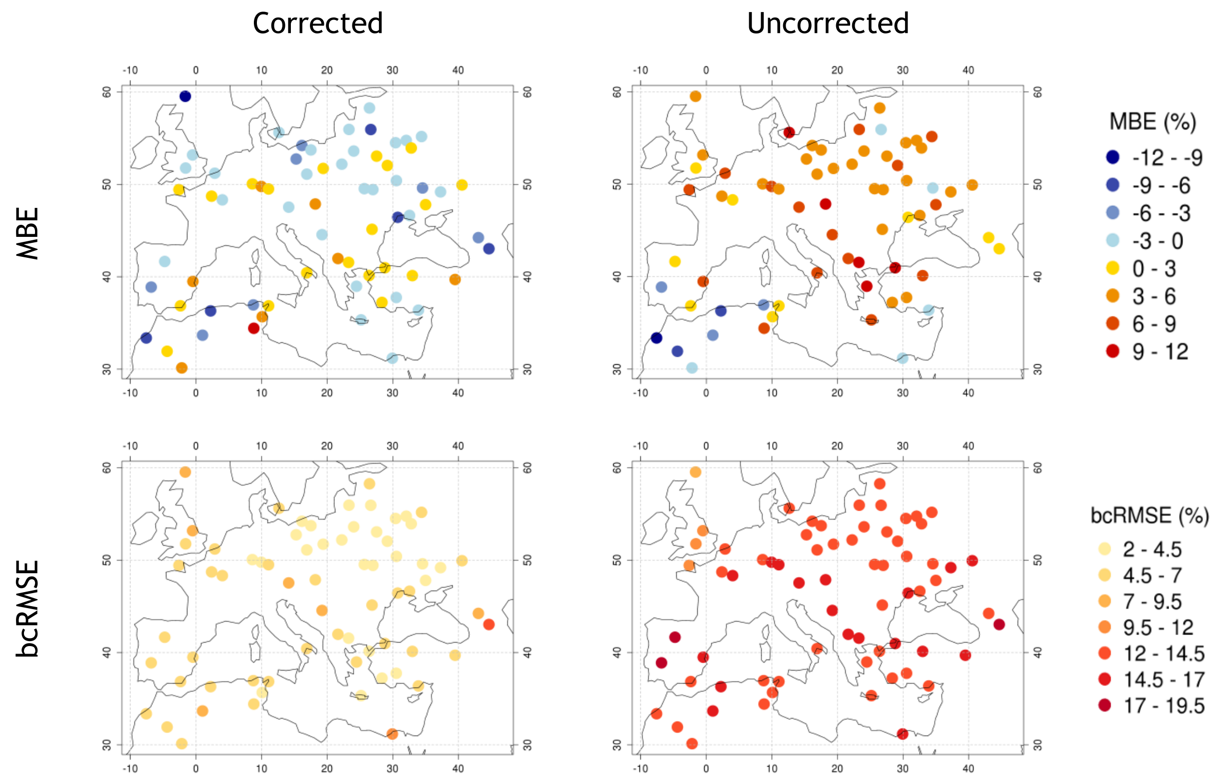


Figure 3-2 *Mean bias error (MBE) and bias-corrected root mean square error (bcRMSE) of CC4CL-AVHRR-PM uncorrected and corrected (debiased) mean monthly cloud fraction as compared to synoptic observations.*

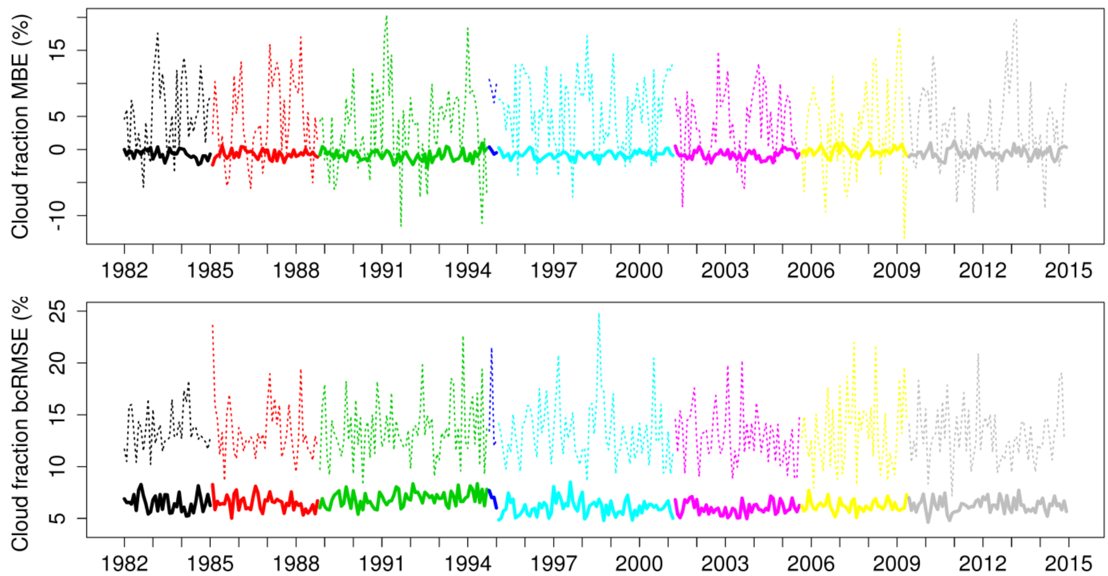


Figure 3-3: Time series of mean bias error (upper panel) and bias-corrected root mean square error (lower panel) of CC4CL-AVHRR-PM as compared to synoptic observations. Dashed line reveals performance of uncorrected data, thick solid line of corrected (debiased) data. Colours represent consecutive satellite missions: NOAA-7 (black), NOAA-9 (red), NOAA-11 (green), NOAA-12 (dark blue), NOAA-14 (light blue), NOAA-16 (purple), NOAA-18 (yellow) and NOAA-19 (grey).

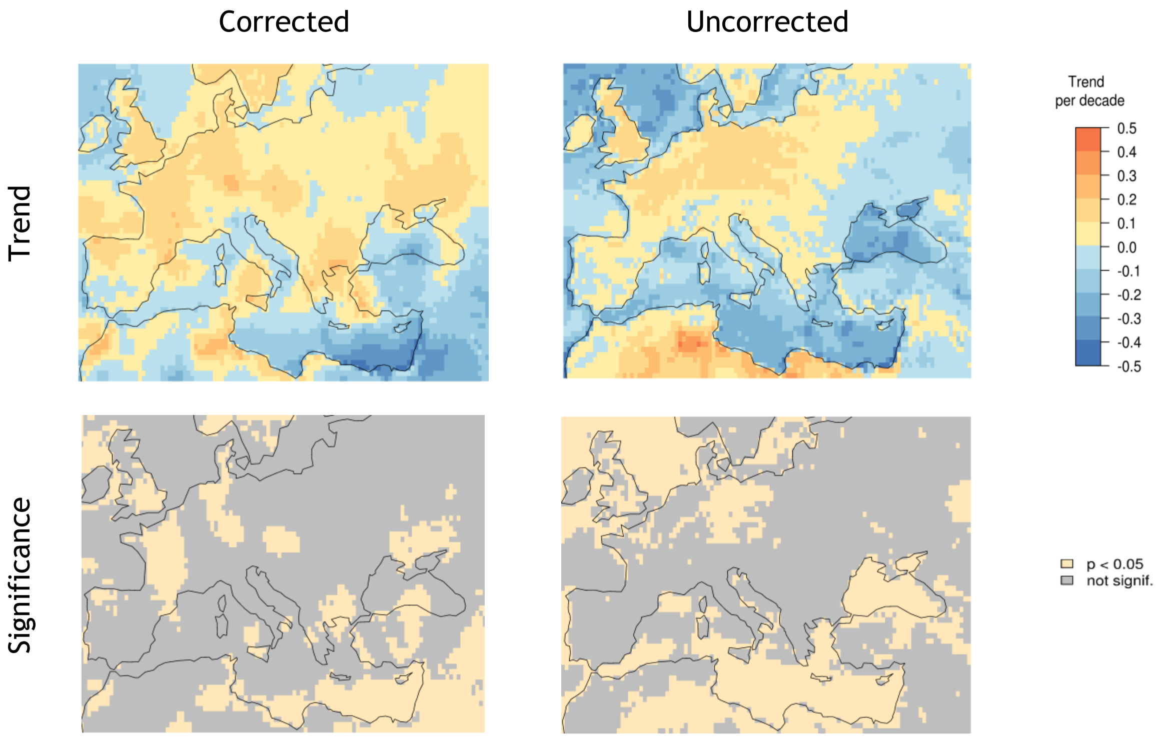


Figure 3-4 *Map of CC4CL-AVHRR-PM Theil-Sen monotonic trend (upper panels) and its statistical significance according to the Mann-Kendall test adjusted using Benjamini-Hochberg method (lower panels) based on the cloud fraction monthly standardized anomalies in 1982-2014.*

4. Comparison to reanalysis data

4.1 ERA-Interim reanalysis

The ERA-Interim global atmospheric reanalysis (Dee et al., 2011) provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) is produced by the Integrated Forecast System (IFS). The IFS contains the forecast model, which fully couples the atmospheric, land, and oceanic components leading to a physically realistic and consistent model. The data is available starting from 1979 and is continuously extended in near-real time.

In large-scales models only the bulk properties of clouds can be taken into account. Hence, in ERA-Interim clouds are described by a fully prognostic cloud scheme, which has been developed by Tiedtke (1993). Cloud cover and cloud water/ice content are derived from prognostic equations following the mass balance equation for cloud water/ice content and cloud air. Source and sink terms related to cloud formation (e.g. condensation/sublimation, cumulus convection) and destruction (e.g. evaporation, precipitation) processes modify the cloud variables as time evolves in the simulation.

In order to evaluate the representation of clouds in ERA-Interim by means of satellite observations, a cloud simulator (described below) is needed to convert the model state into synthetic measurements. This diagnostic tool helps to understand and analyze what a satellite would see if the atmosphere had the clouds of a climate model. In other words, the simulator provides pseudo-satellite cloud products, which can be compared to satellite retrieval results so that differences can be interpreted as model errors. It is important to note that instrument simulators cannot entirely close the gap between models and observations because it is not possible to include everything about the observational process (e.g. satellite view angle effects on cloud detection, artifacts caused by partially cloudy satellite pixels). However, they are essential for carrying out robust inter-comparison studies.

4.2 Simplistic cloud simulator for ERA-Interim

The purpose of the SIMplistic cloud simulator For ERA-Interim (SIMFERA) developed in the framework of ESA Cloud_cci is to evaluate the cloud parameterization used in the IFS. In general SIMFERA consists of three modules: (1) downscaler, which converts the model grid box mean profiles into sub-grid profiles considering the mismatch in spatial scale between that of a model and that of a satellite pixel; (2) pseudo-retrieval, which emulates the pixel-scale cloud parameters based on the sub-grid profiles; and (3) statistical aggregation, which builds the diagnostic output that is comparable to the observational dataset (i.e. temporal averages and histograms, see below).

The general features are:

- SIMFERA uses the three-dimensional (3D) model fields as input (see details below). The simplistic approach in offline mode has the advantage of short computation time (e.g. 33 years of reanalysis data processed in less than 2 days on a HPC system).
- Unlike sophisticated simulators, which are using modelled radiances and brightness temperatures to retrieve cloud optical parameters based on radiative transfer calculations (e.g., COT and CER following Nakajima-King method), SIMFERA stays very close to the original model fields. For instance, it uses the ERA-Interim CER parameterization (Martin et al. 1994, Sun and Rikus 1999, Sun 2001) along with the original 3D variables to convert the model state into comparable synthetic observations. Details are given in Stengel et al. (2017c).
- No satellite overpass is taken into account as ERA-Interim is only available in discrete temporal resolution of several hours. However, day and night conditions are considered for

the calculation of cloud optical parameters (i.e. COT, CER, CWP) that are only available during daytime observations since they are based on visible measurements.

- SIMFERA provides 2 options about how liquid and ice clouds occurring in the same model grid box are treated during the simulations (in the sub-column procedure):
  - mixed phase (i.e. mixed phase clouds if both water/ice contents exists) or
  - no-mixed phase (i.e. considering liquid and ice clouds separately).
- SIMFERA can be used for other model output evaluation after small modifications since there are not instrument/algorithm specifications implemented.

Input:

- The simulator reads 6-hourly (00, 06, 12, 18 UTC) gridded estimates of 3D meteorological upper air parameters on 60 model levels including the following profiles:
  - liquid water content “LWC” [kg/kg],
  - ice water content “IWC” [kg/kg],
  - cloud cover “CC” (0-1),
  - temperature “T” [K], and
  - specific humidity “Q” [kg/kg].
- Additionally, the ERA-Interim file comprises for each grid box two-dimensional (2D) arrays of
  - surface geopotential “Z” [m^2/s^2] and
  - logarithm of surface pressure “LNSP” [Pa].

The latter two parameters are required for the computation of vertical pressure and geopotential profiles by using the provided “A” and “B” coefficients on model levels along with T and Q profiles.

Output:

Grid box monthly means are computed averaging first over all sub-columns per grid box and then averaging over all diagnostic time steps per month. Histograms are based on sub-column values because the downscaled results mimic the spatial resolution of a satellite footprint.

SIMFERA provides the following monthly mean products:

- total, high-, mid-, and low-level CFC (0-1),
- CPH (0-1),
- LWP and IWP [g/m^2],
- CTP [hPa], CTH [km], and CTT [K],
- COT and CER [micron] for liquid and ice phase,
- 2D joint cloud property histograms following the ISCCP classification relating the simulated height and optical thickness of the clouds, and
- 1D histograms for CTP, CTT, CWP, COT, and CER with the cloud phase as additional dimension.

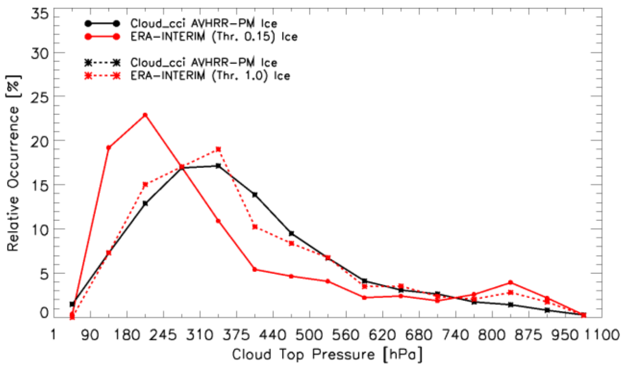
4.3 Results of reanalysis comparisons

In the following subsections the simulated and observed CFC, CPH, and CTP are inter-compared based on zonal monthly means (latitudinal distribution) and time series of monthly averages (temporal variation). The Cloud_cci AVHRR-PM time series are neither trend- nor orbital drift-corrected. All averages are latitude weighted means. The time period of all datasets spans 1982/01 to 2014/12 (33 years), whereby only daytime data been used.

In the case of CFC and CPH a detection threshold of COT > 0.15 was applied for identifying the uppermost cloud layer in the model profile. For CTP a detection threshold of COT > 1.00 was used. The different thresholds are based on validation scores for the Cloud_cci AVHRR-PM dataset against CALIOP on-board CALIPSO. For the binary cloud mask (clear, cloudy) the vertical placement of the cloud does not matter. In the case of a cloudy satellite pixel, the phase is either liquid or ice (binary CPH, since the retrieval is not able considering mixed-phase clouds). A passive imager like AVHRR is capable to detect clouds having an optical thickness larger than 0.15, especially for daytime conditions where more spectral information is available. Since during daytime VIS and IR spectral information is used together, cases of thin cirrus and underlying low clouds are misidentified as low-level clouds. Therefore the number of these cases is being reduced by a applying a threshold of COT > 1.

The investigation of one-dimensional (1D) histograms of CTP [hPa] based on detection threshold larger than 0.15 and 1.00, respectively, confirms the AVHRR validation study using lidar measurements. Figure 4-1 shows the 1D CTP histograms of ice clouds averaged over 33 years excluding the Polar Regions (i.e. ± 60 latitude) for Cloud_cci AVHRR-PM satellite retrievals and both SIMFERA runs (0.15, 1.00). The larger detection threshold applied in SIMFERA leads to a better agreement with observations. The 1D histogram for liquid clouds (not shown here) is very similar for these threshold values because the uppermost clouds are in general ice clouds.

For carrying out a fair inter-comparison between ERA-Interim reanalysis data and AVHRR-based observations, these different detection capabilities have to be addressed in the cloud simulator by applying the appropriate detection threshold values. Otherwise the assessment of the cloud parameterization in the model by means of space-borne data is not reasonable.



Tropics & Mid-latitudes: 1982 - 2014			
	low	mid	high
AVHRR-PM	6.9 %	23.42 %	69.67 %
SIMFERA 0.15	10.87 %	13.34 %	75.78 %
SIMFERA 1.00	9.20 %	22.15 %	68.65 %

Figure 4-1 *Left: One-dimensional (1D) histogram of cloud top pressure [hPa] for ice clouds averaged over 33 years (1982 - 2014) excluding the Polar Regions (± 60 latitude). The black line shows the Cloud_cci AVHRR-PM ice cloud distribution, while the red solid and dashed lines show the simulator results for COT-THV equal 0.15 and 1.00, respectively. Right: table showing the relative occurrences in numbers for low-, mid-, and high-level clouds based on the ISCCP cloud classification scheme.*

4.3.1 Cloud Fraction

Figure 4-2 compares the zonal mean CFC obtained from the cloud simulator applying two different COT thresholds (rose: 0.15, green: 1.00) with AVHRR-PM retrieval results from Cloud_cci (pale blue), PATMOS-x (dark cyan), and CLARA-A2 (olive). It is obvious that ERA-Interim underestimates CFC in the tropics and mid-latitudes, while overestimates in the Polar Regions, especially at the North Pole. It is a known problem that general climate models (GCMs) and reanalysis data underestimate the cloud amount, especially they provide too few mid-level clouds (Kay et al., 2012).

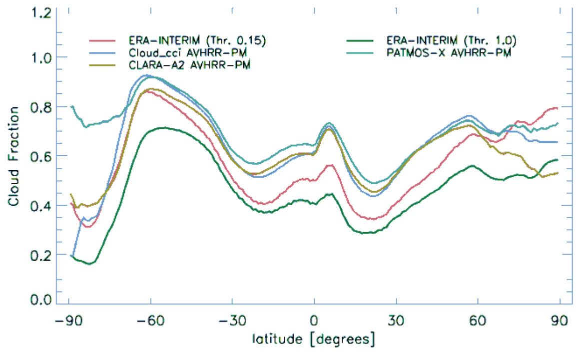


Figure 4-2 *Zonal mean cloud fraction derived from SIMFERA using different detection thresholds (COT-THV = 0.15 / 1.00) and three satellite-based cloud climatologies using AVHRR afternoon (PM) data: Cloud_cci, PATMOS-x, and CLARA-A2. The values are averaged over all monthly means from January 1982 to December 2014.*

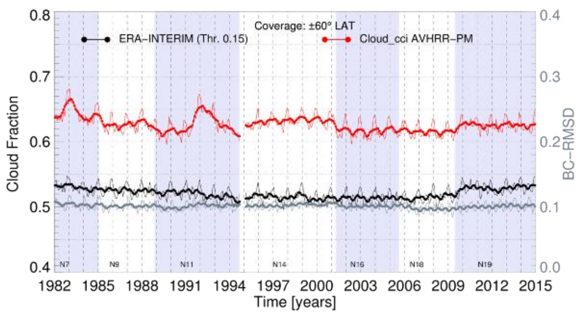


Figure 4-3 *Time series of monthly mean cloud fraction derived from SIMFERA (black; COT-THV=0.15) and Cloud_cci AVHRR-PM satellite retrievals (red) excluding the Polar Regions. Thin lines show monthly averages, while thick lines are running averages. The right y-axis (grey) shows the bias corrected root-mean-square difference (BD-RMSD) between the simulated and observed monthly means.*

In the Polar Regions the space-borne cloud climatologies significantly deviate from each other, with PATMOS-x providing the largest CFC compared to the others. That is why the inter-comparison of the CFC time series between Cloud_cci and SIMFERA (COT-THV > 0.15) is carried out for the mid-latitudes (± 60 lat.) only (see Figure 4-3). Moreover, the AVHRR-PM v2.0 data have some known issues in the Polar Regions, which are already solved for the upcoming v3.0 release. Additionally, at high latitudes the retrieval conditions are difficult for passive imagers due to spectral limitations. Like every model, a simulator is imperfect too and cannot entirely close the gap between simulation and observation and thus, tricky scenarios should be excluded for a robust inter-comparison.

Figure 4-3 also demonstrates that ERA-Interim underestimates CFC by about 10 % compared to AVHRR-PM satellite retrievals. The smaller variability in the simulated CFC is mainly due to the fact that SIMFERA does not account for satellite overpasses. Hence, the number of pseudo-satellite results used for building the diagnostic output is different compared to that used for the observational dataset.

4.3.2 Liquid Cloud Fraction

Figure 4-4 (left) shows the simulated and observed zonal mean CPH. SIMFERA results based on COT threshold of 0.15 agrees very well with the AVHRR-PM satellite retrievals in the tropics and sub-tropics. Beyond ± 40 degree latitude ERA-Interim has more ice clouds compared to the observations, which might be caused by the lack of cloud liquid water in the model, especially in the upper troposphere (Forbes et al., Winter 2015/2016). Pseudo-satellite zonal mean CPH applying a detection threshold of 1.00 leads to a better agreement with space-borne measurements in the high latitudes but worse agreement in tropics and sub-tropics because of "undetected" high and optical thinner ice clouds.

Figure 4-4 (right) inter-compares the time series of monthly mean CPH obtained from SIMFERA and AVHRR Cloud_cci AVHRR-PM satellite retrievals considering only the mid-latitudes. The simulated and observed CPH is in very good agreement for NOAA-16, NOAA-18, and NOAA-19, i.e. reduced orbital drift impact (see Figure 4-4). Between 1982 and 2001 the equatorial crossing time of the NOAA platforms changes significantly in value over time due to orbit degradation, which of course has an influence on the computation of monthly mean CPH. Overall, the global mean of simulated (0.52) CPH corresponds quite well with the observed (0.55) CPH.

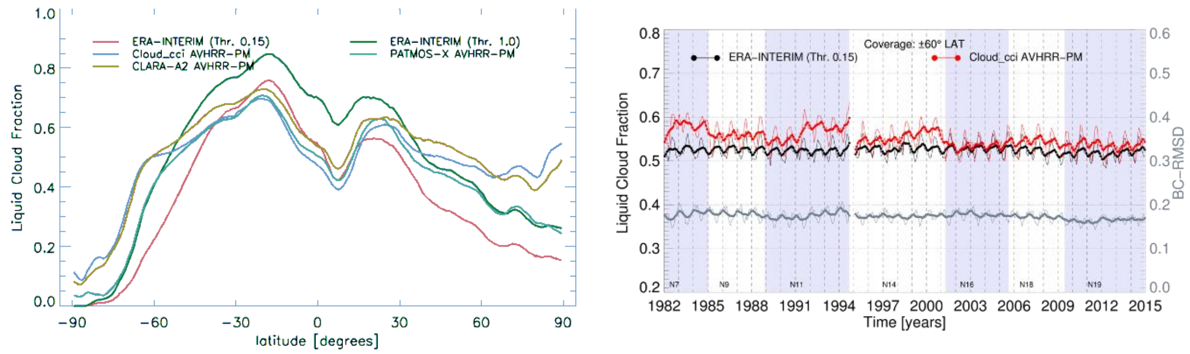


Figure 4-4 *left: Zonal mean cloud phase derived from SIMFERA and satellite-based cloud climatologies using AVHRR-PM data. Right: Time series of monthly mean cloud phase derived from SIMFERA and Cloud_cci AVHRR-PM satellite retrievals for mid-latitudes*

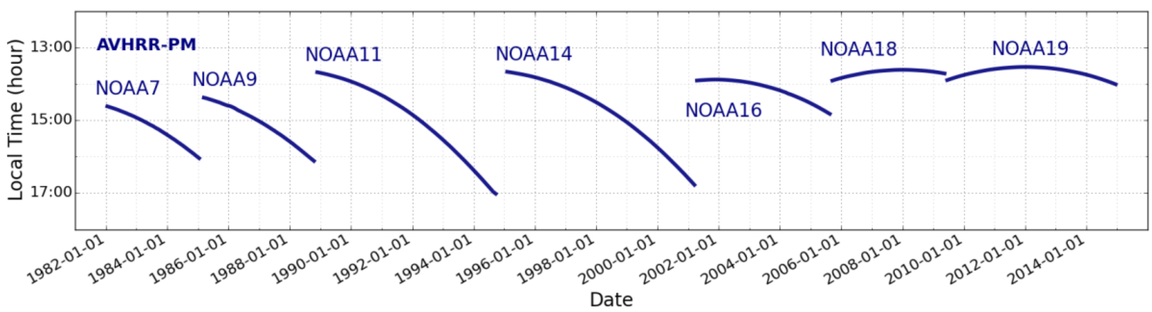


Figure 4-5 *Equatorial crossing times of the ascending nodes of NOAA afternoon (PM) satellites carrying an Advanced Very High Resolution Radiometer (AVHRR).*

4.3.3 Cloud Top Pressure

Figure 4-6 (left) presents the zonal mean CTP derived from the cloud simulator and well-known satellite climatologies. SIMFERA results based on COT threshold of 0.15 coincides quite well with observations; however, these should not be used for the model evaluation (see Figure 4-1 for explanation). The simulated CTP based on detection threshold 1.00 points out that the clouds in ERA-Interim are lower compared to satellite measurements. Figure 4-6 (right) inter-compares the time series of monthly mean CTP derived from SIMFERA and Cloud_cci AVHRR-PM for the mid-latitudes. The global mean standard deviation between the simulated and observed CTP is 92.4 hPa indicating that in the model the clouds are on averaged lower than what satellites observe from space.

From mean values of CTP it is very difficult to conclude on an evaluation. As the histograms in Fig. 2.13 have shown, the difference might also originate from the fact that CTP of low-level clouds in the Cloud_cci datasets is slightly overestimated.

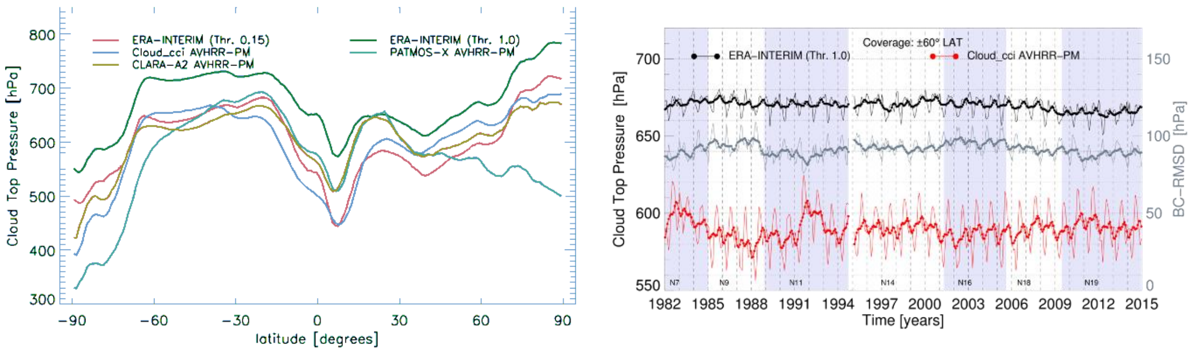


Figure 4-6 *left: Zonal mean cloud top pressure [hPa] derived from SIMFERA satellite-based cloud climatologies using AVHRR-PM data. Right: Time series of monthly mean cloud top pressure [hPa] derived from SIMFERA and Cloud_cci for mid-latitudes.*

5. Comparisons with climate models

Clouds strongly affect the Earth's radiative balance and temperature but are difficult to accurately model and observe, leading to major uncertainties in understanding climate variability and change. Further model development and evaluation using long consistent observational data records can increase both the understanding of the present climate and the confidence in climate model scenarios. Modelled and observed clouds are different, as on one hand the observed clouds are determined by the satellite instrument sensitivity, the temporal and spatial sampling and the vertical overlap of the cloud layers, while climate model clouds are assumed to be plane-parallel and are of coarse horizontal and vertical resolution. Ideally, a satellite simulator should be used (e.g. Bodas-Salcedo et al., 2011) on the model data to mimic the satellite view and sampling.

However, since many CMIP5 climate models historical and scenario simulations have been run without a simulator, cloud observations are directly compared with models. Total cloud cover is the model cloud parameter that most readily can be compared directly to the satellite derived cloud fraction without a simulator, even though models can have substantial cloud cover but very little cloud condensate making those clouds too optically thin to be detected by the satellite instrument.

Therefore, in this section Cloud_cci data is compared with regional and global climate models with and without simulators. Comparisons are presented using simulators for a kilometre-scale forecast model in section 5.1, a regional climate model in section 5.2, a global climate model in section 5.3, and finally direct comparisons with CMIP5 models in section 5.4.

5.1 Regional case study: kilometre-scale

In this sub section a comparison is presented of Cloud_cci level 3 data (L3U) with a kilometer-scale weather forecast model during a convective period.

5.1.1 The regional weather forecast and climate model COSMO

The non-hydrostatic model COSMO is operationally used for weather forecast at MeteoSwiss and other national weather services, mainly in Europe. In climate mode and referred to as COSMO-CLM, the model has been used to contribute to the CORDEX project (e.g. Kotlarski et al. 2014).

In this study, COSMO is used with Version 4.25, in climate mode, and at kilometre-scale resolution (Baldauf et al. 2011). The setup follows previous studies (e.g. Ban et al., 2014; Keller et al., 2016) and convection-resolving simulations in numerical weather prediction mode at MeteoSwiss (e.g. Weusthoff et al., 2010). A radiative transfer scheme based on the δ-two-stream approach (Ritter and Geleyn, 1992), a convection scheme for shallow and deep convection (Tiedtke, 1989), or alternatively a convection scheme only for shallow convection (Tiedtke, 1989; Theunert and Seifert, 2006), a one-moment bulk cloud-microphysics parameterization (1M) (Reinhardt and Seifert, 2006), or alternatively a two-moment bulk cloud-microphysics parameterization (2M), which includes ice sedimentation, (Seifert and Beheng, 2006) are used.

5.1.2 The COSMO simulator

The cloud simulator, which was used in Keller et al. (2016) and for the CAR, was built to compare ESA Cloud_cci L3U data (cloud optical thickness (COT) and cloud top pressure (CTP)) with the COSMO model when no official cloud simulator for this data was available. The calculation of COT is implemented in the radiation code of the COSMO model (Ritter and Geleyn, 1992). It is calculated from direct solar radiation at ToA (I0) and at certain atmospheric layers (IL) with COTL = − ln(IL/I0). COT over the full atmospheric column is calculated as COT = COT lowest level. Between ToA and the atmospheric layers, direct solar radiation is reduced considering absorption and scattering due to cloud water and cloud ice at 0.24 µm to 0.7 µm. Water vapor, aerosols, or Rayleigh scattering are not considered, which differs from the standard radiation calculations in the code. Radiation through subgrid clouds is calculated with the maximum random overlap assumption (Geleyn and Hollingsworth, 1979). Compared to COSP, one major difference is that for this cloud simulator no sub-columns are used.

Since COT is calculated inside the simulations, input for COT is taken directly from there. CTP is calculated with an NCL (NCAR command language) script, which needs COT and pressure at every level as input. CTP is defined at the pressure level where COTL =0.3. While calculating COT does not lead to a considerably longer run time of the model, the script to calculate CTP and the COT/CTP histograms needs several minutes per data input (for 500x500x60 grid points). If this cloud simulator would be used further in the future, it would be worth to also implement the CTP calculation into the radiation code of the model or to improve the performance of the script. But this cloud simulator is intended to be replaced by the official cloud simulator. A more detailed explanation on how COT is calculated can be found in Appendix B of Keller (2016).

5.1.3 Experimental framework

Three simulations are undertaken. The first has 12 km spatial resolution, a parameterization for deep convection, uses 1M and is called 12km1M. The other two simulations have 2.2 km resolution, a parameterization only for shallow convection, use 1M or 2M (see 5.1.1), and are called 2km1M or 2km2M, respectively. 12km1M is driven by ERA interim, and the 2 km simulations by 12km1M.

The level 3 data (L3U) from the MODIS instrument on the Aqua platform is used from the Cloud_cci project. The data is masked to be equal to the analysis domain (Figure 5-1). For the observations, clouds are excluded which have a cloud mask uncertainty larger than 35%. The uncertainty threshold is chosen at 35 % where a minimum is found between two maxima in the cloud mask uncertainty histogram over Europe for June 2007. The excluded clouds are discussed at the end of the results. For the model, CTP and COT are calculated at 13 UTC to match the overpass time of 1:30 PM local time of the Aqua satellite. The histograms are calculated and normalized over all data points (including cloud-free points). Comparisons are carried out for 11 days in June 2007, a period with a strong diurnal cycle of deep convection, to highlight differences between simulations with parameterized (12km1M) and explicit (2km1M, 2km2M) deep convection.

5.1.4 Results for regional case studies on kilometre scale

In Figure 5-1 CTP is shown over the model domain at 5 June 2007. Compared to observations, 12km1M has a large overestimation of high clouds (< 400 hPa), which is improved with 2km1M and further with 2km2M. This is confirmed by the COT/CTP histograms over all 11 days.

The improvement in high clouds from 12km1M to 2km1M shows the added value of the higher resolution without parameterization for deep convection, during a period with frequent deep convection. The further reduction of high clouds with 2km2M demonstrates the added value of ice sedimentation in association with the two-moment cloud-microphysics. Despite the differences in clouds, the impact of the two-moment cloud-microphysics on area averaged accumulated precipitation (not shown) is small with less than 1 mm for this period.

If including clouds with cloud mask uncertainties > 35%, mainly low clouds with CTP > 800 hPa appear (not shown). A lot of these clouds are isolated (surrounded by cloud-free areas), which indicates that these clouds are shallow cumulus clouds. Further investigation of the model grid points, which cause the maxima in the COT/CTP histograms of the 2 km simulations, show that the model simulates shallow cumulus clouds as parameterized subgrid clouds. These subgrid clouds are treated as very thin grid-scale clouds in the cloud simulator instead of as partly cloudy and partly cloud-free. Therefore, in the observations and also in the cloud simulator, small low-level clouds, which represent probably shallow cumulus clouds, are treated with some uncertainty and therefore their comparison is difficult.

5.1.5 Conclusions for regional case studies on kilometre scale

The use of the Cloud_cci data to evaluate the model demonstrates the importance of increasing the spatial resolution to avoid parameterization of deep convection. Further, the possibility to consider uncertainty (e.g. for cloud mask) helps to decide which cloud types are important for model evaluation and which should be treated with care.

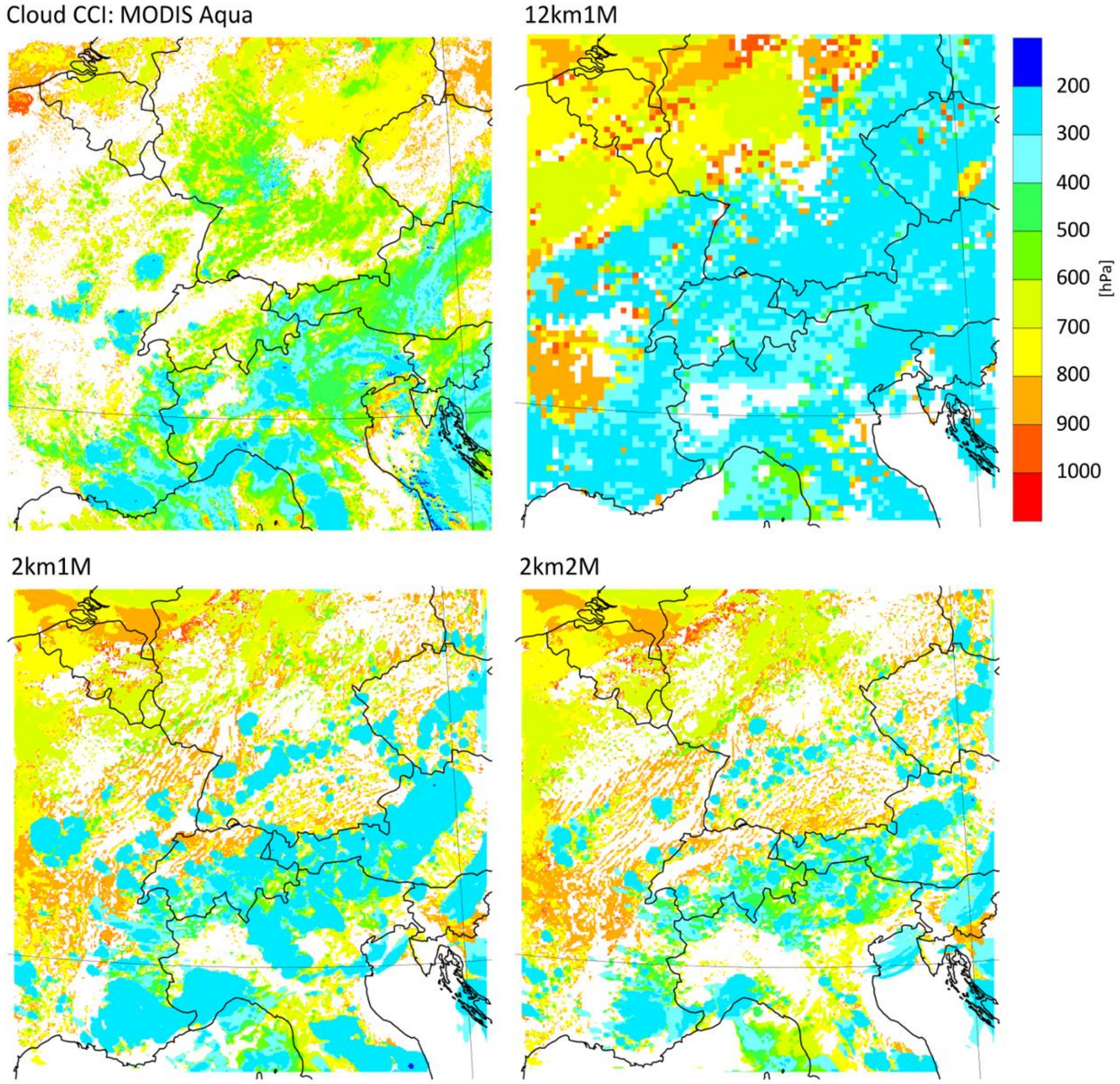


Figure 5-1 *Cloud top pressure on 5 June 2007 from the ascend node of Aqua and at 13 UTC in the simulations 12km1M, 2km1M, and 2km2M over Central Europe.*

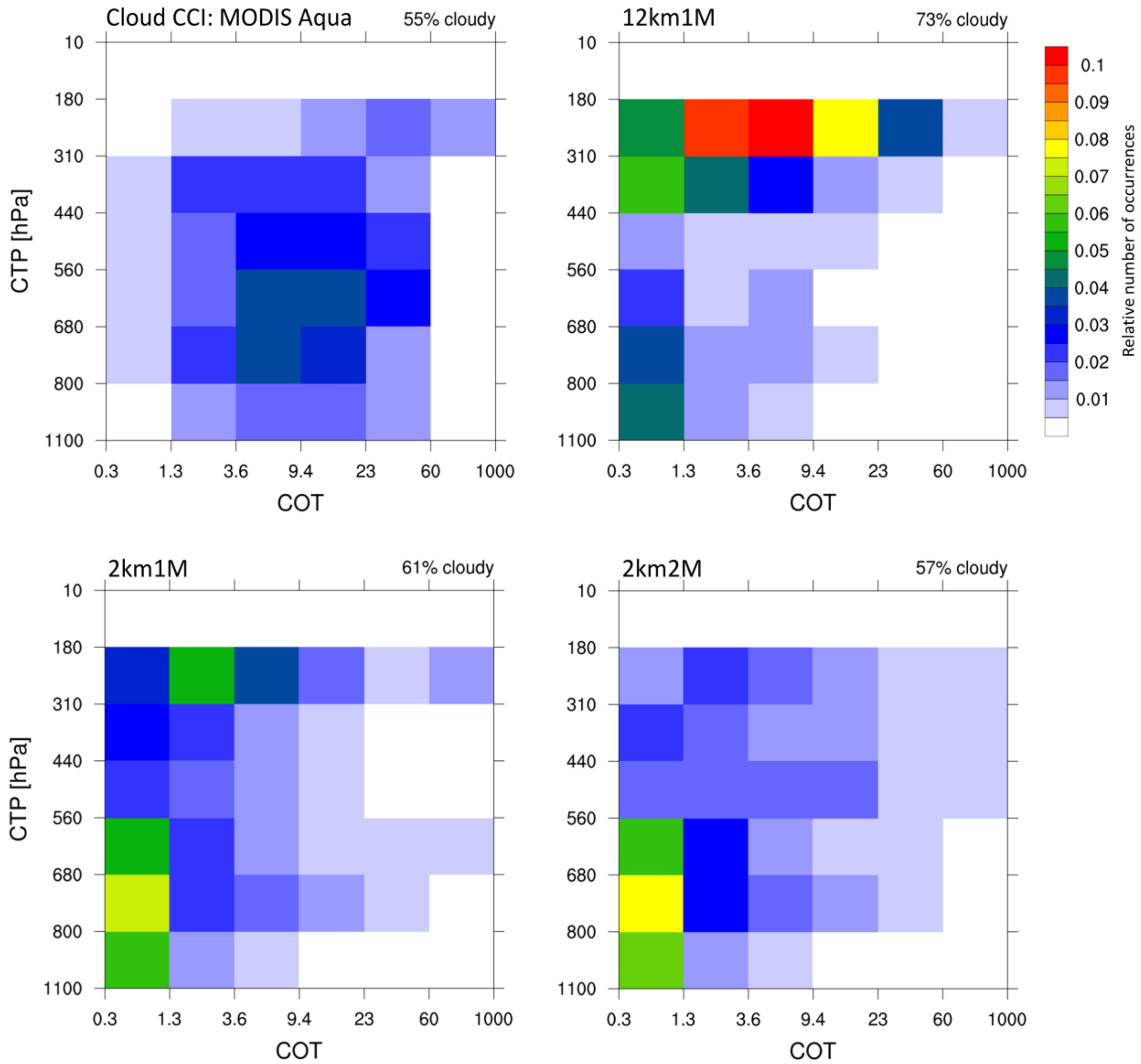


Figure 5-2 *Two-dimensional histograms of cloud optical thickness (COT) and cloud top pressure (CTP) as averages over the period of 3 to 13 June 2007. For observations, the average is over local time of the ascending node of the Aqua satellite (approx. 13:30 PM), for the simulations, it is at 13 UTC. Fractional cloud cover (defined by COT > 0.3) is indicated in the right upper corner of all panels.*

5.2 Regional case studies

In this sub section simulations applied to model output from a regional climate model operated at 25 km horizontal resolution are compared with Cloud_cci level 3 data (L3U).

5.2.1 The regional climate model KNMI-RACMO

The hydrostatic regional climate model (RCM) RACMO2 is originally built from the parameterization package of physical processes employed in the ECMWF physics merged with the dynamical kernel of the High Resolution Limited Area Model (HIRLAM; Undén et al., 2002) NWP. RACMO2.3 used in this study is based on ECMWF cycle 33r2 (ECMWF-IFS, 2009), which constitutes large overlaps with the ECMWF physics of cy36r1 used in EC-EARTH3.2. Its predecessor RACMO2.2 (van Meijgaard et al., 2012) was to a large extent based on ECMWF physics of cy31r2, also used in the ERA-Interim reanalysis (Dee et al., 2011) and EC-EARTH2.3 (based on Hazeleger et al., 2012). Differences between RACMO2.3 and 2.2 primarily constitute the substitution of the short wave radiation scheme by a rapid radiation transfer module (RRTM; short wave radiation Clough et al., 2005; long wave radiation, Mlawer et al., 1997) and the introduction of the independent column approximation (ICA; Morcrette et al., 2007) to better represent the effect of clouds on radiation at the sub-grid scale. Other physics components include a TKE driven eddy-diffusivity mass-flux scheme (Siebesma et al., 2007; Lenderink and Holtslag, 2004) for mixing and cloud processes in the boundary layer, a scheme for deep convection (originated by Tiedtke, 1989), a prognostic cloud scheme (Tiedtke, 1993; Tompkins et al., 2007), the land surface/soil scheme HTESSEL (Balsano et al., 2009).

In the past five years, RACMO2.2 has been frequently used in transient climate simulations with different GCM drivers (EC-EARTH, HadGEM2-ES) in the framework of CORDEX (Giorgi et al., 2009) (European scale at 50 and 12km resolution) and in preparation of the KNMI'14 climate scenarios for The Netherlands (van den Hurk et al., 2014).

5.2.2 Experimental setup

Integrations with RACMO in the framework of this study are carried out at 25km horizontal resolution and with 40 model levels using a hybrid vertical coordinate. The model domain, configured by employing a rotated pole coordinate, roughly encompasses the region between 30°W and 50° E, and between 30° and 70° N (see ***Figure 5-1***). The integrations are driven by ERA-Interim atmospheric fields at the lateral boundaries and temperature and sea-ice extent at the sea surface. Integrations are carried out in hindcast mode (instead of climate mode) implying short-term runs of 36 hours. In the chosen set up runs are re-initialized daily at 12 UTC from ERA-Interim including the land surface state. To avoid effects from spin up, in particular in the cloud fields, the first 12 hours of each run are not considered in the further processing. In this way a quasi-continuous long-term time series of non-overlapping model output is constructed. The benefit of employing hindcast mode rather than climate mode is that the model atmospheric state stays close to the quasi-observed large-scale flow imposed by ERA-Interim, facilitating the comparison of the simulated cloud parameters with those inferred from satellite observations. Multi-level output, including cloud parameters like instantaneous cloud fraction, cloud liquid water content and cloud ice content is archived in 3-hourly resolved files.

The direct (or native) model output is subsequently used to drive the cloud simulator. Here the Cloud_cci simulator (version 1.3) is utilizied which has been specifically developed in the framework of this project (as described in the next sub section 5.3.2). The purpose of applying the cloud simulator is to mimic the response of the AVHRR instrument on board the respective NOAA satellites in viewing a model atmospheric state instead of a real atmosphere. Settings are chosen such that the simulated response is collocating and synchronous to ascending AVHRR nodes, allowing a once-daily direct comparison with retrievals of observed cloud parameters.

Prior to comparison, the L3U data retrieved from the NOAA-AVHRR measurements and composited from all ascending nodes in daily images on a nominal 5 km spatial resolution have been aggregated to 25 km resolution in order to match the model resolution. Observed ice cloud fraction and liquid water cloud fraction is calculated from the number of 5 km pixels indicating ice and liquid water clouds, respectively, in the designated 25 km cell. The aggregation of remaining cloud parameters (ice/liquid water path; ice/liquid water effective particle size, ice/liquid water cloud top height, etc.) is carried out accordingly. All cloud parameters are defined as all-sky (rather than cloudy-sky) averages.

5.2.3 Qualitative comparison for a single month

Here first results are presented regarding the comparison of direct model output, simulated model output, and retrievals. It is focused on a single month, i.e. September 1986, when the AVHRR on NOAA-9 was active.

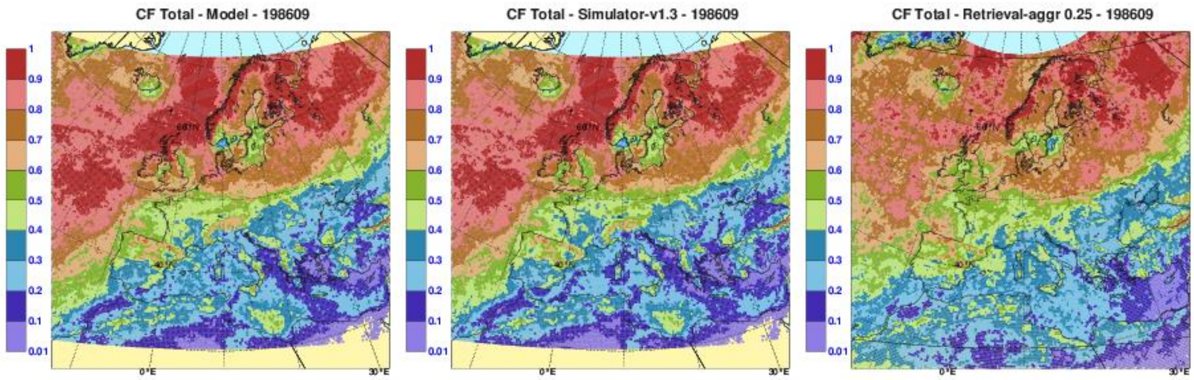


Figure 5-3 *Total cloud fraction: native model (left), simulated (middle), retrieval (right) for September 1986. The (faintly visible) square rectangle in the right panel indicates the limits of the modelling domain.*

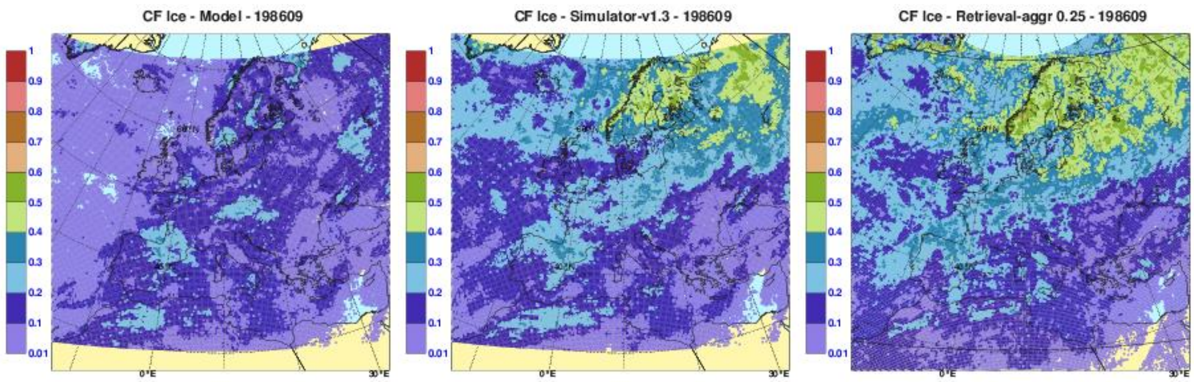


Figure 5-4 *Like* **Figure 5-3**, *but for ice cloud fraction.*

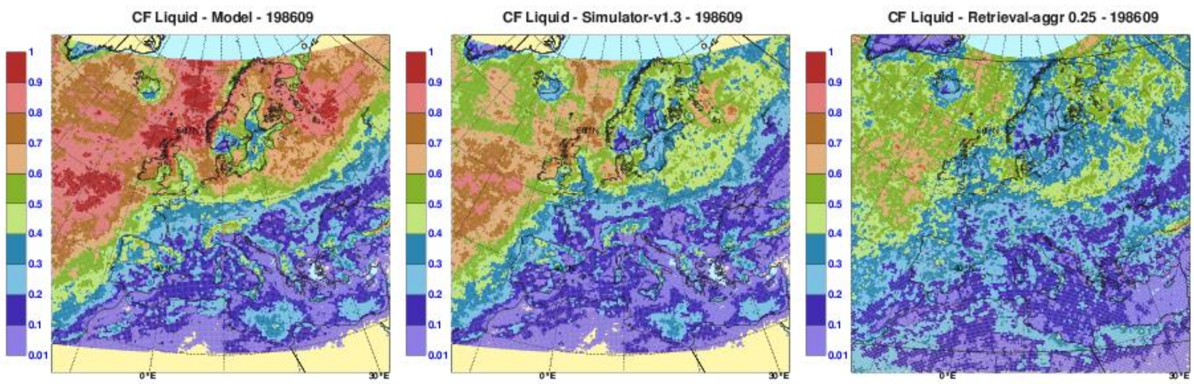


Figure 5-5 *As Figure 5-3 but for liquid water cloud fraction.*

The Figure 5-3 to Figure 5-5 show the monthly mean total cloud fraction (Figure 5-3), the monthly mean ice cloud fraction (Figure 5-4), and the monthly mean liquid water cloud fraction (Figure *5-5*) derived from direct model output (left), after applying the cloud simulator (centre), and the retrieval (right). For total cloud fraction, the simulator and direct model output appear very similar. In fact the simulated amount is slightly smaller because the simulator removes all model total cloud amounts representing clouds with an optical thickness less than 0.2. The overall spatial structure in simulated total cloud fraction resembles that of the retrieval, yet there is a tendency in the model to overestimate total cloud fraction over the Atlantic Ocean. On the other hand, for the separate cloud phase applying the cloud simulator results in a distinct increase of ice cloud amount at the expense of liquid water cloud amount. The reason is that while mixed-phase profiles are abundantly present in the direct model output, the simulator labels many of these clouds as pure ice clouds once the phase at cloud top is identified as ice. For example, after sub-column decomposition of a mixed-phase cloud profile topped by a unity fraction ice cloud layer with optical thickness larger than the threshold, the clouds in all sub-columns will be interpreted as ice clouds.

Evidently, the simulator has a substantial and beneficial effect on the cloud fractions derived from the direct model output for both phases, yet the retrieval indicates that ice water cloud amount is still underestimated by the model, while liquid water cloud amount is still overestimated, the latter in particular over the ocean.

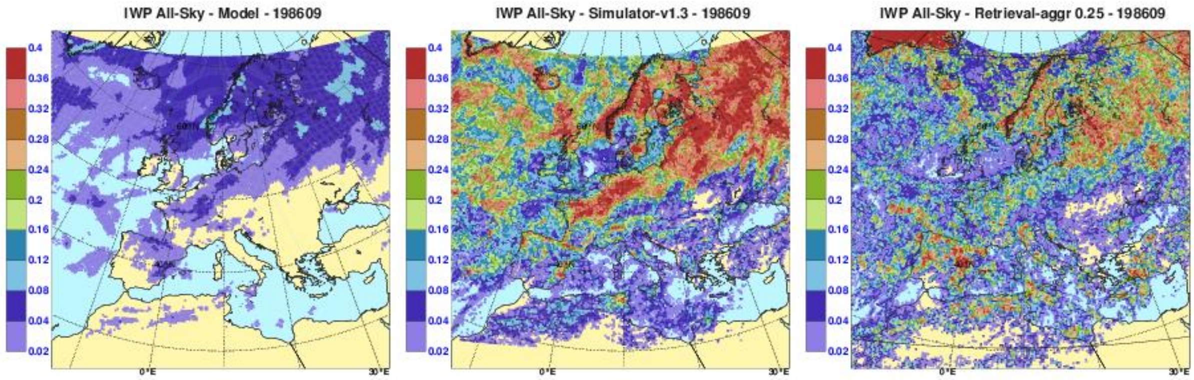


Figure 5-6 *As Figure 5-1 but for all-sky ice water path.*

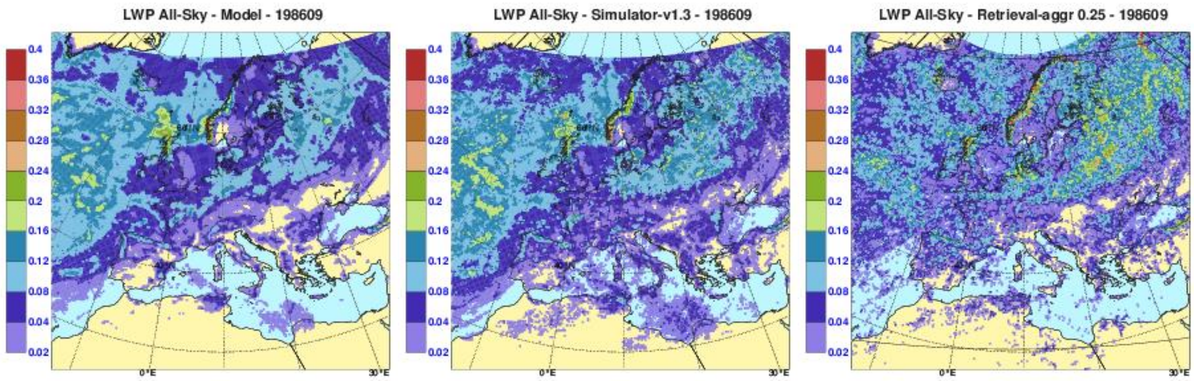


Figure 5-7 *As Figure 5-1 but for all-sky liquid water path.*

The Figure 5-4 to Figure 5-5 show ice water path (Figure 5-4) and liquid water path in the same sequence as for the cloud fraction parameters. For ice water path the simulated amounts substantially overestimate the retrieved amounts in a large part of the domain, while the direct model ice water path is much smaller than the retrieved amounts, consistent with the ice cloud fraction. There are two reasons for the high simulated ice water path values: (i) the large simulated ice cloud fraction as discussed before, and (ii) the relatively larger simulated particle effective radius, which resembles the (ice) particle effective radius near the top of the clouds, compared to the model effective radius, which is on average closer to that of the liquid phase often occurring in the lower part of the clouds. In contrast to ice water path, simulated and native model liquid water path values are very similar, which is a compensating effect between simulated liquid water cloud fraction being smaller than native model output and simulated cloud droplet effective radius being larger than the native model value (not shown). The simulated liquid water path tends to be larger than the retrieved values over ocean, but smaller over land.

5.2.4 Qualitative comparison for a single day

As mentioned before, owing to the experimental setup the large-scale flow in the model integrations is very close to the observed state which allows a direct comparison between simulated cloud fields and retrieved cloud fields on a daily basis. Bearing in mind one image is a daily composite of a sequence of overpasses, it is interesting to see how well large-scale weather patterns, in particular low pressure systems and associated frontal bands can be recognized in the images in Figure 5-8:

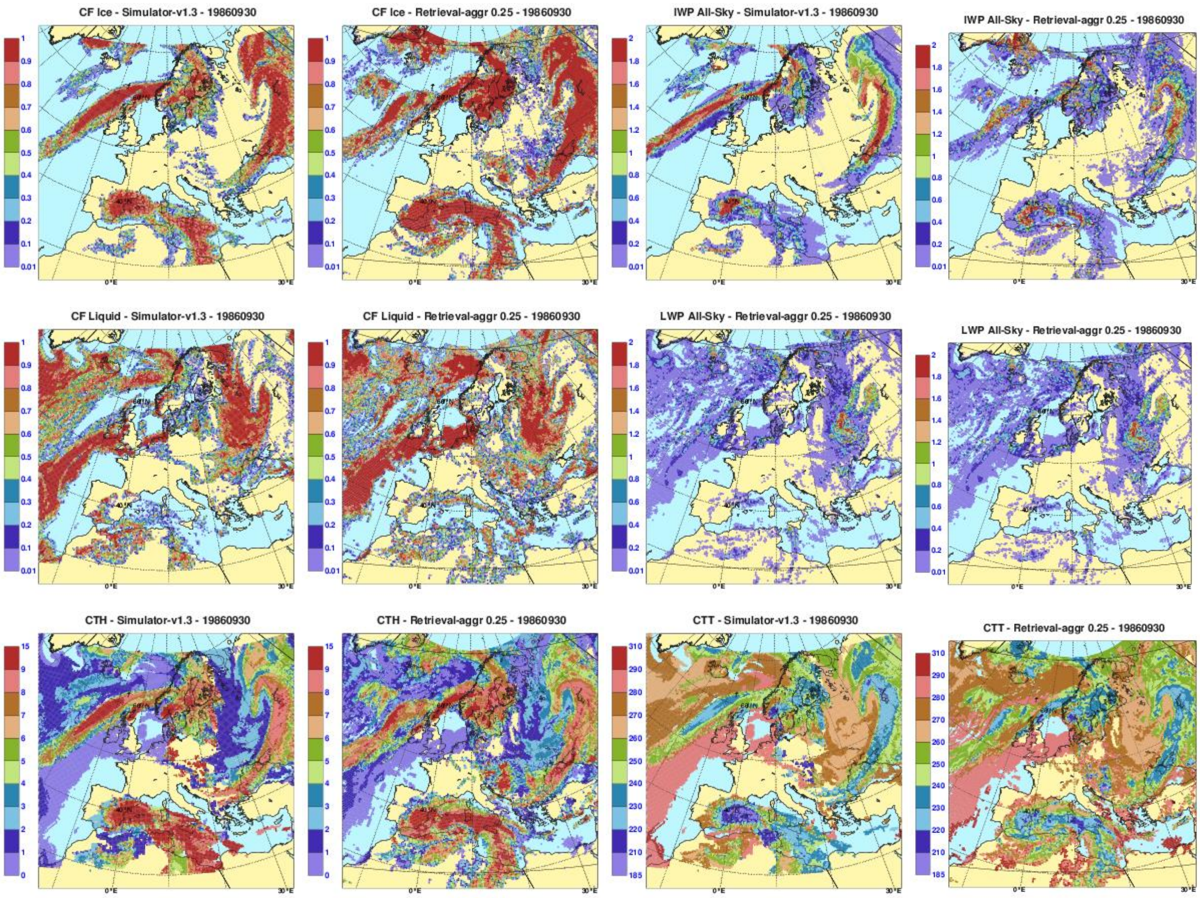


Figure 5-8 *Top row: simulated and retrieved ice cloud fraction, and simulated and retrieved ice water path (from left to tight), all derived for a single day: 30 September 1986. Middle row: same as upper row but for liquid water. Bottom row: simulated and retrieved cloud top height (left two panels), and simulated and retrieved cloud top temperature (right two panels). Retrieved parameters are inferred from AVHRR-NOAA9 measurements and aggregated to 0.25 degree.*

In particular, the cyclonic disturbance in the western part of the Mediterranean with abundant ice clouds according to the retrievals is nicely captured by the model fields processed with the cloud simulator, although the model produces higher and colder cloud tops in this region and a higher ice water path than the retrieval indicates. Another feature worth mentioning is the presence of a large region in the south-western part of the model domain with warm boundary-layer clouds according to the model which is only seen to a limited extent in the retrieved fields. As a more general remark it is noted that the retrieved fields look somewhat less coherent than the simulated fields.

5.2.5 Conclusion and outlook for regional case studies

The Cloud_cci simulator has successfully been applied to 25 km resolved multi-level cloud fields obtained with integration in hindcast mode of the regional climate model RACMO2. Simulated cloud fields have been compared with CC4CL L3U-based retrievals, aggregated at 25 km resolution. Simulated fields are consistently (much) closer to the retrievals than the original model fields, the exception being the ice water path field. It is shown that driving the RCM with information from a reanalysis like ERA-Interim, and integrating the model in hindcast mode allows a one-to-one comparison on a daily basis.

This work is still in progress and a quantitative analysis of a comparison for a longer period (3-5 years) is ongoing.

5.3 Global climate studies with the Cloud_cci simulator

In this sub section Cloud_cci L3 data is compared with simulations using the Cloud_cci simulator applied to model output from a global climate model operated at 125km and 60km horizontal resolutions.

5.3.1 The EC-Earth model and experimental set-up

The climate model used in this study is the global coupled climate model EC-Earth. The Integrated Forecast System (IFS) of the European Centre for Medium Range Weather Forecasts (ECMWF) constitutes the atmosphere component, which also includes the HTESSEL land-surface model (Balsamo et al., 2009; Boussetta et al., 2013). Here, two model versions are used, EC-Earth2.3 (Hazeleger et al., 2012) and EC-Earth3.2beta (Bousetta et al 2016). Version 2.3 is based on IFS cycle 31r1, but also includes some improvements from later cycles. The most important improvements are the convection scheme by Bechtold et al. (2008), the land surface scheme H-TESSEL (Balsamo et al. 2009), and a new snow scheme (Dutra et al. 2010). EC-Earth 3.2beta is based on CY36R4 release of the IFS in addition to few important physical adjustments mainly related to the non-orographic gravity wave drag parameterization (latitudinal and resolution dependency of the launch momentum flux) and the diurnal cycle of convection (Bechtold et al. 2014).

To reproduce present day climate variability the model was run with prescribed observed SST and Sea-Ice using ERA-Interim monthly data as the lower boundary condition. The two model versions were run for the time period 1979-2015 with 125km horizontal resolution and 60 vertical levels and 60km with 91 vertical levels, respectively. The simulator will be installed in the model but for these experiments the full model vertical fields were stored and the model parametrisation formula for Reff and overlap were used to reproduce the in-model variables needed for the simulator as described in the next sub-section 5.3.2.

5.3.2 The Cloud_cci simulator

The Cloud_cci simulator emulates the Cloud_cci observational cloud dataset using atmospheric variables from model output. One main component of the satellite simulator is to create sub-columns in order to make smaller grids, more comparable with the satellite measurement surface “footprint”, and to ensure the models cloud properties can be properly represented according to the models’ cloud overlap assumptions. Neglecting this step and only mapping the grid-average cloud properties to the sub column causes a miss-representation of the model atmosphere, and make it very hard to take the satellite cloud sensitivity into account in a meaningful way. A “cloud sensitivity”-filter is applied to the sub-columns so that they are comparable to observations. This is done by treating all sub-columns that have an optical depth less the 0.2 as cloud free. For consistency, the simulator uses the same lookup tables used in CC4CL to translate layer effective radius, optical depth, and solar zenith angle to single scattering albedo, effective radius and water path.

The simulator takes into account the solar zenith angle to ensure that retrievals only during daytime conditions are simulated. The simulator samples the model atmosphere to be consistent with the local measurement time of the observations. Since each satellite has different equatorial overpass times and most “drift” over time (in terms of equatorial overpass time), the simulator makes use of look up tables of monthly average overpass times for each satellite to ensure a temporal sampling consistent with observations. The simulator calculates a variable number of sub-columns, which decrease in number as a function of absolute latitude, in order to keep the “footprint area” consistent with observations. The time the simulator takes depends heavily on the settings and the resolution of the model. Important factors are

• The model resolution

• The satellite overpass time, which affects the number if illuminated grids cells to work on (note1)

• The assumed footprint size, which directly affects the number of sub-columns calculated per model grid.

• Whether or not to interpolate sampled model data so that longitude has exactly the same equatorial overpass time as that of the satellite. The interpolating between time steps adds up to 10% extra computational time, but interpolation will not be needed if the simulator is integrated into the model via COSP.

Table 5-1 shows how long it takes to simulate one day of model data (4 time steps), timed using Fortran’s CPU TIME(), which does not include IO.

The Cloud_cci simulator produces daily NETCDF files containing the following output on the native model resolution:

• cloud fraction, (ice, liquid, all)

• optical depth (ice, liquid, all)

• cloud albedo

• cloud top height, pressure, temperature

• corrected cloud top height, pressure, temperature

• cloud effective radius (ice, liquid)

• cloud water path (ice, liquid)

• cloud top pressure–cloud optical thickness 2D histograms

Notes:

1: For example in 2007, noaa18 has an EOT around 13:30, and noaa17 has an EOT of around 17:30, and therefore noaa18 will take considerably longer to run in this year

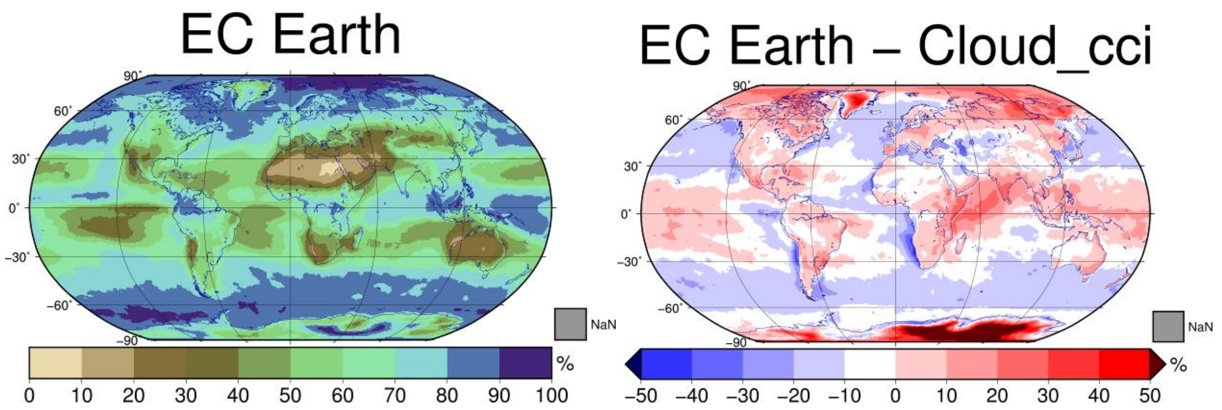
2: These times are dependent on the computer environment and may differ if run in other environments.

***Table 5-1**: CPU time spent simulating 4 time steps.*

Model resolution	Footprint (#grids at equator)	Time [s]
320x160x61	25 km (25)	14
320x160x61	5 km (617)	176
512x256x91	25 km (10)	44
512x256x91	5 km (242)	342

5.3.3 Results for global climate studies with the cloud simulator

To demonstrate the simulator, EC Earth-based Cloud_cci simulations are compared to the Cloud_cci dataset. Figure 5-9 shows the impact the simulators have on the comparison to observations. Based on data from 1982, the top row in this figure compares the model cloud cover (TCC) directly to the observations without any temporal sampling or steps to take satellite measurement sensitivity to clouds into account. The second row shows when the model is sampled so that the local time in the model matches the satellite equatorial overpass time. The third row shows the comparison of EC Earth to Cloud_cci using the full Cloud_cci simulator. Comparing row 1 to row 3 it is clear that the simulated cloud fraction is much closer to the observations compared to directly comparing the 'native' model cloud fraction to observations. The difference between row 1 and 2 demonstrates that sampling the model by local time consistent with observations, has a strong impact too. In fact, judging from these results, about half of difference between the simulated cloud fraction and the total cloud cover directly from the model, comes from correctly temporal sampling alone.



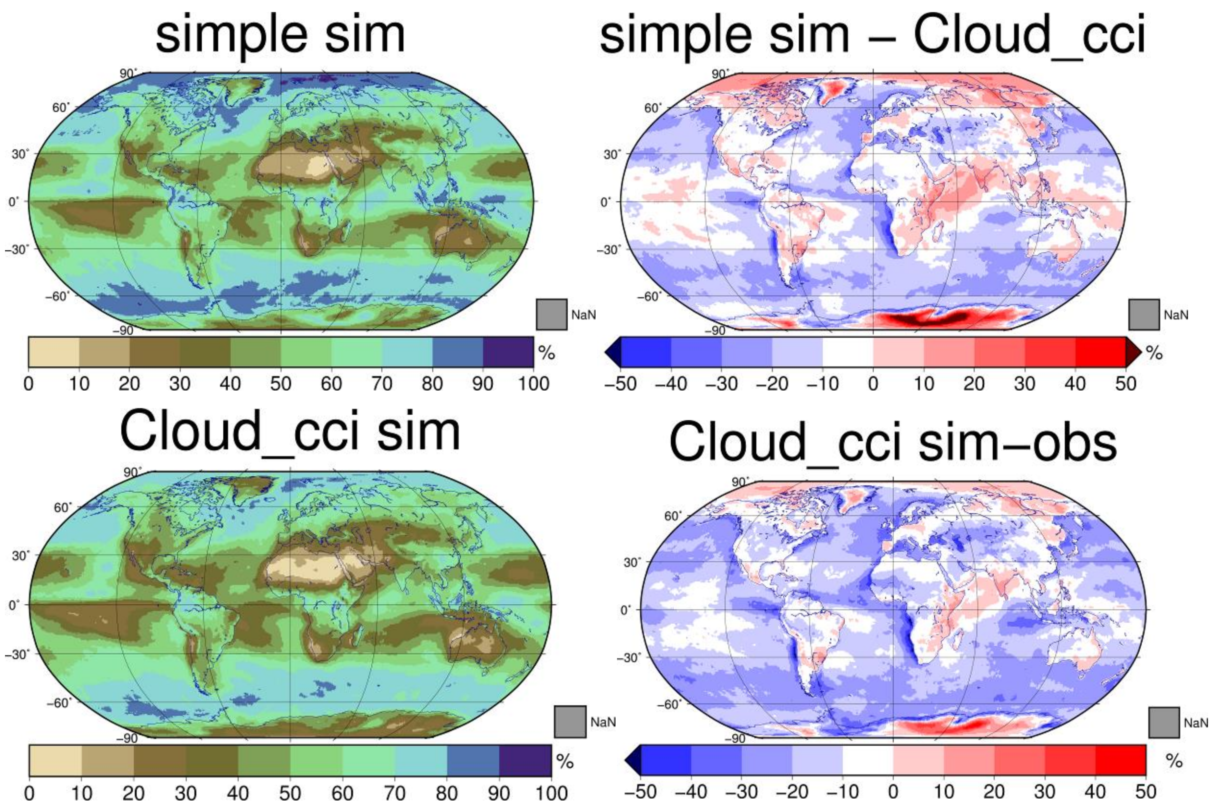


Figure 5-9 *Cloud fraction compared to observations. The top row shows the cloud cover directly from the model (TCC) (left) and model TCC minus Cloud_cci cloud cover (right). The second row shows the same as above, but using model data sampled to match the satellite overpass time. The third row shows the full Cloud_cci simulator (left) and the difference to observations (right). The equatorial overpass time of the satellite using in 1982, NOAA-7, is close to 15:00 local time. The values are the average cloud fraction in 1982.*

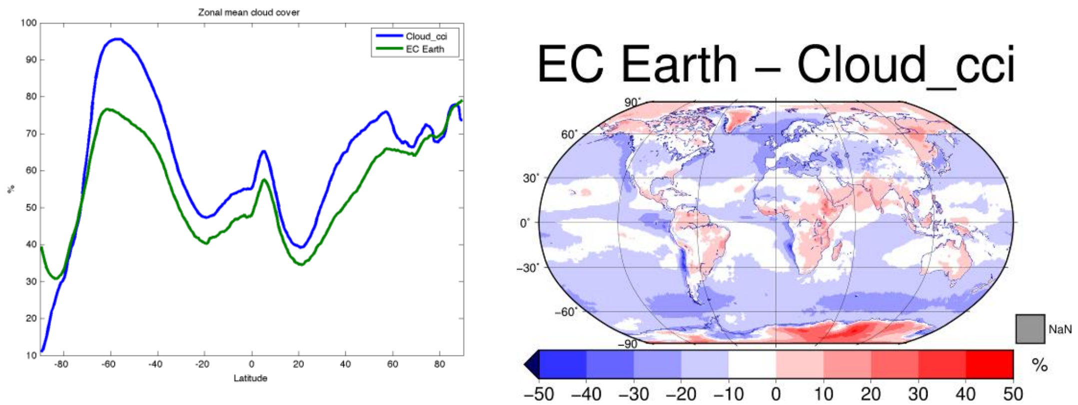


Figure 5-10 *Left: Zonal mean cloud fraction derived from the Cloud_cci simulator based on an EC Earth model atmosphere and the Cloud_cci climatology using AVHRR afternoon (PM) data. The values are the average retrievals from 1982–2012. Right: The spatial distribution of the difference in cloud fraction between EC Earth and Cloud_cci. The values are the average retrievals from 1982-2012.*

Figure 5-10 (left) shows using zonal averages that EC Earth has consistently less cloud cover than indicated by observations, although the general patterns appear quite similar. However, the spatial distribution of the difference in cloud amount shown in Figure 5-10 (right part) tells more of a story, for instance that EC Earth appears to have considerably more clouds over continental land, whilst the reverse is true over Ocean.

In terms of cloud effective radius, the differences are substantial. Figure 5-11 shows that there is a general and large offset between EC Earth and Cloud_cci in terms of effective radius for ice clouds, and disturbingly a somewhat opposite behaviour in some regions. For instance the Sahara is a local minimum in the observations, whilst the opposite relationship exists at high latitudes. These results are preliminary and need further investigation into the causes, and fidelity of the results. Figure 5-11 shows the effective radius of liquid clouds. Here the differences are also large, and especially due to the very distinct land-sea relationship in the model effective radius. This difference is probably explained by the parameterization used for the model effective radius which depends heavily on the land sea mask. However, the differences can also be due to problems for Cloud_cci retrieved effective radius during daytime, since thin cirrus clouds overlying low-level clouds could be misidentified as low-level cloud.

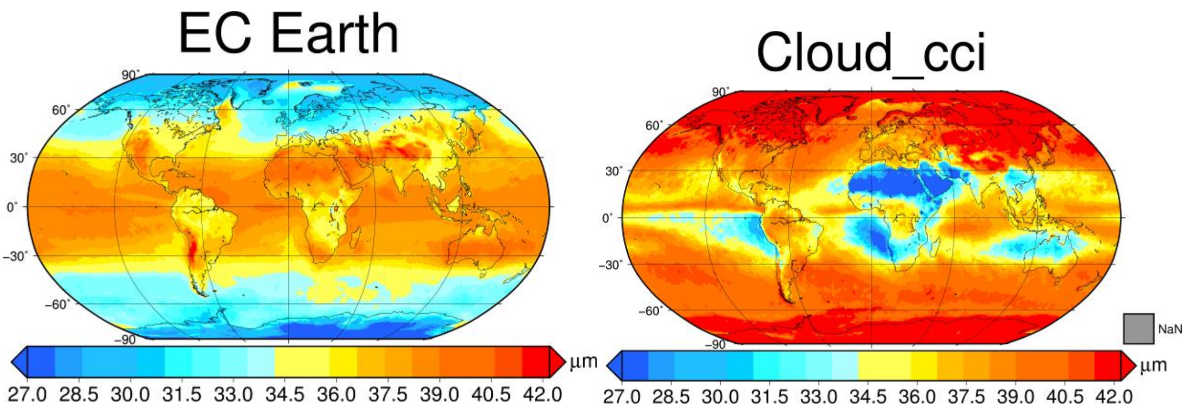


Figure 5-11 *Effective radius for ice clouds. The values are the average retrievals from 1982-2012. Left: Cloud_cci simulator based on an EC Earth model atmosphere. Right: Cloud_cci climatology using AVHRR afternoon (PM) data.*

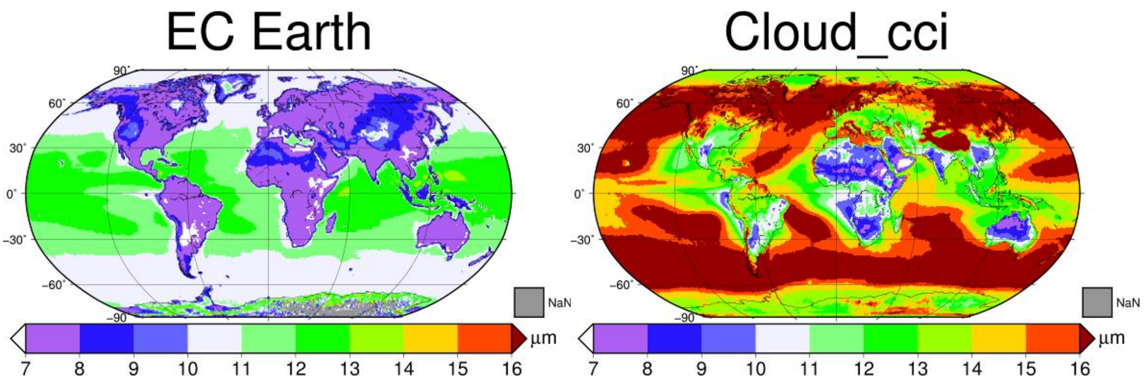


Figure 5-12 *As Figure 5-11, but for Effective radius for liquid clouds.*

The cloud albedo shown in Figure 5-13 shows some common patterns between the model and observations, such as a decreased cloud albedo over China and Northern Europe, yet the datasets completely disagree at the high latitudes and desert regions. These differences naturally require more investigation than presented here.

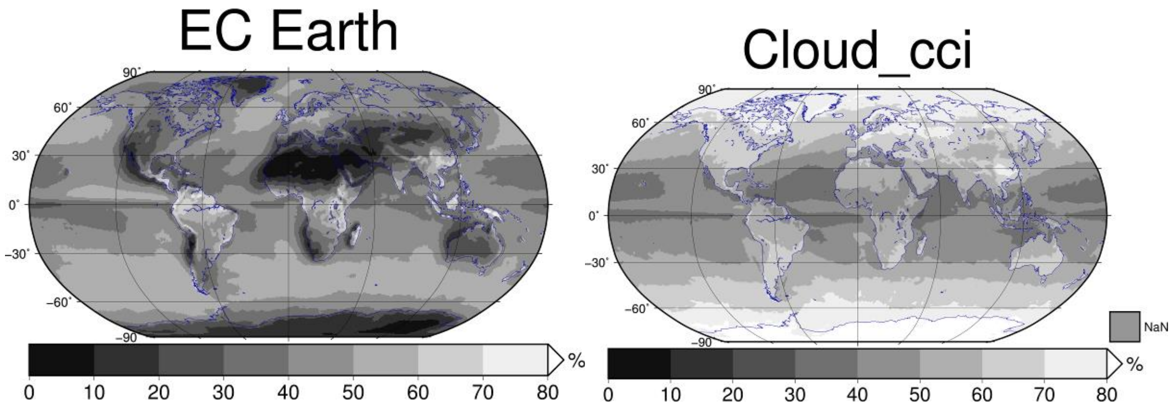
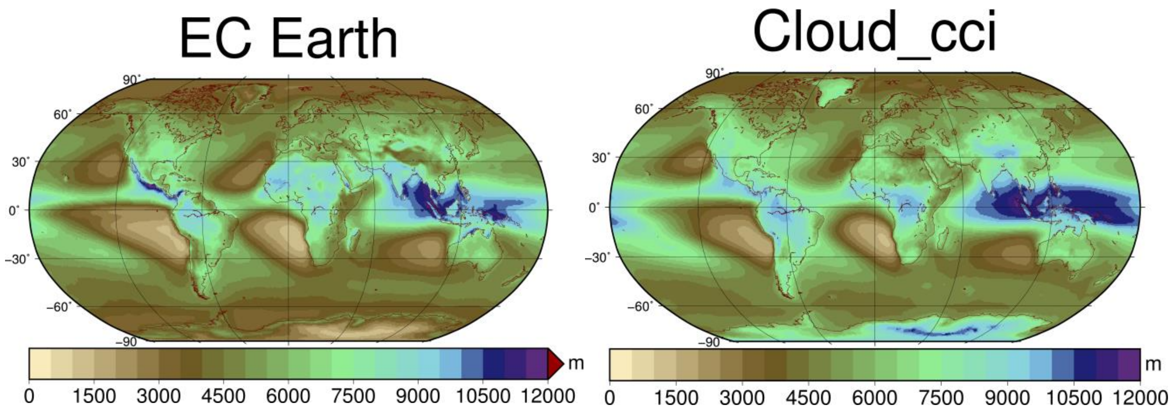


Figure 5-13 *As Figure 5-11, but for Cloud albedo.*

The corrected Cloud Top Height (CTH) is simulated by finding the altitude where the model cloud exceeds a visible optical depth of 0.3 integrated from the model cloud top, and the "non-corrected" CTH is found at a cloud optical depth is 1.0 from the model cloud top. The "non-corrected" cloud top is actually exactly the same simulation as made by the MODIS simulator in COSP. The corrected CTH is shown in Figure 5-14, and overall the model generally underestimates the CTH compared to Cloud_cci. Notably, the cloud top is considerably higher in the central Pacific in the observations than in the model. Very large differences are also seen in the high altitude regions of the Tibetan plateau, Greenland, and Antarctica. However, over a large region in North Africa and Arabia the clouds in the model are instead higher to much higher in EC Earth compared to the Cloud_cci retrievals.



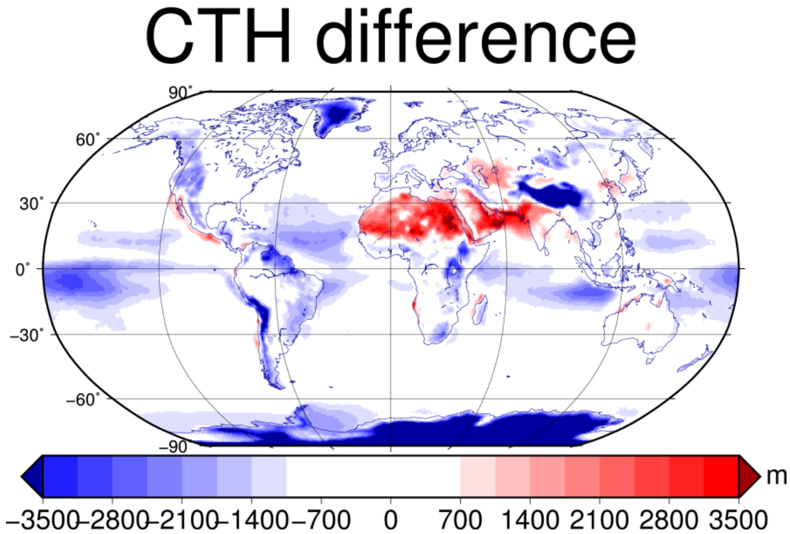


Figure 5-14 *Cloud Top Height-corrected. The values are the average retrievals from 1982-2012. The top left panel shows the simulated corrected CTH based on the EC Earth atmosphere, the top right panel shows the corrected CTH retrievals from Cloud_cci, and the bottom panel shows the difference EC Earth – Cloud_cci.*

5.3.4 Conclusions for global climate studies with the cloud simulator

A Cloud_cci satellite simulator has been developed that can be run for future CMIP simulations. It can be used to study many cloud variables in addition to the total cloud cover e.g. cloud top pressure, optical thickness, effective radius, albedo and liquid/ice water path.

The main mechanisms behind bringing the cloudiness from the observations and climate model closer together are due to a correct temporal sampling of the model data and the removal of thin clouds from the model. Using one year of data, 1982, the impact of the simulator on cloudiness comparisons is shown by contrasting observations-to-model output directly, compared to using the cloud_cci simulator inbetween. For this year at least, the difference in cloudiness between EC Earth and Cloud_cci reduces considerably, by up to 20% when using the simuator.

Overall for the period of 1982—2012, EC Earth has 10—20% less clouds at mid-latitudes over Ocean, and higher total cloud cover over land in general. The highest values are seen in Eastern Africa, Eastern South America, Eastern Australia, and Southern Asia. For the cloud top height variables, EC Earth appears to have lower cloud tops in the Tropics compared to Cloud_cci, but much higher clouds over the Sahara. Outside the tropics, except for over ice-capped regions, the EC Earth's cloudiness matches the observations fairly closely.

In terms of the microphysical variables, effective radius, water path, and cloud albedo, the difference between the model and observations is large. It is very difficult to attribute the differences to one or the other, thus more in depth studies are needed to understand the origins of these differences.

The Cloud_cci simulator is shown to be a valuable tool for model inter-comparisons, and will be included in COSP version 2 so that other climate models can have easier access to Cloud_cci's satellite simulator.

5.4 CMIP5 Global climate studies

To ensure the Cloud_cci data will be used by a larger community the data have been included into the Earth System Model Evaluation Tool (ESMValTool, *Eyring et al., 2016*), which is a community diagnostics and performance metrics tool for the systematic evaluation of Earth system models (ESMs). ESMvaltool is used for routine evaluation the Coupled Model Intercomparions Project (CMIP ref) models. Here CMIP5 models are compared with Cloud_cci and other satellite data sets (Lauer *et al., 2017*).

5.4.1 The CMIP5 models

In this study output from 24 global climate models is used that participated in CMIP5 (Taylor et al. 2012). The model data were obtained from the World Climate Research Programme's (WCRP) CMIP5 data archive made available through the Earth System Grid Federation. The concentration driven CMIP5 historical simulations have been analyzed which are available until 2005. For the time period from 2005 to 2014 results have been used from simulations extended in time with forcing under the Representative Concentration Pathways (RCP) 4.5. RCP4.5 is a scenario applied within CMIP5 prescribing future greenhouse gas concentrations and resulting in a radiative forcing of 4.5 W m^{-2} in the year 2100 relative to pre-industrial values (Clarke et al., 2007).

5.4.2 The Cloud_cci data

Monthly mean cloud fraction data (inferred from Level 3C data product with 0.5° resolution on a latitude-longitude grid) have been used for comparison with the CMIP5 model results. Total Cloud fraction (clt) represents the monthly summary of the results of Community Cloud retrieval for CLimate (CC4CL) cloud detection scheme (Sus et al., 2017; McGarragh et al., 2017). The monthly mean cloud detection uncertainty, cltu, is also inferred from Level 3C products. The cloud mask uncertainty is based on hit rate scores against measurements from the Cloud-Aerosol Lidar with Orthogonal Polarization,(CALIOP). All pixel level uncertainties are propagated in a mathematically consistent way into the Level 3C products (Stengel et al., 2017).

As shown in the PVIRv2, CC4CL cloud detection results have been validated against CALIOP space-based lidar measurements, with a global Hanssen-Kuiper skill score of 0.66 and a global hit rate of 81 % demonstrating the high quality of the cloud detection in the AVHRR-PM v2.0 data set. It needs to be noted, that the Cloud_cci AVHRR-PM data set has a few limitations of which particularly the underrepresentation of optically very thin clouds (with optical thicknesses of below 0.15) and the sparse temporal sampling (twice a day for non-polar regions) is of relevance when using this data set for model evaluation. Particularly difficult conditions for cloud detection are polar night periods, for which the detection scores decrease significantly in the current version of the data set. Furthermore, the monthly cloud fraction data and the corresponding uncertainties of the Cloud_cci AVHRR-PM data set used in this study have not undergone any further processing such as satellite drift correction.

5.4.3 Results of CMIP5 global climate studies

The ESMvaltool basic statistics (mean, bias and variability) are used to compare Cloud_cci AVHRR-PM (CCI hereafter) total cloud cover to CMIP5 models and to other AVHRR satellite-based cloud datasets PATMOS-x (Heidinger et al., 2014) and CLARA-A2 (Karlsson et al., 2013) and the MODIS cloud dataset (Platnick et al., 2003). The 1982-2014 annual mean cloud cover for all CMIP5 models, ERA-Interim and satellite data sets are shown in Figure 5-15. The cloudiness of most CMIP5 models compare well with CCI clt and the alternative reference datasets on a global scale, but there are geographical differences. The CMIP5 models deviating mostly CCSM4, CESM1-BGC, HadCM3, MIROC-ESM and MIROC-ESM-CHEM, underestimate cloud amount on a global scale. The CMIP5 model mean bias compared to CCI (Figure 5-16) show an underestimation of cloud amount especially for the subtropical stratocumulus regions off the west coasts of North and South America as well as off the coast of Australia as known from many previous studies (e.g. Nam et al 2012). In contrast, the CMIP5 model mean and most individual models overestimate cloud amounts by 20 % over the sub-tropical high pressure cloud minima regions. These biases remain but are smaller (10-15 %) if the models instead are compared to PATMOS-x and CLARA-A2, since their cloud cover are larger than for CCI for these regions.

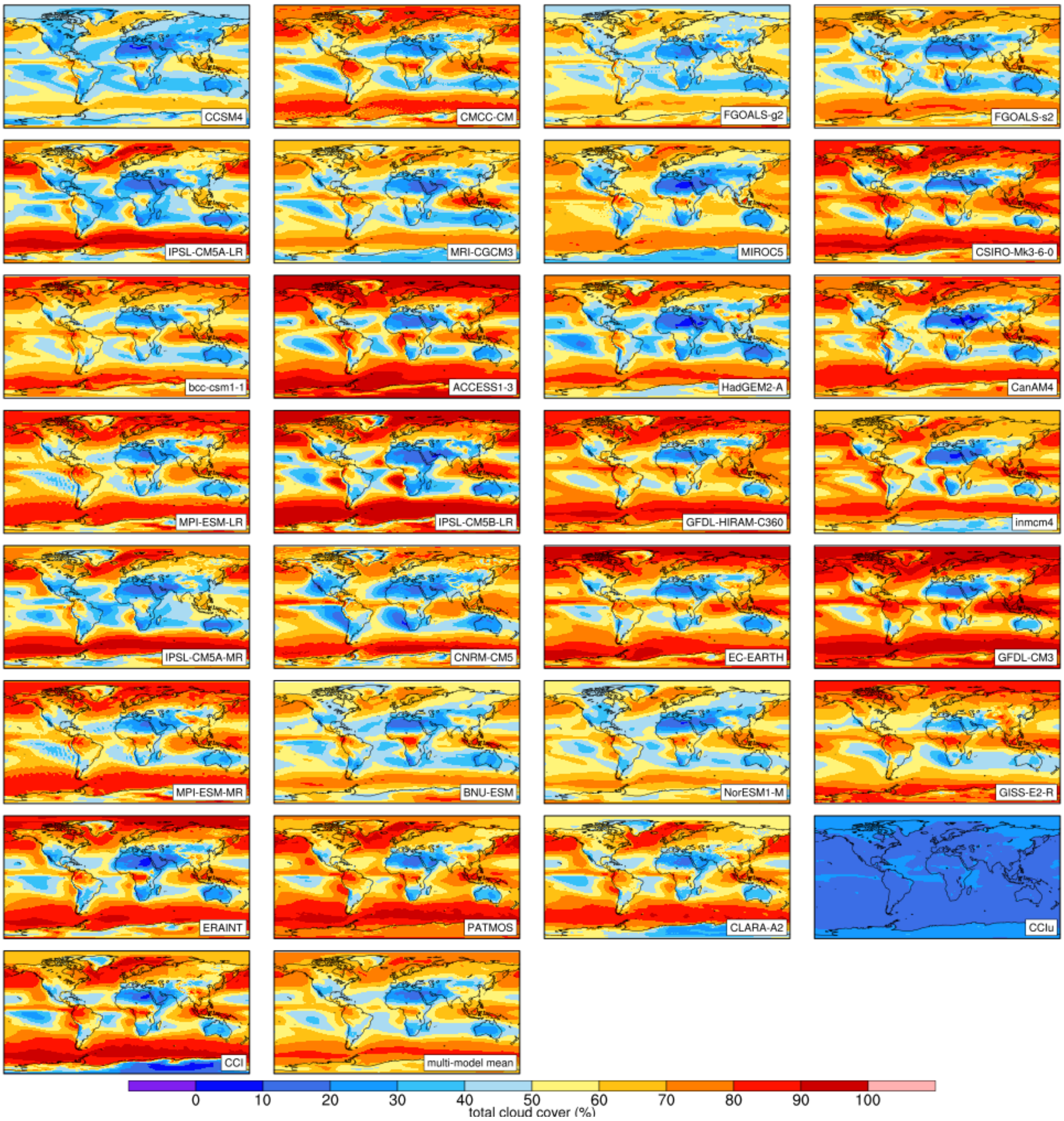


Figure 5-15 *The top 6th rows show all the 24 CMIP5 model mean clt, the 7th row show ERAINT, CLARA-A2, PATMOS-X clt, and CCI cltu and the bottom row show CCI clt and the 24 model clt mean for the years 1982-2014. The CCI uncertainty is shown in second last bottom row, 4th column.*

Figure 5-16 shows CCI total cloud cover in boreal winter and summer and the associated uncertainties. The inherent AVHRR difficulties in detecting clouds for polar night and high altitude snow covered areas (North Canada, North East Asia and Himalayas) give CCI uncertainties of more than 20 % for these regions. Comparing CCI zonal mean cloudiness to other AVHRR cloud datasets PATMOS-x and CLARA-A2 and the MODIS cloud dataset also show the largest observational spread (40-50 %) for high latitudes in the winter hemisphere. The zonal uncertainty in Figure 5-16 is a simple area weighted average, which overestimate the uncertainty. How the zonal mean uncertainties should be calculated as well as how L3 uncertainty is derived from L2 uncertainty warrant further work in CCI+ also for CCI ECV's in general. The largest model spread (60 %) occurs for high latitudes in polar winter, where also the observational datasets have their largest uncertainties as seen in the zonal mean plots in Figure 5-16. For these cold conditions the amount of cloud condensate is small and the model clouds are often thinner than the satellites can detect, using a simulator removes part of these model clouds. The CCI uncertainties are also high with up to 20 % for the subtropical high pressure dry zones.

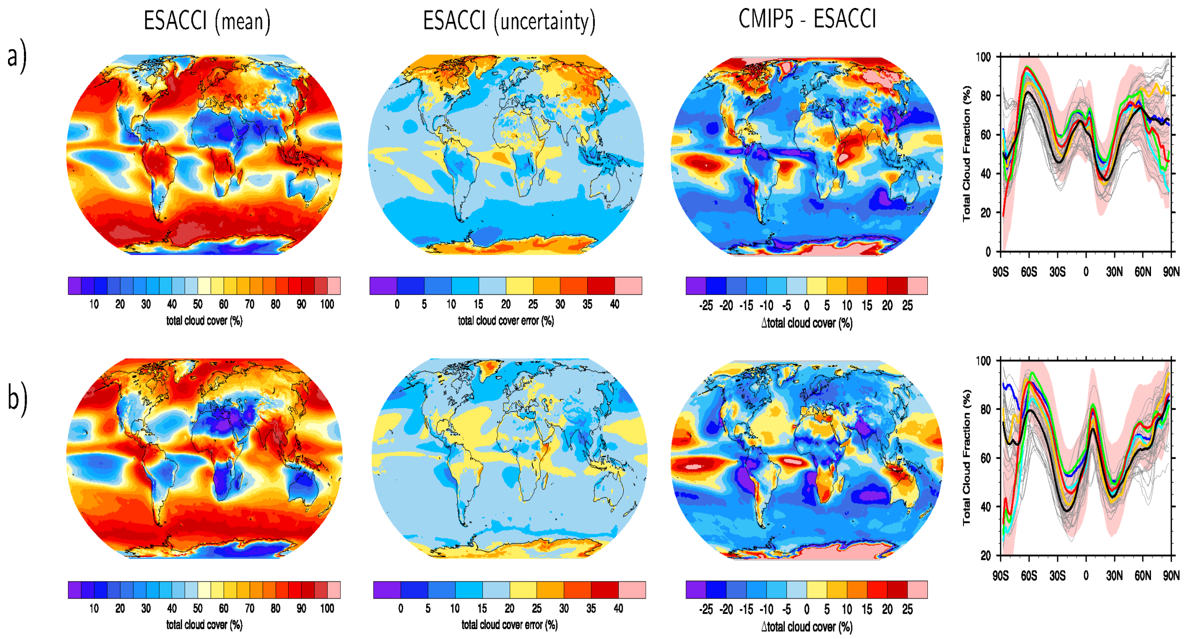


Figure 5-16 *Maps of the multi-year seasonal mean of total cloud cover and 1-sigma uncertainty from ESA CCI cloud for a) December-January-February (DJF) and b) June-July-August (JJA) 1982-2014. The figure also shows the differences between the ESA CCI data and the CMIP5 multi-model mean as well as zonal means. The zonal mean panels show averages from ESA CCI (red), PATMOS-x (blue), CLARA-A2 (cyan), MODIS (green), ERA-Interim (orange), and the CMIP5 multi-model mean (black). The individual CMIP5 models are shown as thin grey lines and the observational uncertainties of the ESA CCI data (±1-sigma) are shaded in light red. The MODIS data are only available for the years 2003-2014.*

Figure 5-17 shows the interannual variability for the satellite datasets and the CMIP5 model mean and for ERA-Interim. All the AVHRR datasets have their largest variability (30-40%) for the dry tropical high pressure regions over the oceans, over North Africa, south Africa and Australia, reflecting the annual shift of the ITCZ and the El Niño/ Southern Oscillation (ENSO). MODIS tropical Pacific Ocean variability is smaller than for the AVHRR datasets, since MODIS data is available only for 2003-2014, years that do not include the major El Nino's in the 1980 and 1990'ties, which illustrates the importance of using long term observational records when evaluating ENSO. CCI has larger variability over the tropical Pacific Ocean than the other AVHRR satellite datasets. Time series (Figure 5-18) reveal that Cloud_cci clt is of similar magnitude to PATMOS-x and CLARA-A2 for El Niño years when the cloud cover is maximum, while CCI has less cloud amount (5-15%) for La Niña years when the cloud cover has its minima, resulting in larger CCI variability which needs further understanding. PATMOS-x has less variability and higher cloud amount over Antarctic than CCI and CLARA-A2.

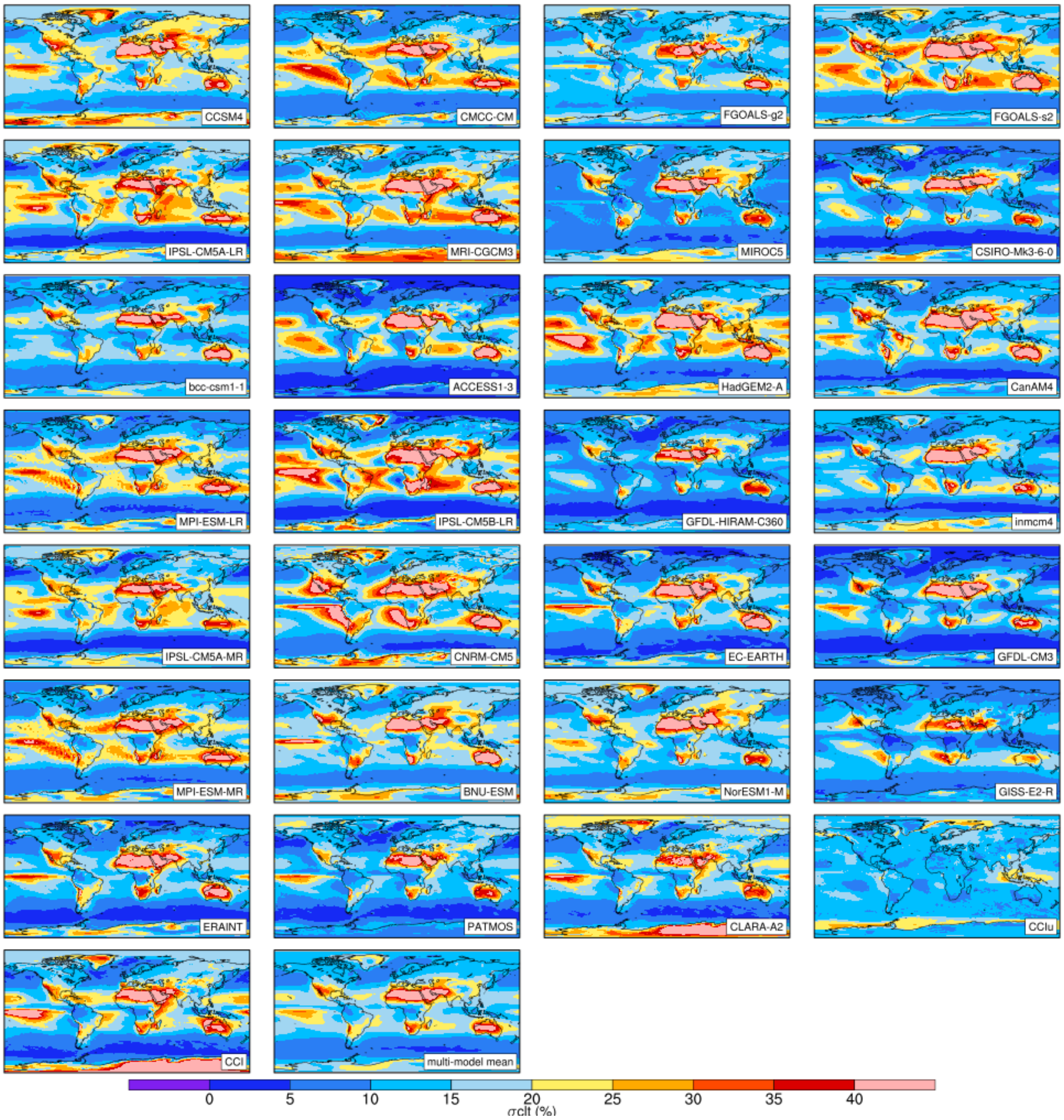


Figure 5-17 *The top 6th rows show all the 24 CMIP5 model clt interannual variability, the 7th row show ERAINT, CLARA-A2, PATMOS variability, and CCI cltu and the bottom row show CCI clt and the 24 model mean.*

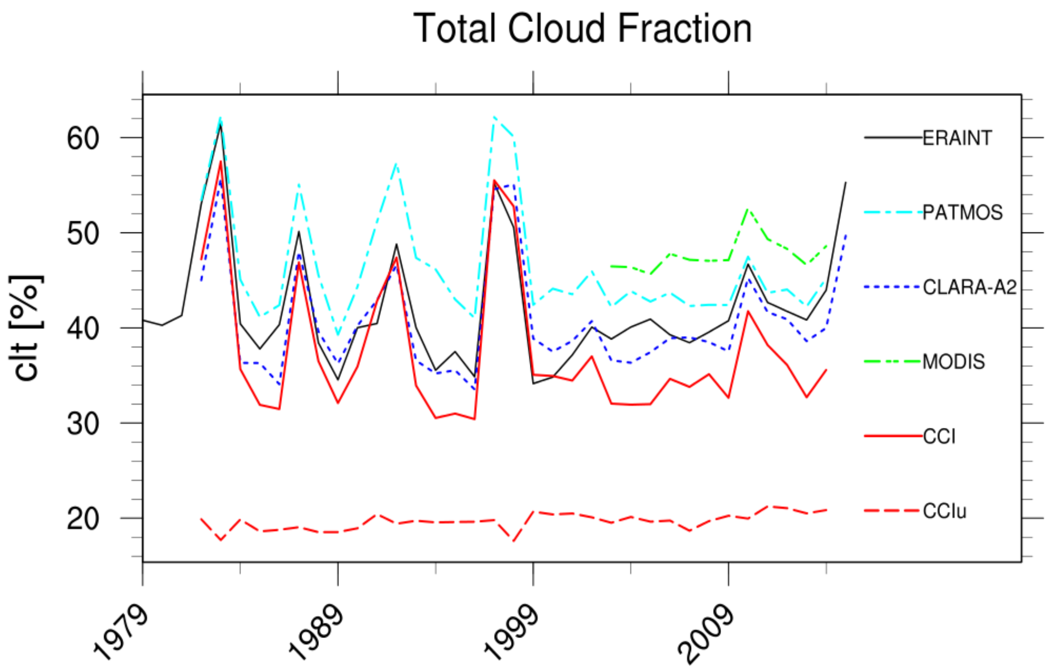


Figure 5-18 *The time series of clt for the Nino3.4 area (5°S-5°N, 190°E-240°E) for ERAINT (black), PATMOS-X (cyan), CLARA-A2 (blue), MODIS (green), CCI clt (red) and CCI cltu (red hatched).*

The CMIP5 model means show less variability than the observations, especially over the sub-tropical high pressure regions, where most of the individual CMIP5 models overestimate the total cloud cover. Contrary the models that underestimate clt for the dry regions (CCSM4, CESM1-BGC, HadCM3, MIROC-ESM, MIROC-ESM-CHEM) have larger variability. The tropical high pressure regions also show the largest difference between Cloud_cci clt compared to other AVHRR datasets (Figure 5-19).

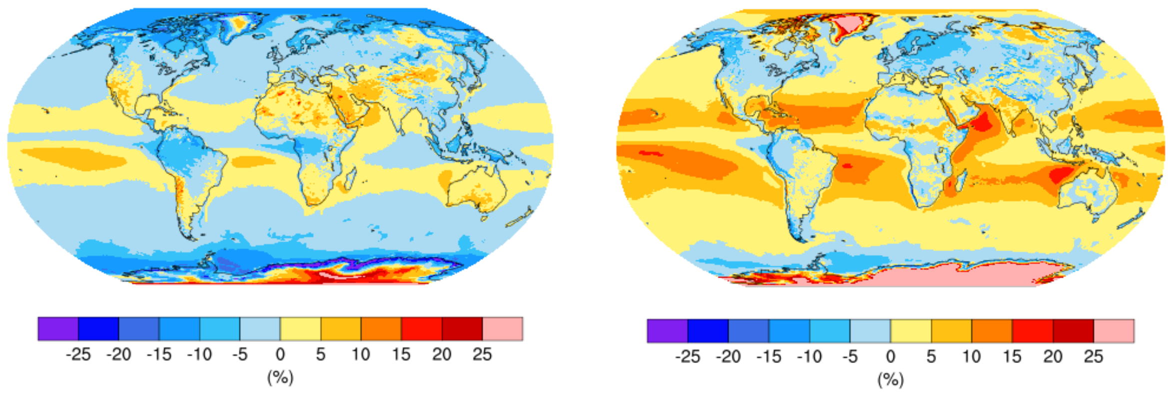


Figure 5-19 *The annual clt difference between Cloud_cci and CLARA (left side) and between Cloud_cci and PATMOS (right side).*

5.4.4 Conclusions of CMIP5 global climate studies

The Cloud_cci AVHRR-PM total cloud cover compares well with other existing long term AVHRR cloud datasets. The Cloud_cci pixel based uncertainties show the user which areas should be carefully treated, e.g. polar and high altitude snow covered regions where the passive satellites have problems detecting clouds. Cloud_cci has lower cloud minima than the other AVHRR datasets for the tropical Pacific, which should be investigated. The other Cloud_cci datasets with shorter time records, MODIS, ATSR-2, AATSR and MERIS can be used for process studies and for narrowing the observational uncertainties. The CMIP5 models cloud cover show typical error patterns compared to CCI and the other satellite datasets, underestimating clouds in the stratocumulus regions and overestimating clouds in the subtropical dry regions. More detailed analysis of the individual models and the interaction with radiation are needed to understand these biases.

6. Analysis of tropical cloud properties

In this section we report and discuss the response of cloud properties to changes in atmospheric water vaour and sea surface temperature in the Tropics.

6.1 Background

It is well known that total columnar water vapour (TCWV) over ice free ocean shows a quantitatively positive feedback to the changes in sea surface temperature (SST) (e.g. Stephens, 1990). This response follows well theoretical considerations.

In contrast to the SST-TCWV link, the relation between TCWV and cloud properties is, due to the variety of cloud processes, less obvious. One common hypothesis is that moist atmospheres have more and deeper clouds while dry atmospheres have fewer and thinner clouds. However, due to the complexity of the development of clouds in different dynamical regimes this hypothesis is difficult to be assessed.

For an investigation of the TCWV-cloud relation on global scales satellite measurements are exclusively suited due to the near-global coverage of related remote sensing applications. A recent study of Forsythe et al. (2012) used the TCWV product of NOAA based on blending of different sources and related the TCWV anomalies to vertical occurrence profiles of CloudSat-CALIPSO. For the three investigated regions they found different relations between TCWV and clouds depending on cloud type, season and region.

Besides the Forsythe et al. (2012) study, the TCWV – clouds relation has only been investigated to a limited extent in the past. One reason for that is the unavailability of global, long-term and high-quality cloud property datasets which are stable enough for this type of analysis. This topic, however, is very important in particular in the light of the found, varying trends over ocean (Figure 3-1), which corresponds to changes in SST.

6.2 Data used

6.2.1 Observations

In this study we used the Cloud_cci AVHRR-PM datasets (See Section 1.2), i.e. the cloud fraction and high cloud fraction. TCWV data is taken from the from HOAPS dataset (Schröder et al., 2013) which was be procured from EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF). HOAPS data also contain observational datasets of SST.

6.2.2 Models

Enhancing the link to the climate modelling and reanalysis community, ERA-Interim fields of clouds, water vapour and SST are used. In terms of cloud properties, cloud fraction and high cloud fraction were used which were generated using the simplistic simulator documented in Section 4.2. It needs to be noted that for ERA-Interim water vapour and SST significant contributions from observations are assimilated. On the other hand, no direct cloud observations is assimilated, making the ERA-Interim clouds being based on model diagnostics exclusively.

6.3 Methodology

For all datasets a common period was defined 1988-2008. Monthly mean properties were used exclusively and all data over land were removed. All data fields were brought onto a common latitude/longitude grid with 1°x1° resolution. The tropical area was defined to be within 30°S to 30°N.

6.4 Results

6.4.1 Trends in the Tropics

Figure 6-1 shows the linear trends of SST and TCWV for the tropical ocean for 1988 – 2008 from both obervations and ERA-Interim. Generally, very similar patterns are found between the observation and ERA-Interim, also reflecting the assimilated satellite information in ERA-Interim. Although, the observational datasets have a tendency to more positive trends. Quite some agreement is also found between the spatial patterns of SST trends and TCWV trends in both the observation and ERA-Interim highlighting the natural connection that higher SST values lead to higher TCWV in many regions cases. However, for some regions TCWV and SST trends have different signs. Although with relatively small trend values.

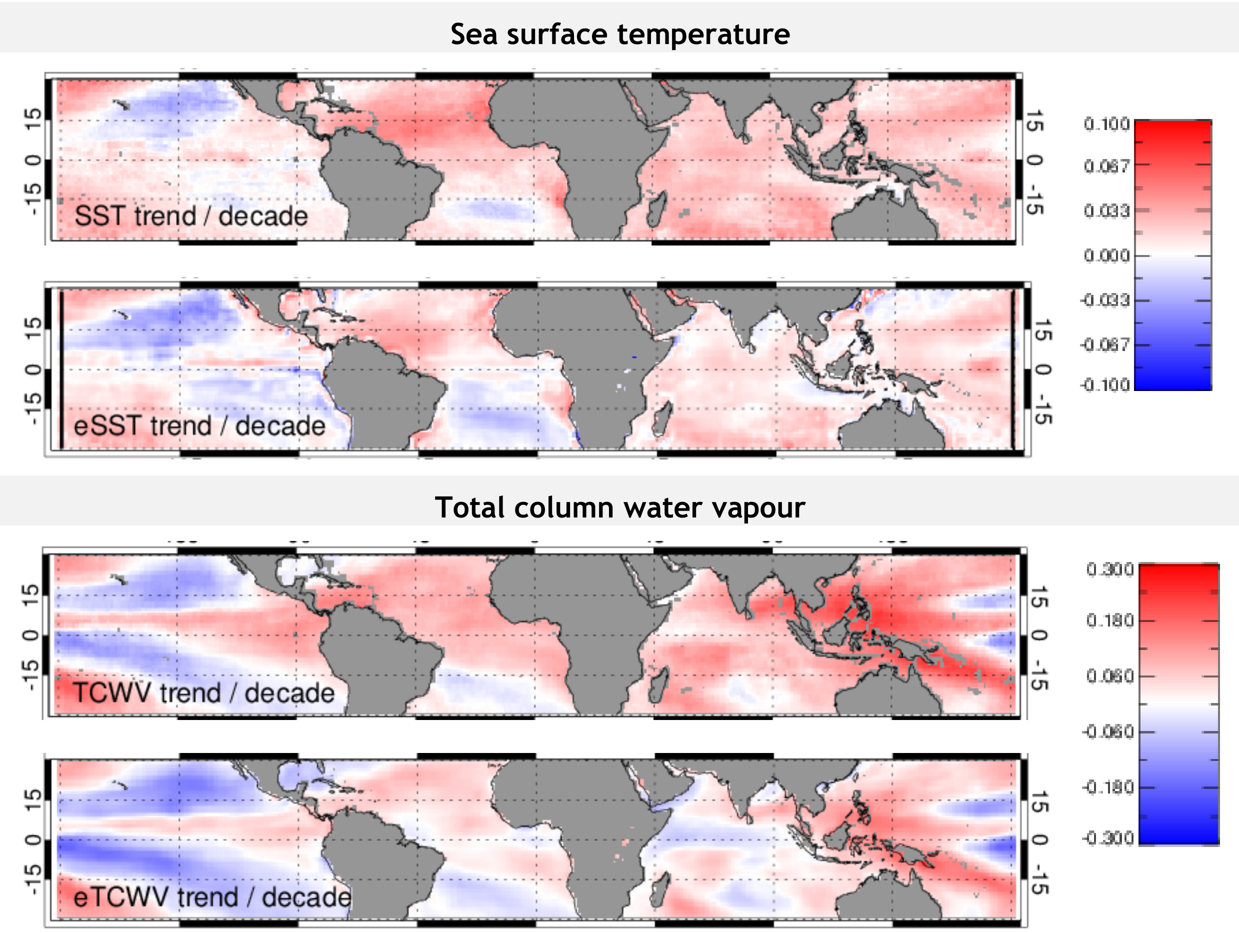


Figure 6-1 *Maps of decadal trends of sea surface temperature (SST) and total column water vapour (TCWV) taken from HOAPS (SST/TCWV) and ERA-Interim (eSST and eTCWV). Figure from Stengel et al. (2017b).*

Figure 6-2 shows the trends for total cloud fraction and for fraction of high clouds for Cloud_cci and ERA-Interim. Unlike TCWV and SST, the ERA-Interim cloud information comes excluvively from model diagnostics and not from assimilation of clouds observations. For total cloud fraction, both regions with agreement and disagreement are found between model and observations. Looking at the trend of the fraction of high clouds, a surprisingly good agreement is found. Some spatially limited regions are found with diffferences, e.g. near the coast of South America, where ERA-Interim exhibits a large positive trend in fraction of high clouds, which is not seen in Cloud_cci. For more detailed comparisons, one needs to consider maps of the significance of these trends or, look more closely into the trends, as in the next section.

In the next section the Tropics are broken down into three selected regions where the temporal variability of SST, TCWV and CFC_{high} and the correlation among them are investigated.

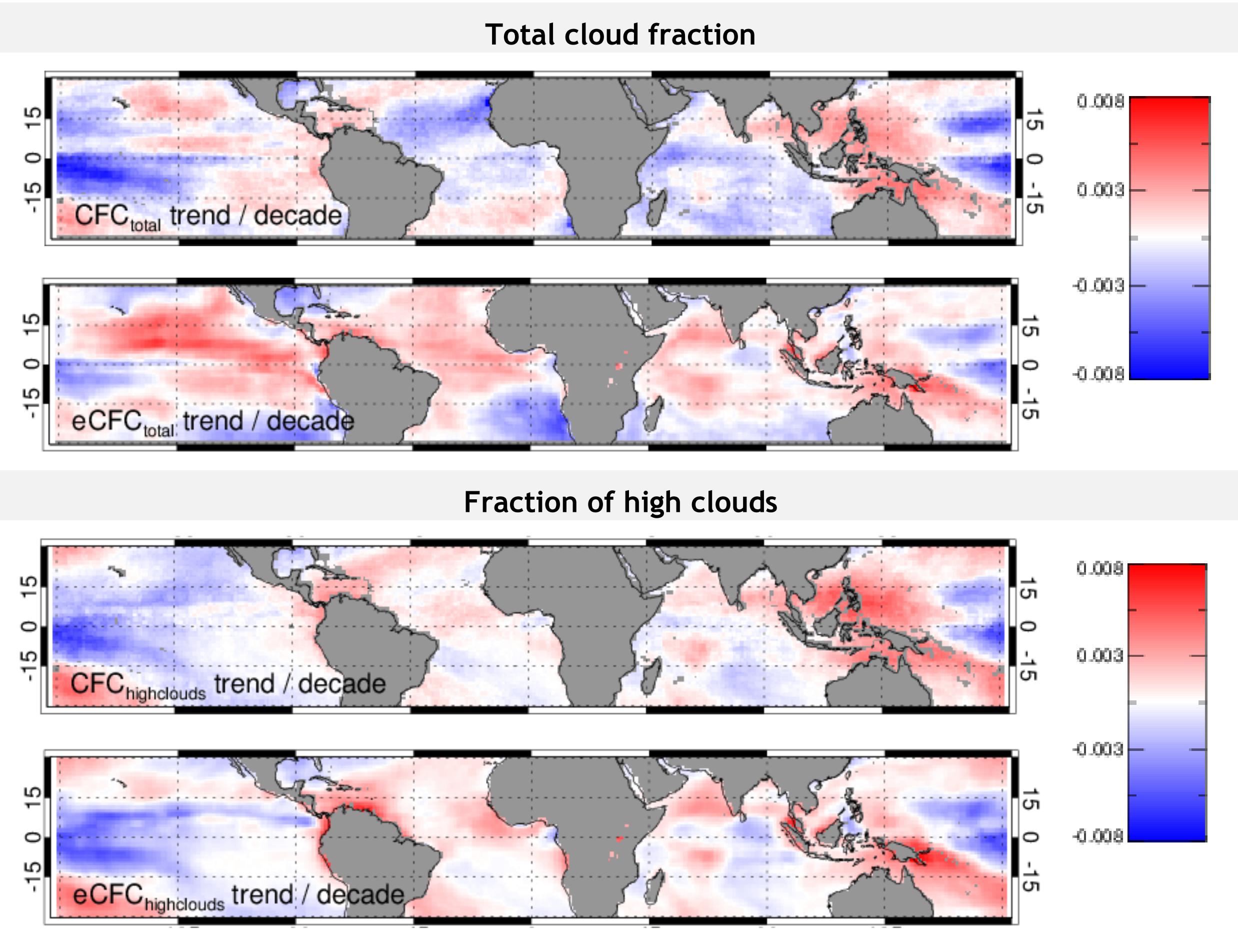


Figure 6-2 *Maps of decadal trends of total cloud fraxtion (CFC) and high cloud fraction (CFC_{high}) taken from Cloud_cci (CFC, CFC_{high}) and ERA-Interim (eCFC, $eCFC_{high}$). Figure from Stengel et al. (2017b).*

6.4.2 Regional studies

Figure 6-3 shows the location of the three selected regions under consideration in this section. Region 1 representes the warm pool region around Indonesia, Region 2 a region in the centre Pacific also offen refered to as ENSO region, and Region 3 - a small Pacific region off the west coast of Middle America. All three regions are characterized by a more or less significant trend in fraction of high clouds.

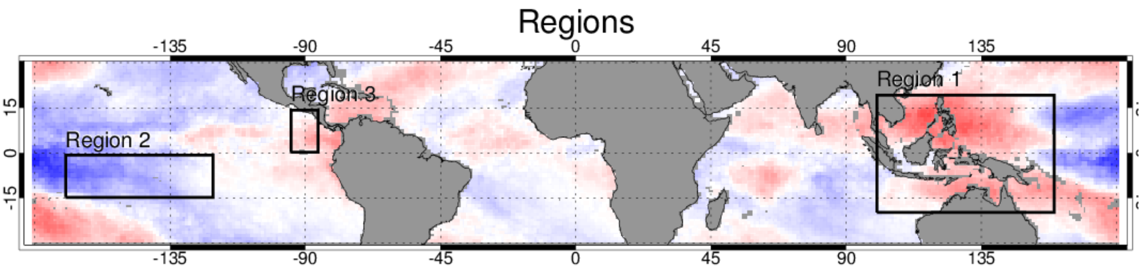


Figure 6-3 *Repeating map of decadal trends of Cloud_cci high cloud fraction (CFC$_{high}$) from Figure 6-2 now overlaid with locations of the three regions of interest. Figure from Stengel et al. (2017b).*

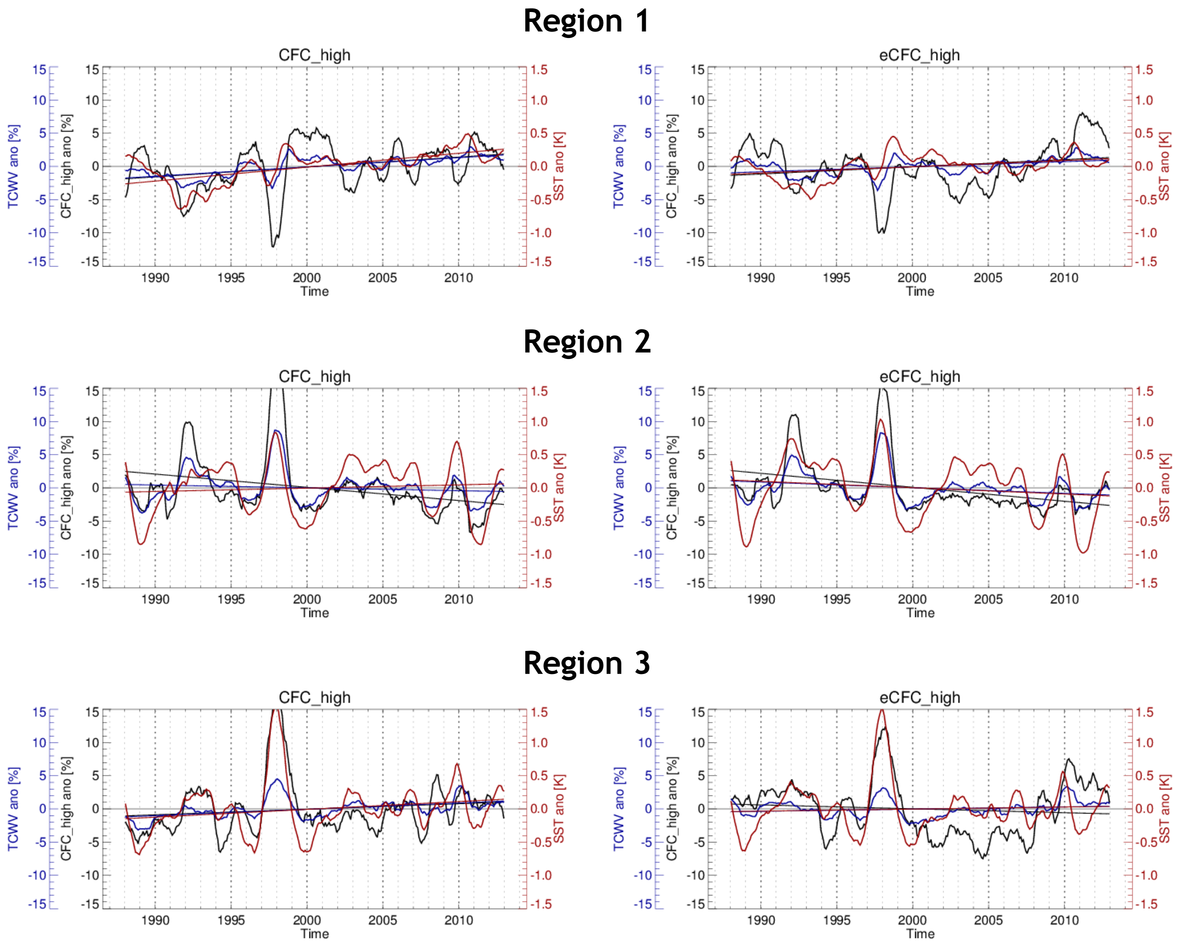


Figure 6-4 Left column : *Time series of fraction of Cloud_cci high clouds (CFC high), HOAPS TCWV and HOAPS SST for regions 1 to 3. Right column: As left column but for ERA-Interim data. Figure from Stengel et al. (2017b).*

Figure 6-4 reports the time series of monthly SST, TCWV and CFC$_{high}$ anomalies for the three tropical regions under considerations for both observations and reanalysis data. For all regions and data sources we find a strong correlation between the SST and TCWV anomalies, but also between the

TCWV and CFC_{high}. The latter are also reported in Table 6-1. While the tropical trend maps (Figures Figure **6-1** and Figure **6-2**) already indicated that there is a lot in common between the observations and ERA-Interim, the time series plots confirm this also for the monthly anomalies in all three regions. In addition, the signs of the trends of TCWV and CFC_{high} agree for observations and ERA-Interim: with positive trends for TCWV and CFC_{high} in Region 1 and negative trend in Region 2, fostering the indication that more water vapour also leads to more high clouds on average. The only aspect in which this is not supported are ERA-Interim trends of TCWV and CFChigh, which are of opposite sign. It remains to be investigated what the main reasons for the findings are.

Table 6-1 *Slope of linear trend per year for monthly fraction of high clouds (CFC_{high}) and total column water vapour (TCWV) as well as the correlation betwwen the monthly data of both soruce for the period 1888 and 2008 for three selected tropical regionsindicated in Figure 6-3. Table from Stengel et al. (2017b).*

Region1	CFC_{high} slope [%/year]	WV slope [%/year]	Correlation CFC_{high}/WV
	Region1		
Cloud_cci/HOAPS	0.15	0.14	0.7362
ERA-Interim	0.11	0.08	0.7324
	Region2		
Cloud_cci/HOAPS	-0.20	-0.04	0.8980
ERA-Interim	-0.21	-0.08	0.8984
	Region3		
Cloud_cci/HOAPS	0.08	0.09	0.7144
ERA-Interim	-0.05	0.03	0.7136

7. Conclusions

Here the main findings from the different assessments are summarized, providing examples of application and a basis for the ready-to-use of the Cloud_cci data sets. In addition to the Product and Validation report this document provides remarks on issues that can be investigated in further re-processing events or projects.

ESA Cloud_cci cloud retrievals are based on an Optimal Estimation technique and use similar instruments (AVHRR, MODIS and AATSR), with spectral information ranging from visible to infrared. Key characteristics of the Cloud_cci cloud datasets are as follows:

- Spectral consistency of derived parameters, which is achieved by an optimal estimation (OE) approach based on fitting a physically consistent cloud model to satellite observations simultaneously from the visible to the mid-infrared.
- Uncertainty characterization, which is inferred from OE theory on pixel level, 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 mathematical framework.
- Potential to combine AVHRR-heritage datasets to achieve increased temporal resolution by including multiple polar-orbiting satellite instruments, which also allows for 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 the exploitation of applicability for evaluation of climate models and reanalyses.

Based on the comparison with the GEWEX Cloud Assessment data base it can be summarized that the total cloud amount (0.68±0.03) compares well to the reference datasets from other passive remote sensing cloud datasets, with a similar good performance of Cloud_cci AVHRR and MODIS.

Cloud amount of AATSR is slightly lower (0.66), and the one of MERIS-AATSR (only during tay-time available) is underestimated (0.60). The data show a good coherence of latitudinal variation and the seasonal cycle, except for MERIS-AATSR. Overall it can be observed that total cloud amount is underestimated over parts of the ocean, especially for AATSR and MERIS-AATSR.

Furthermore, the comparison shows that the identification of high-level clouds is worse during daytime than during night-time. During daytime the amount of high-level clouds, in particular over land, is underestimated for all Cloud_cci datasets. This might be explained by the fact that VIS information in combination with spectral IR information in the Optimal Estimation retrieval method leads to a misidentification in case of multi-level clouds. During night-time, when only spectral IR information is available, Cloud_cci relative high-level cloud amounts compare well to IR sounders, though total cloud amount seems to be slightly underestimated mostly by missing low-level clouds.

The vertical distribution of cloud pressure which is supposed to be bimodal in the tropics with peaks around 250 and 950 hPa, reveal that the height of low-level clouds from the Cloud_CCI data sets seems to be slightly overestimated. The peak for high-level clouds, which is located at the right height, is smaller than for other other datasets during daytime / night-time.

To overcome this situation and to improve the performance for high-level clouds, one path might be to adapt the OE method using daytime spectral information to the one using night-time spectral information (which excludes VIS information). By applying both methods during daytime, keeping the VIS information for the cloud mask, one could even get information of multi-layer cloud situations from a comparison of VIS optical depth to IR emissivity.

Cloud_cci data sets have been applied in the evaluation of (a) the ERA-Interim Reanalyses, (b) regional climate models, (c) EC-Earth global climate model and (d) for CMIP-5 models.

Based on the comparisons on the regional scale, the use and application of the Cloud-CCI simulator is essential in setting up a direct comparison between model predicted cloud parameters and Cloud_cci based retrievals. Without the simulator, quantitative comparison of model cloud liquid water and cloud ice parameters with the corresponding values inferred from the retrievals is meaningless. An application of the Cloud-CCI simulator to 25-km resolved multi-level cloud fields obtained with a one month integration in hindcast mode of the regional climate model RACMO2 shows that the simulated fields are consistently (much) closer to the Cloud_cci AVHRR cloud property retrievals than the original model fields, the exception being the ice water path field.

When the regional atmospheric (climate) model is driven by large-scale (re-)analyses at the lateral boundaries and operated in quasi-NWP mode, a direct comparison between simulated model output and retrievals is feasible on the basis of a daily composite, allowing the evaluation of model cloud parameters corresponding to individual weather systems. First results indicates that the model produces higher and colder cloud tops and a higher ice water path than the retrieval indicates.

The conducted comparison yield that the possibility to consider the provided uncertainty for cloud mask and other variables helps to decide which cloud types are more important for model evaluation and which should be treated with care. Here, based on the comparison with the global climate model, polar regions as well as high altitude snow covered regions became visible where passive instrument satellites have in general problems detecting clouds. However, to optimize the use of the Cloud_cci uncertainties it would be useful to have more guidance from the data providers of how average uncertainties should be calculated and compared to model estimates and to have insight in the derivation of L3 uncertainties from L2 data.

Also the comparisons show that Cloud_cci has lower cloud minima than the other AVHRR datasets for the tropical Pacific, which need further investigation. This could be related to underestimation of low level clouds as found in the GEWEX study.

8. Acronyms

Acronym	Explanation
1M	One-moment bulk cloud-microphysics parameterization
2M	Two-moment bulk cloud-microphysics
AATSR	Advanced Along Track Scanning Radiometer
AIRS	Atmospheric Infrared Sounder
ANN	Artificial Neural Network
ATSR	Along-Track Scanning Radiometer
AVHRR	Advanced Very High Resolution Radiometer
bcRMSE	Bias-corrected Root Mean Square Error
BRDF	Bi-directional Reflectance Distribution Function
CA	Cloud Amount, Cloud Assessment (depends on the context)
CAH	High Cloud Amount
CAHR	Relative High Cloud Amount
CAL	Low Cloud Amount
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CALR	Relative Low Cloud Amount
CAM	Middle Cloud Amount
CAMR	Relative Middle Cloud Amount
CC4CL	Community optimal estimation Cloud retrieval For CLimate
CC	CM SAF cloud, albedo and radiation dataset -AVHRR based Version 1
CCI	Climate Change Initiative
CDO	Climate Data Operators
CEM	Cloud Emissivity
CIWP	Cloud Ice Water Path
CLWP	Cloud Liquid Water Path
CM	Cloud Mask
COD	Cloud Optical Depth
CODI	Cloud Optical Depth for Ice clouds
CODW	Cloud Optical Depth for Water clouds
COT	Cloud Optical Thickness (equivalent to COD, Cloud Optical Depth)
CP	Cloud Pressure
CPS	Cloud Particle Size
CRE	Effective Radius of Cloud particles
CT	Cloud top temperature

Acronym	Explanation
CTH	Cloud-Top Height
CTP	Cloud-Top Pressure (equivalent to CP, Cloud Pressure)
CWP	Cloud Water Path
CZ	Cloud Height
DCOMP	Daytime Cloud Optical and Microphysical Properties algorithm
ECMWF	European Centre for Medium-Range Weather Forecasts
EnviSat	Environmental Satellite
ERA	ECMWF Re-analysis
ESA	European Space Agency
FAME-C	Freie Universität Berlin AATSR-MERIS Cloud retrieval algorithm
GCM	General Circulation Model
GDAP	GEWEX Data and Assessment Panel
GEWEX	Global Energy and Water EXchanges project
GMAO	Global Modeling and Assimilation Office
HIRS	High resolution Infrared Radiation Sounder
HTESSEL	Hydrology-TESSEL
IFS	Integrated Forecast System
IR	Infrared
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Intertropical Convergence Zone
IWP	Ice Water Path
LMD	Laboratory of Dynamic Meteorology
LUT	Look-Up Table
LWP	Liquid Water Path
LWRTM	Longwave Radiative Transfer Model
MBE	Mean bias error
MERIS	Medium Resolution Imaging Spectrometer
MISR	Multi-angle Imaging SpectroRadiometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MOMO	Matrix Operator Model
NCEP	National Centers for Environmental Prediction
NEMO	Nucleus for European Modelling of the Ocean model
NH	Northern Hemisphere
NIR	Near Infrared
NOAA	National Oceanic and Atmospheric Administration

Acronym	Explanation
NWP	Numerical Weather Prediction
OASIS	Coupling software, https://verc.enes.org/oasis
ORAC	Oxford RAL retrieval of Aerosol and Cloud
PARASOL	Polarization and Anisotropy of Reflectances for Atmospheric science coupled with Observations from a Lidar
PATMOS-x, PATMOSX	Pathfinder Atmospheres Extended
PDF	Probability density function
POLDER	POLarization and Directionality of the Earth's Reflectances
REF	Effective radius
RMSE	Root-Mean-Square Error
RTM	Radiative Transfer Model
RTTOV	Radiative Transfer for TOVS
SH	Southern Hemisphere
SMHI	Swedish Meteorological and Hydrological Institute
SNHT	Standard Normal Homogeneity Test
SST	Sea Surface Temperature
STDD	STanDard Deviation
SWIR	Short-Wave Infrared
SWRTM	Shortwave Radiative Transfer Model
TCC	Total Cloud Cover
TESSEL	Tiled ECMWF Scheme for Surface Exchanges over Land
TOA	Top of atmosphere
TOC	Top of cloud
TOVS	TIROS Operational Vertical Sounder
VIS	Visible

9. References

ATBDv5, 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

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: 1, date of Issue: 28/08/2017, Available at: http://www.esa-cloud-cci.org/?q=documentation

Alexandersson, H., 1986. A homogeneity test applied to precipitation data. J. Climatol. 6, 661–675. doi:10.1002/joc.3370060607

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: description and sensitivities. Monthly Weather Review, 139, 3887-3905.

Balsamo, G., P. Viterbo, A. Beljaars, B.J.J.M. van den Hurk, M. Hirschi, A. Betts and K. Scipal, 2009: A revised hydrology for the ECMWF model: Verification from field site to terrestrial water storage and impact in the Integrated Forecast System. J. Hydrometeor., 10, 623-643, doi:10.1175/2008JHM1068.1

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. Journal of Geophysical Research: Atmospheres, 119, 7889-7907.

Bechtold, P., N. Semane, P. Lopez, J.-P. Chaboureau, A. Beljaars, and N. Bormann (2014). Representing equilibrium and non-equilibrium convection in large-scale models. J. Atmos. Sci. 71 (2), 734–753.

Bechtold, P., M. Köhler, T. Jung, M. Leutbecher, M. Rodwell, F. Vitart, and G. Balsamo, 2008: Advances in simulating atmospheric variability with the ECMWF model: From synoptic to decadal time scales. Quart. J. Roy. Meteor. Soc.,134,1337–1351.

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.

Bodas-Salcedo, A., M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John (2011), COSP Satellite simulation software for model assessment, Bull. Amer. Meteorol. Soc., 92, 1023-1043, doi: 10.1175/2011BAMS2856.1.

Boussetta, S. Simarro, C. and Lucas, D., 2016, Exploring EC-Earth 3.2-Beta performance on the new ECMWF Cray-Broadwell Technical Memorandum 779.

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

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

Chepfer H., S. Bony, D. Winker, G. Cesana, J.L. Dufresne, P. Minnis, C.J. Stubenrauch,and S. Zeng, The GCM Oriented Calipso Cloud Product (CALIPSO-GOCCP). J. Geophys. Res., 115, D00H16, doi:10.1029/2009JD012251, 2010.

Clarke, L. E., J. A. Edmonds, H. D. Jacoby, H. Pitcher, J. M. Reilly, and R. Richels (2007), Scenarios of greenhouse gas emissions and atmospheric concentrations, Sub-report 2.1a of Synthesis and

Assessment Product 2.1, Climate Change Science Program and the Subcommittee on Global Change Research, Washington DC.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfer modeling: A summary of the AER codes. J. Quant. Spectrosc. Radiat. Transfer, 91, 233–244.

CMIP5 Coupled Model Intercomparison Project: http://cmip-pcmdi.llnl.gov/cmip5.

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, L., 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., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, doi:10.1002/qj.828, 2011.

Devasthale, A., Karlsson, K., Quaas, J., & Grassl, H., 2012. Correcting orbital drift signal in the time series of AVHRR derived convective cloud fraction using rotated empirical orthogonal function. Atmospheric Measurement Techniques, 5, 267-273.

Dutra, E., G. Balsamo, P. Viterbo, P. M. A. Miranda, A. Beljaars, C. Schär, and K. Elder, 2010: An improved snow scheme for the ECMWF land surface model: Description and off line validation. J. Hydrometeor., doi: 10.1175/2010JHM1249.1

ECWMF-IFS, 2009. IFS documentation-Cy33r1. PART IV: Physical Processes. http://www.ecmwf.int/sites/default/files/elibrary/2009/9227-part-iv-physical-processes.pdf

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

Eyring, V., M. Righi, A. Lauer, M. Evaldsson, S. Wenzel, C. Jones, A. Anav, O. Andrews, I. Cionni, E. L. Davin, C. Deser, C. Ehbrecht, P. Friedlingstein, P. Gleckler, K.-D. Gottschaldt, S. Hagemann, M. Juckes, S. Kindermann, J. Krasting, D. Kunert, R. Levine, A. Loew, J. Mäkelä, G. Martin, E. Mason, A. S. Phillips, S. Read, C. Rio, R. Roehrig, D. Senftleben, A. Sterl, L. H. van Ulft, J. Walton, S. Wang, and K. D. Williams (2016), ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth System Models in CMIP, Geosci. Model Dev., 9, 1747-1802, doi: 10.5194/gmd-9-1747-2016.

Feofilov, A. G., and C. J. Stubenrauch, 2017: LMD Cloud Retrieval using IR Sounders: Algorithm Theoretical Basis Document for CIRS-LMD software package (V2), DOI: 10.13140/ RG.2.2.15812.63361

Forbes, R. M., Geer, A., Lonitz, K., and Ahlgrimm, M.: Reducing systematic errors in cold-air outbreaks, ECMWF Newsletter, 146, 17–22, Winter 2015/2016.

Forsythe, J. M., Dodson, J. B., Partain, P. T., Kidder, S. Q., & Vonder Haar, T. H. (2012). How total precipitable water vapor anomalies relate to cloud vertical structure. Journal of Hydrometeorology, 13(2), 709-721.

Foster, M. J., Heidinger, A., 2013. PATMOS-x: Results from a Diurnally Corrected 30-yr Satellite Cloud Climatology. J. Clim., 26, 414–425.

Frey, R.A., S.A. Ackerman, Y. Liu, K.I. Strabala, H. Zhang, J. Key and X. Wang, 2008: Cloud Detection with MODIS, Part I: Recent Improvements in the MODIS Cloud Mask. J. Atmos. Oceanic Tech., 25, 1057-1072.

Geleyn J. and A. Hollingsworth. An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. Beitr. Phys. Atm., 52:1–16, 1979.

Giorgi, F., Jones, C., and Asrar, G.R., 2009: Addressing climate information needs at the regional level: the CORDEX framework, WMO Bulletin, 58(3), 175-183.

Guignard, A., C.J. Stubenrauch, A.J. Baran, and R. Armante, 2012: Bulk microphysical properties of semi-transparent cirrus from AIRS: a six year global climatology and statistical analysis in synergy with geometrical profiling data from CloudSat-CALIPSO. Atmos. Chem. Phys., 12, 503-525.

Hazeleger W, Severijns C, Semmler T, Stefanescu S, Yang S, Wang X, Wyser K, Baldasano JM, Bintanja R, Bougeault P, Caballero R, Dutra E, Ekman AML, Christensen JH, van den Hurk B, Jimenez P, Jones C, Kållberg P, Koenigk T, MacGrath R, Miranda P, van Noije T, Schmith T, Selten F, Storelvmo T, Sterl A, Tapamo H, Vancoppenolle M, Viterbo P, Willén U (2010) EC-Earth: a seamless earth system prediction approach in action. Bull Am Meteor Soc 91:1357–1363.

Hazeleger, W., X. Wang, C. Severijns, S. Stefanescu, R. Bintanja, A. Sterl, K. Wyser, T. Semmler, S. Yang, B. van den Hurk, T. van Noije, E. van der Linden and K. van der Wiel, 2012 EC-Earth V2.2: description and validation of a new seamless Earth system prediction model. Clim. Dyn., 39, 2611-2629, doi:10.1007/s00382-011-1228-5.

Heidinger, A.K., A.T. Evan, M. Foster, and A. Walther, 2012: A Naïve Bayesian Cloud Detection Scheme Derived from CALIPSO and Applied within PATMOS-x. J. Appl. Meteor. Clim., 51,1129-1144.

Heidinger, A. K., M. J. Foster, A. Walther, and X. Zhao (2014), The Pathfinder atmospheres-extended AVHRR climate dataset, Bull. Am. Meteor. Soc., 95(6), 909-922, doi: 10.1175/BAMS-D-12-00246.1.

Hendricks J., A. Falb, C. J. Stubenrauch, U. Lohmann, and C. Emde, 2010: A Method for comparing properties of cirrus clouds in global climate models with those retrieved from IR sounder satellite observations. Meteorol. Z.19, 577-589.

Hiemstra, P.H., Pebesma, E.J., Twenhöfel, Ch.J.W., Heuvelink, G.B.M., 2009. Real-time automatic interpolation of ambient gamma dose rates from the Dutch radioactivity monitoring network, Computers & Geosciences 35, 1711-1721.

Hilton, F., R. Armante, T. August, C. Barnet, A. Bouchard, C. Camy-Peyret, V. Capelle, L. Clarisse, C. Clerbaux, P.-F. Coheur, A. Collard, C. Crevoisier, G. Dufour, D. Edwards, F. Faijan, N. Fourrié, A. Gambacorta, M. Goldberg, V. Guidard, D. Hurtmans, S. Illingworth, N. Jacquinet-Husson, T. Kerzenmacher, D. Klaes, L. Lavanant, G. Masiello, M. Matricardi, A. McNally, S. Newman, E. Paveli, S. Payan, E. Péquignot, S. Peyridieu, T. Phulpin, J. Remedios, P. Schlüssel, C. Serio, L. Strow, C. J. Stubenrauch, J. Taylor, D. Tobin, W. Wolf, D. Zhou, Hyperspectral Earth Observation from IASI, Bull. Amer. Meteor. Soc., 93, 347-370, doi:10.1175/BAMS-D-11-00027.1, (2012).

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.

Jakob C. and S. A. Klein. The role of vertically varying cloud fraction in the parametrization of microphysical processes in the ECMWF model. Q. J. R. Meteorol. Soc., 125(555):941–965, 1999. doi:10.1002/qj.49712555510.

Karlsson, K.-G., A. Riihelä, R. Müller, J. F. Meirink, J. Sedlar, M. Stengel, M. Lockhoff, J. Trentmann, F. Kaspar, R. Hollmann, and E. Wolters (2013), CLARA-A1: a cloud, albedo, and radiation dataset from 28 yr of global AVHRR data, Atmos. Chem. Phys., 13, 5351-5367, doi: 10.5194/acp-13-5351-2013.

Kay, J. E., Hillman, B. R., Klein, S. A., Zhang, Y., Medeiros, B., Pincus, R., Gettelman, A., Eaton, B., Boyle, J., Marchand, R., and Ackerman, T. P.: Exposing Global Cloud Biases in the Community Atmosphere Model (CAM) Using Satellite Observations and Their Corresponding Instrument Simulators, J. Climate, 25, 5190–5207, doi:10.1175/JCLI-D-11-00469.1, 2012.

Keller, M., O. Fuhrer, J. Schmidli, M. Stengel, R. Stöckli, and C. Schär, 2015: Evaluation of convection-resolving models using satellite data: The diurnal cycle of summer convection over the Alps. Meteorologische Zeitschrift.

Keller. M. The diurnal cycle of alpine summer convection in a convection-resolving model: evaluation with satellite data and sensitivity to atmospheric forcing. 2016. PhD Thesis.

Keller, M. O. Fuhrer, J. Schmidli, M. Stengel, R. St ̈ockli, and C. Schr. Evaluation of convection-resolving models using satellite data: The diurnal cycle of summer convection over the alps. Met. Zeit., 25:165–175, 2016. doi: 10.1127/metz/2015/0715.

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

Kotlarski, S., K. Keuler, O. B. Christensen, A. Colette, M. Déqué, A. Gobiet, K. Goergen, D. Jacob, D. Lüthi, and E. van Meijgaard, 2014: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. Geoscientific Model Development, 7 (4), 1297–1333.

Lauer, A., V. Eyring, M. Righi, M. Buchwitz, P. Defourny, M. Evaldsson, P. Friedlingstein, R. de Jeu, G. de Leeuw, A. Loew, C. J. Merchant, B. Müller, T. Popp, M. Reuter, S. Sandven, D. Senftleben, M. Stengel, M. Van Roozendael, S. Wenzel, U, Willén. 2017 Bechmarking CMIP5 models with a subset of ESA CCI Pahse 2 data using the ESAvalTool. RSE, DOI: 10.1016/j.rse.2016.12.027

Lenderink, G. and Holtslag, A. A. M., 2004. An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. Q.J.R. Meteorol. Soc., 130: 3405–3427. doi: 10.1256/qj.03.117

Mann, H.B., 1945. Nonparametric Tests Against Trend. Econometrica 13, 245–259.

Martin, G. M., Johnson, D. W., and Spice, A.: The Measurement and Parameterization of Effective Radius of Droplets in Warm Stratocumulus Clouds., J. Atmos. Sci., 51, 1823–1842, doi:10.1175/1520-0469(1994)051<1823:TMAPOE<2.0.CO;2, 1994.

McGarragh, G., Poulsen, C., Thomas, G., Povey, A., Sus, O., Schlundt, C., Stapelberg, S., Proud, S., Christensen, M., Stengel, M., and Grainger, R.: The Community Cloud retrieval for CLimate (CC4CL). Part II: The optimal estimation algorithm, submitted to Atmospheric Measurement Techniques Discussions, pp. –, 2017.

Menzel, W.P., R.A. Frey, H. Zhang, D.P. Wylie, C.C. Moeller, R.E. Holz, B. Maddux, B.A. Baum, K.I. Strabala, and L.E. Gumley, 2008: MODIS Global Cloud-Top Pressure and Amount Estimation: Algorithm Description and Results. J. Appl. Meteor. Climl., 47, 1175-1198.

Minnis, P., Q.Z. Trepte, S. Sun-Mack, Y. Chen, D.R. Doelling, D.F. Young, D.A. Spangenberg, W.F. Miller, B.A. Wielicki, R.R. Brown, S.C. Gibson, and E.B. Geier, 2008a: Cloud detection in non-polar regions for CERES using TRMM VIRS and Terra and Aqua MODIS data. IEEE Trans. Geosci. Remote Sens., 46, 3857-3884.

Minnis, P., S. Sun-Mack, D.F. Young, P.W. Heck, D.P. Garber, Y. Chen, D.A. Spangenberg, R.F. Arduini, Q.Z. Trepte, W.L. Smith, Jr., J.K. Ayers,S.C. Gibson, W.F. Miller, V. Chakrapani, Y. Takano, K.-N. Liou, Y. Xie, and P. Yang, 2011: CERES Edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data, Part I: Algorithms. IEEE Trans. Geosci. Remote Sens., 49, 11, 4374-4400.

Mlawer, E.J., S.J. Taubman, P.D. Brown, M.J. Iacono and S.A. Clough, 1997. Radiative transfer for inhomogeneous atmospheres : RRTM, a validated correlated-k model for the longwave. J. Geophys. Res., 102D, 16663-16682

Morcrette, J-J., H. W. Barker, J. N. S. Cole, M. J. Iacono, R. Pincus, 2008: Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System. Mon. Wea. Rev., 136, 4773–4798, doi: http://dx.doi.org/10.1175/2008MWR2363.1

Nam, C., S. Bony, J. L. Dufresne, and H. Chepfer (2012), The 'too few, too bright' tropical low-cloud problem in CMIP5 models, Geophys. Res. Lett., 39, L21801, doi: 10.1029/2012GL053421.

Norris, J. and A. Evan, 2015. Empirical Removal of Artifacts from the ISCCP and PATMOS-x Satellite Cloud Records. J. Atmos. Oceanic Technol., 32, 691–702.

Pincus,R., S. Platnick, S. A. Ackerman, R. S. Hemler, and R. J. P. Hofmann. Reconciling simulated and observed views of clouds: Modis, isccp, and the limits of instrument simulators. J. Climate, 25:4699–4720, July 2012. doi: 10.1175/JCLI-D-11-00267.1.

Reinhardt, T., and A. Seifert, 2006: A three-category ice scheme for LMK. Cosmo Newsletter, 6, 115-120.

Ritter, B., and J.-F. Geleyn, 1992: A comprehensive radiation scheme for numerical weather prediction models with potential applications in climate simulations. Monthly Weather Review, 120, 303-325.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riedi, and R. A. Frey (2003), The MODIS cloud products: algorithms and examples from Terra, IEEE Transactions on Geoscience and Remote Sensing, 41, 2, 459-473, doi: 10.1109/TGRS.2002.808301.

Poulsen, C. A., Watts, P. D., Thomas, G. E., Sayer, A. M., Siddans, R., Grainger, R. G., Lawrence, B. N., Campmany, E., Dean, S. M., Arnold, C., 2011. Cloud retrievals from satellite data using optimal estimation: evaluation and application to ATSR. Atmospheric Meas. Tech. Discuss. 4, 2389–2431.

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

PVIRv4, Product Validation and Intercomparison Report (PVIR) – ESA Cloud_cci, Issue 4, Revision: 1, Date of Issue: 18/04/2017 Available at: http://www.esa-cloud-cci.org/?q=documentation

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

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

Sayer, A.M., C.A. Poulsen, C. Arnold, E. Campmany, S. Dean, G.B.L. Ewen, R.G. Grainger, B.N. Lawrence, R. Siddans, G.E. Thomas, and P.D. Watts, 2011: Global retrieval of ATSR cloud parameters and evaluation (GRAPE): dataset assessment. Atmos. Chem. Phys., 11, 3913-3936.

Schröder, M., Jonas, M., Lindau, R., Schulz, J., and Fennig, K. (2013): The CM SAF SSM/I-based total column water vapour climate data record: methods and evaluation against re-analyses and satellite. Atmospheric Measurement Techniques, 6(3), 765-775.

Seifert, A., and K. Beheng, 2006: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. Meteorology and atmospheric physics, 92, 45-66.

Siebesma, A.P., Soares, P. M. M., and Teixeira, J., 2007: A Combined Eddy-Diffusivity Mass-Flux Approach for the Convective Boundary Layer, J. Atmos. Sci., 64, 1230-1248.

Stengel, M., Mieruch, S., Jerg, M., Karlsson, K.-G., Scheirer, R., Maddux, B., Meirink, J. F., Poulsen, C., Siddans, R., Walther, A., Hollmann, R., 2015. The Clouds Climate Change Initiative: Assessment of state-of-the-art cloud property retrieval schemes applied to AVHRR heritage measurements. Remote Sens. Environ. 162, 363–379.

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, D. G., Meirink, J. F., Feofilov, A., Bennartz, R., Bojanowski, J., 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 Discuss., https://doi.org/10.5194/essd-2017-48, in review, 2017.

Stengel et al., 2017b: Relating tropical changes in sea surface temperature and water vapour to changes in cloud top properties using satellite observations and reanalysis data.

Stengel M., Schlundt C., Willén U. Eliasson S., 2017c: Comparing ERA-Interim clouds with satellite observations using a simplistic satellite simulator. In preparation, to be submitted to ACP.

Stephens, Graeme L., 1990: On the Relationship between Water Vapor over the Oceans and Sea Surface Temperature. J. Climate, 3, 634–645. doi: http://dx.doi.org/10.1175/1520-0442(1990)003<0634:OTRBWV>2.0.CO;2

Stubenrauch, C. J., A. D. Del Genio, and W. B. Rossow, 1997: Implementation of Sub-Grid Cloud Vertical Structure inside a GCM and its Effect on the Radiation Budget. J. Climate, 10, 273-287.

Stubenrauch, C.J., W.B. Rossow, F. Chéruy, N.A. Scott, A. Chédin, 1999: Clouds as seen by Infrared Sounders (3I) and Imagers (ISCCP): Part II) A New Approach for Cloud Parameter Determination in the 3I Algorithms. J. Climate, 12, 2214-2223.

Stubenrauch. C.J., S. Cros. A. Guignard, and N. Lamquin, 2010: A six-year global cloud climatology from the Atmospheric InfraRed Sounder aboard the AQUA Satellite: statistical analysis in synergy with CALIPSO and CloudSat. Atmos. Chem. Phys., 10, 7197-7214.

Stubenrauch, C., Rossow, W., Kinne, S., and the Cloud Assessmet Group, A Project of the World Climate Research Programme Global Energy and Water Cycle Experiment (GEWEX) Radiation Panel, WCRP Report No. 23/2012

Stubenrauch, C. J., W. B. Rossow, S. Kinne, S. Ackerman, G. Cesana, H. Chepfer, L. Di Girolamo, B. Getzewich, A. Guignard, A. Heidinger, B. C. Maddux, W. P. Menzel, P. Minnis, C. Pearl, S. Platnick, C. Poulsen, J. Riedi, S. Sun-Mack, A. Walther, D. Winker, S. Zeng, G. Zhao, 2013: ASSESSMENT OF GLOBAL CLOUD DATASETS FROM SATELLITES: Project and Database initiated by the GEWEX Radiation Panel, Bull. Amer. Meteor. Soc., doi: 10.1175/BAMS-D-12-00117.

Stubenrauch, C. J., A. G. Feofilov, S. E. Protopapadaki, and R. Armante, 2017: Cloud climatologies from the InfraRed Sounders AIRS and IASI: Strengths, Weaknesses and Wpplications, Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-422, in review, 2017.

Sun, Z.: Reply to comments by Greg M. McFarquhar on 'Parametrization of effective sizes of cirrus-cloud particles and its verification against observations'. (October B, 1999,125, 3037 3055), Q. J. Roy. Meteor. Soc., 127, 267–271, doi:10.1002/qj.49712757116, 2001.

Sun, Z. and Rikus, L.: Parametrization of effective sizes of cirrus-cloud particles and its verification against observations, Q. J. Roy. Meteor. Soc., 125, 3037–3055, doi:10.1002/qj.49712556012, 1999.

Sus, O., Jerg, M., Poulsen, C., Thomas, G., Stapelberg, S., McGarragh, G., Povey, A., Schlundt, C., Stengel, M., and Hollmann, R.: The Community Cloud retrieval for CLimate (CC4CL). Part I: A framework applied to multiple satellite imaging sensors, submitted to Atmospheric Measurement Techniques Discussions, pp. –, 2017.

Taylor, K. E., R. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, Bull. Amer. Meteor. Soc., 93, 485-49.

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.

Theunert, F., and A. Seifert, 2006: Simulation studies of shallow convection with the convection-resolving version of the DWD Lokal-Modell. COSMO-Newsletter, 6, 121-128.

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models, Mon. Weather Rev., 117, 1779–1799.

Tiedtke, M., 1993. Representation of clouds in large-scale models. Mon. Wea. Rev., 121, 3040-3061

Tompkins, A.M., K. Gierens and G. Rädel, 2007. Ice supersaturation in the ECMWF Integrated Forecast System. Q.J.R. Meteorol. Soc., 133: 53-63.

Undén, P. et al., 2002: HIRLAM-5 Scientific Documentation, 144 pp. Available from SMHI, S-601 76 Norrköping, Sweden

URDv2, User Requirements Document (URD) – ESA Cloud_cci, Issue 2, Revision: 0, Date of Issue: aa/bb/cccc Available at: http://www.esa-cloud-cci.org/?q=documentation

van den Hurk, B., P. Siegmund, A. Klein Tank (Eds), J. Attema, A. Bakker, J. Beersma, J. Bessembinder, R. Boers, T. Brandsma, H. van den Brink, S. Drijfhout, H. Eskes, R. Haarsma, W. Hazeleger, R. Jilderda, C. Katsman, G. Lenderink, J. Loriaux, E. van Meijgaard, T. van Noije, G.-J. van Oldenborgh, F. Selten, P. Siebesma, A. Sterl, H. de Vries, M. van Weele and R. de Winter en G.-J. van Zadelhoff, KNMI'14: Climate Change scenarios for the 21st Century - A Netherlands perspective. KNMI publication: WR-2014-01, 26/5/2014, pp120.

van Meijgaard, E., L.H. van Ulft, G. Lenderink, S.R. de Roode, L. Wipfler, R. Boers, R.M.A. Timmermans, 2012: Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe, Climate changes Spatial Planning publication: KvR 054/12, ISBN/EAN 978-90-8815-046-3, pp 44, http://climexp.knmi.nl/publications/FinalReport_KvR-CS06.pdf

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

Walther, A., and A. Heidinger, 2012: Implementation of the Daytime Cloud Optical and Microphysical Properties Algorithm (DCOMP) in PATMOS-x. J. Appl. Meteor. Climatol., 51, 1371-1390.

Weusthoff, T., F. Ament, M. Arpagaus, and M. W. Rotach, 2010: Assessing the benefits of convection-permitting models by neighborhood verification: Examples from MAP D-PHASE. Monthly Weather Review, 138, 3418-3433.

Winker, D.M., M.A. Vaughan, A.H. Omar, Y. Hu, K.A. Powell, Z. Liu, W.H. Hunt, and S.A. Young, 2009: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. J. Atmos. Oceanic Techn., 26, 2310–2323, doi:10.1175/2009JTECHA1281.1.

Annex A

In this section we provide a comparison between the output of the simplistic satellite simulator (SIMFERA, see Section 4.2) and the Cloud_cci satellite simulator as described in Section 5.3.2. For this purpose both simulator were applied to the same model data: EC-Earth (Section 5.3.1) output for January 2007. A brief summary on the approach of both simulators (for each core cloud variable) is given in Table A-1.

Figures Figure ***A-1*** and Figure ***A-2*** present the monthly mean maps and monthly zonal means for both simulator versions for the core cloud properties. There seems a generally good agreement for cloud fraction (total and low/mid/high layers) and cloud phase. While also CTP still shows a reasonable agreement the differences are found increased for the other variables CER, COT, LWP and IWP.

Table A-1 *Summary of the simulator approach for each core cloud variable for SIMFERA and the Cloud_cci satellite simulator.*

	SIMFERA	Cloud_cci satellite simulator
CFC (total, low/mid/high)	Removing all columns with COT < COT_{thresh} (COT_{thresh}=0.2 in this case)	As SIMFERA
CPH	Each column gets one phase assigned (liq or ice) based on LWC <> IWC of the uppermost cloud layer. Top cloud layers with COT< COT_{thresh} are removed breforehand.	Liquid or ice phase is determined from the uppermost cloud/s down to 1 optical depth. The phase is decided based on the ratio of liquid and ice optical depth. Compared to SIMFERA, all clouds in the column above the COT=1-altitude are taken into account.
CTP	Each column gets one cloud top pressure assigned based on layer top pressure of the uppermost cloud layer. Top cloud layers with COT< COT_{thresh} are removed breforehand.	The cloud top is determined from the attitude where COT=1 from space. All clouds in the column above the COT=1-altitude are taken into account.
CER (liq/ice)	Each column is assigned a liquid CER or ice CER depending on phase assignment. The CER is calculated as in ERA-Interim radiation: Liquid clouds: CER is calculated fowollowing the method of Martin et al. (1994) as a function of LWC of the uppermost cloud layer in the column. In addition the number of cloud condensation nuclei (over land: 300; over sea: 100) is used Ice clouds: ice crystal effective radius is calculated as a function of	CER is derived from the best fit to the estimated model TOA reflectivity. 1) The simulated reflectivity in each column is estimated based on the model effective radius and optical depth vertical profiles from the model. It is estimated by the two stream reflectance approach calculated using the single scattering albedo (SSA), and asymmetry parameter (ω0) read from CC4CL-LUTs as a function of CER for liquid and ice.

	temperature and IWC of the uppermost cloud layer in the column based on Sun and Rikus (1999), which has been revised by Sun (2001). Top cloud layers with COT< COT_{thresh} are removed breforehand.	2) From these LUTs of SSA and ω0, a corresponding table of reflectivity as a function of CER is calculated. 3) The simulated CER is found by minimizing the reflectivity calculated in (1) against the tabled reflectivity from (2) for the simulated CPH. This is exactly the same CER-simulation-approach as used by the MODIS simulator in COSP
COT (liq/ice)	The cloud optical thickness (COT) per layer is obtained by the method of Han et al. (1994) $COT = \frac{3}{4}\frac{CWP \cdot Q_{ext}}{CER \cdot \rho}$ where CWP is the sum of LWP and IWP. Q_{ext} denotes the extinction coefficient, which is assumed to be 2 for water and 2.1 for ice. The density is set to 1 g/m^3 for water and 0.9167 g/m^3 for ice. The COT per column is then the sum of the layer COTs. Top cloud layers with COT< COT_{thresh} are removed breforehand.	The same as SIMFERA except the no clouds are excluded from the column.
LWP	Entire CWP of a column becomes LWP when cloud top phase in that the column is liquid. Top cloud layers with COT< COT_{thresh} are removed breforehand.	The same as SIMFERA except the no clouds are excluded from the column.
IWP	Entire CWP of a column becomes IWP when cloud top phase of the column is ice. Top cloud layers with COT< COT_{thresh} are removed breforehand.	The same as SIMFERA except the no clouds are excluded from the column.

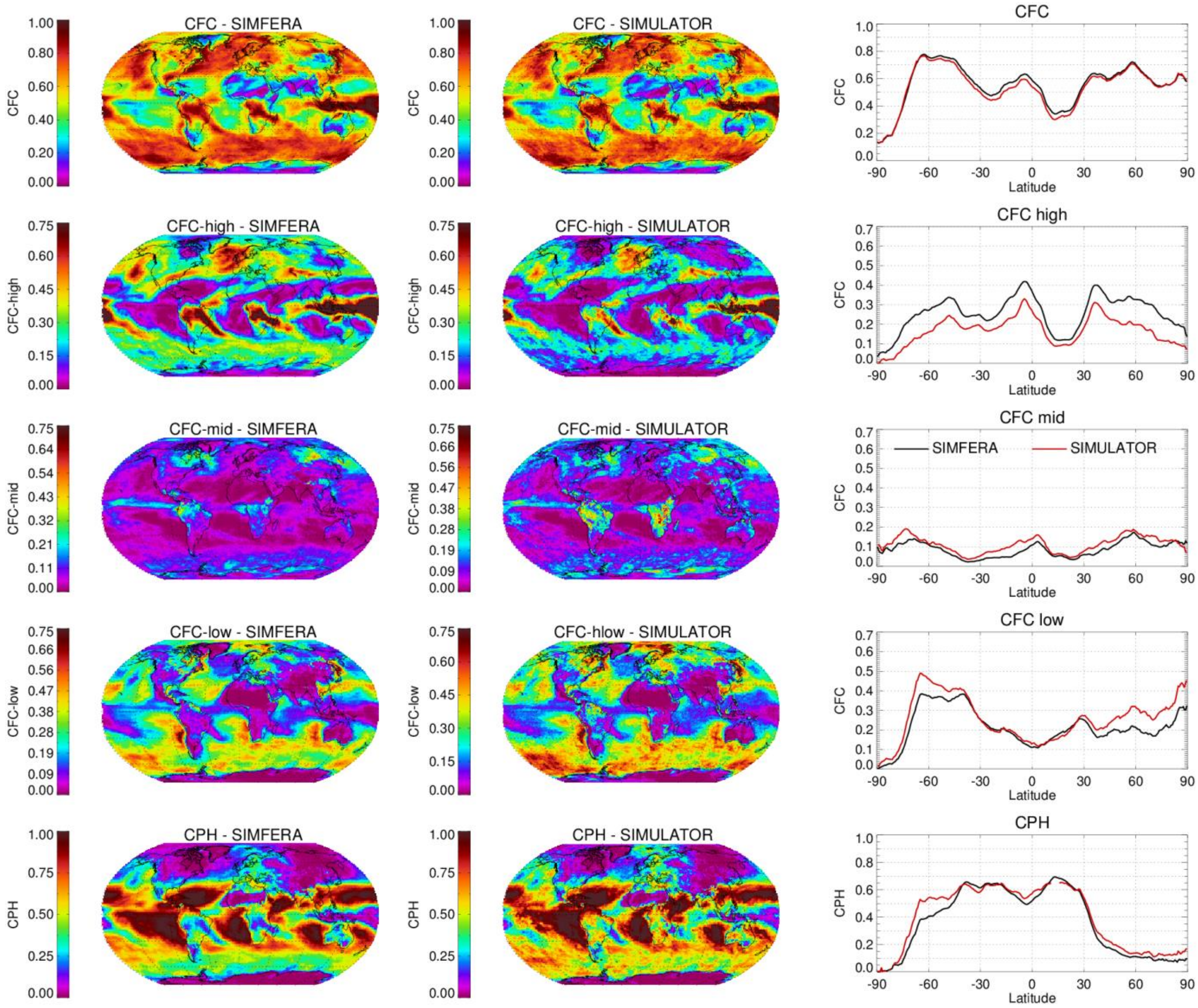


Figure A-1 *Monthly mean values for (rows from top to bottom): total cloud fraction (CFC), cloud fraction of low/mid/high clouds (CFC-low, CFC-mid, CFC-high) and cloud phase (CPH). Data shown is for January 2007. Left column: SIMFERA output, centre column: Cloud_cci simulator output, right column znoal mean plots of both.*

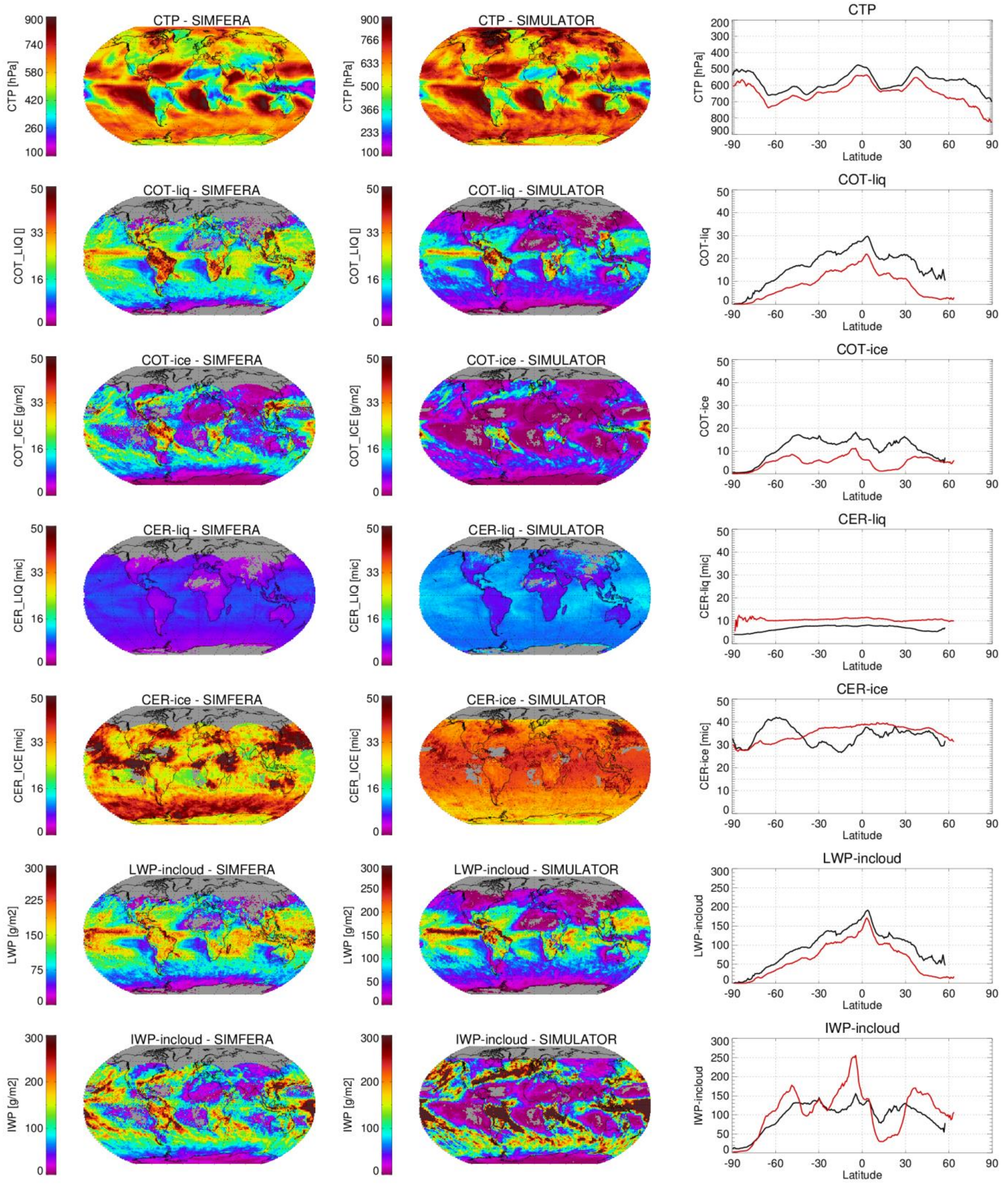


Figure A-2 *Monthly mean values for (rows from top to bottom): cloud top pressure (CTP), cloud optical thickness for liquid/ice clouds (COT-liq/COT-ice), cloud effective radius for liquid/ice clouds (CER-liq/CER-ice), in-cloud liquid water path (LWP-incloud) and in-cloud ice water path (IWP-incloud). Data shown is for January 2007. Left column: SIMFERA output, centre column: Cloud_cci simulator output, right column znoal mean plots of both.*