Satellite Application Facility on Climate Monitoring

Scientific Report

Final validation of CM-SAF cloud products derived from MSG/SEVIRI data

Reference Number: SAF/CM/DWD/KNMI/SMHI/SR/CLOUDS-ORR/3

Issue/Revision Index: 1.0

Date: 12/01/2007

Document Signature Table

	Name	Function	Signature	Date
Author	Hartwig Deneke	CM-SAF scientist		12/01/07
	Sheldon Johnston	CM-SAF scientist		12/01/07
	Max Reuter	CM-SAF scientist		12/01/07
	Rob Roebeling	CM-SAF scientist		12/01/07
	Anke Tetzlaff	CM-SAF scientist		12/01/07
	Werner Thomas	CM-SAF scientist		12/01/07
	Erwin Wolters	CM-SAF scientist		12/01/07
Approval	Karl-Göran Karlsson	WP manager clouds		12/01/07
	Annegret Gratzki	CM-SAF science co-ordinator		
Release	Martin Werscheck	Project manager		
Eumetsat Approval				

Document Change Record

Issue/Revision	Date	DCN No.	Changed Pages/Paragraphs
1.0	12/01/06	SAF/CM/DWD/KNMI/SMHI/SR/CLOUDS-ORR/3	First version

Distribution List

Internal Distribution	
Name	**No. Copies**
DWD CM-SAF document archive	1
KNMI	2
SMHI	2

External Distribution		
Company	**Name**	**No. Copies**
EUMETSAT		1
Review board		4

Table of Contents

List of Tables

List of Figures

List of Acronyms

AVHRR	Advanced Very High Resolution Radiometer
BT	Brightness Temperature
CALIOP	Cloud-Aerosol LIDAR with Orthogonal Polarization
CALIPSO	Cloud-Aerosol LIDAR and Infrared Pathfinder Satellite Observation
CDOP	Continuous Development and Operations Phase
CESAR	Cabauw Experimental Site for Atmospheric Research
CFC	Cloud Fractional Cover
CM-SAF	Satellite Application Facility on Climate Monitoring
CLWP	Cloud Liquid Water Path
COT	Cloud-Optical Thickness
CTx	Cloud-top parameters (height, temperature, pressure)
CTY	Cloud Type
DAK	Doubling-Adding KNMI radiative transfer model
DEM	Digital Elevation Model
DWD	Deutscher Wetterdienst
EUMETCAST	Eumetsat data distribution system
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FUB	Free University of Berlin (Germany)
GERB	Global Earth Radiation Budget
GME	Global NWP model of DWD
IOP	Initial Operations Phase
ISCCP	International Satellite Cloud Climatology Project
ITCZ	Inner-Tropical Convergence Zone
KNMI	Royal Dutch Meteorological Institute
KSS	Kuiper Skill Score
LWP	Liquid Water Path
METOP	Meteorological Operational Platform
MODIS	Moderate Resolution Imaging Spectrometer
MSG	Meteosat Second Generation
MWR	Microwave Radiometer
NRT	Near-Real-Time
NWCSAF	Satellite Application Facility on Nowcasting and Very Short-range Forecasting
PPS	Polar Platform System
ORR	Operations Readiness Review
RAM	Random Access Memory
RMI	Royal Meteorological Institute of Belgium
SCIAMACHY	Scanning Imaging Absorption Spectrometer for Atmospheric Cartography
SEVIRI	Spinning Enhanced Visible and Infrared Imager
SIRR	Software Integration Readiness Review
SIVVRR	System Integration Verification and Validation Readiness Review
SG	Steering Group
SMHI	Swedish Meteorological and Hydrological Institute
URD	User Requirements Document
UMP	User Manual Products
WP	Work Package
WUI	Web User Interface
WV	Water Vapour

1 Introduction

1.1 Purpose

This document contains the final scientific validation report of CM-SAF cloud products based on MSG/SEVIRI data of the full visible MSG disk.

The document follows the EUMETSAT guidelines.

1.2 Scope

This document describes the final validation of CM-SAF cloud products (CFC, CTY, CTx, COT, CLWP, CPH) derived from MSG/SEVIRI data against ground-based measurements (synoptic data, radar data) for several months in 2004 and 2006. In addition, MSG/SEVIRI retrieval results were compared against corresponding MODIS and CALIPSO results for a limited number of cases.

1.3 Applicable Documents

Table 4.1 *List of Applicable Documents*

Reference	Title	Code
[AD.1.]	CM-SAF User Requirements Document	SAF/CM/DWD/URD/1
[AD.2.]	CM-SAF Software Requirements Document	SAF/CM/DWD/SRD/1
[AD.3.]	CM-SAF Interface Control Document	SAF/CM/DWD/ICD/1
[AD.4.]	CM-SAF Configuration Management Plan	SAF/CM/DWD/SCMP/1
[AD.5.]	CM-SAF Software Integration Verification and Validation Plan	SAF/CM/DWD/SIVVP/1
[AD.6.]	CM-SAF Detailed Design Document	SAF/CM/DWD/DDD/1
[AD.7.]	Scientific Report, ORR V2 Validation of CM-SAF Cloud Products using MSG/SEVIRI Data	SAF/CM/DWD/SMHI/KNMI/SR/CLOUDS/2
[AD.8.]	Report of the Operation Readiness Review for the System Version 2 (ORR 2)	EUM/PPS/REP/05/0061
[AD.9.]	Initial validation of CM-SAF cloud products based on MSG/SEVIRI data	SAF/CM/DWD/SR/CLOUDS/1
[AD.10.]	Validation of CM-SAF cloud products using MSG/SEVIRI data	SAF/CM/KNMI/SIVVRR_V3/1
[AD.11.]	User Manual for the PGE01-02-03 of the SAFNWC / MSG: Scientific part	SAF/NWC/IOP/MFL/SCI/SUM/01

1.4 Document status

The document is based on the project status of CM-SAF as reported by work package managers and project scientists at the beginning of January 2007.

1.5 Document Overview

This document describes the validation tasks that were performed to assess the performance of CM-SAF cloud products derived from MSG/SEVIRI radiance data over the full visible MSG disk. The present document is structured in the following sections:

This section contains an introduction to the document, providing scope and purpose and documentation references. Section 2 describes the overall status and provides information about validation data sets. Section 3 briefly describes the SEVIRI retrieval methods while sections 4-9 contain detailed descriptions of tasks that were performed. The overall summary and conclusions are given in section 10.

2 Validation data sets

This section is divided into the following paragraphs:

- Introduction
- Overall status
- Validation period
- Product list and resolution
- Cloud observation instruments
- Measurement campaigns
- Other satellite observations

2.1 Introduction

The goal of the validation activity is to provide validation results for several months of data for which cloud products are available from different measurement systems. A major limiting factor is the availability of independent measurements of cloud parameters from e.g., cloud radars and microwave radiometers at reasonable temporal and spatial distribution. Especially this latter requirement limits the validation of SEVIRI-based cloud parameters over sea surfaces and over land surfaces outside Europe. To partly compensate the lack of ground-based validation data we included results of two satellite – satellite intercomparison studies using MODIS and CALIPSO data.

2.2 Overall status

CM-SAF passed successfully the ORR-V2 in July 2005 and started operational processing of MSG-based products in September 2005. The first version was confined to the initial baseline area (30°N to 80°N, 60°W to 60°E) while the enhancement to the full visible MSG disk is subject of this review. An initial validation of “full disk” results of one month was already performed for the intermediate SIVVRR-V3 review that took place in October 2006. The retrieval of macrophysical cloud parameters (CFC, CTY, CTx) is based on version MSGv1.2 of the NWC-SAF that was released in 2005. It is an upgraded version of MSGv1.1 which was validated during ORR-V2.
Cloud microphysical properties (COT, CWP, CPH) were derived using Cloud Physical Products (CPP) software version 2.0 which was released by KNMI in spring 2006. Major improvements (compared to software used at ORR-V2) are the inclusion of monthly surface albedo maps (rather than fixed values) and recalculated look-up tables of simulated top-of-atmosphere SEVIRI radiances for different cloud models.

Official coud products over the baseline area are available from 1st September 2005 onwards (subject of ORR-V2) and new cloud products over the MSG full disk (subject of this validation study) are now available from July 2006 onwards. In addition, four months in 2004 (May, July, October, December) were processed on the full disk.

The software is installed on the DWD mainframe system that is based on Power P5 CPU’s (1.9 GHz) and operated under IBM/AIX 5.3. Processing time per MSG slot on a single CPU is larger than 15 minutes (the sampling rate of MSG/SEVIRI) but slots can be operated in parallel. Several hours (CPU time) per month are required to process monthly mean products and the monthly mean diurnal cycle of selected products (see Table 2.1). Results are available on a slightly reduced “full disk” (3636 × 3636 pixels, edges cutted) which is due to hardware limits of the IBM Power architecture for 32-bit programmes and the amount of RAM needed.

2.3 Validation period

2.3.1 Comparison period for the full disk

Validation results are available for the months May, July, October, December 2004 and from July - October 2006. We used again data of 2004 to allow a comparison of retrieval results over the same geographic area and the same time period as analysed for ORR-V2 in July 2005.

2.3.2 Validation for the CloudNET sites

Validation results are available for the period May 2004 until April 2005 for the CloudNET sites of Cabauw in the The Netherlands, Chilbolton in the UK and Palaiseau in France. The CPH, COT and CWP products were retrieved at a 15 minutes resolution. For validation of CPH around the Chilbolton site, a four month period (May-August 2004) was used. The CPP products were retrieved with version 2.1 that was run at KNMI, which is a more recent version than version 2.0 that was operated at DWD during the Initial Operational Phase (IOP). The KNMI retrieved CPP products were used because these products were available at a much higher temporal resolution (15 minutes) than the DWD products (1 hour). Note that the decision to use the products of KNMI version 2.1 instead of DWD version 2.0 was made following the suggestion of the SIVVRR_V3 review board.

From version 2.0 to version 2.1, we have modified the calibration settings and the line to band conversion coefficients. These modifications have an impact on the retrieval results. In order to verify the correspondence between the cloud properties retrieved with version 2.1 at KNMI, and version 2.0 at DWD, we compared the COT and CLWP products of both versions for 7 June 2004, 11:45 UTC over Europe. For this observation Figure 2.1 and 2.2 present the frequency distributions and the relationship between DWD and KNMI retrieved COT and CWP values, respectively. The frequency distributions of both versions are very similar. The median COT and CWP values show that the modifications of the calibration settings and the line to band conversion coefficients results in a slight increase of both the COT (~5%) and CWP (~10%) values. The density plots show that the retrievals of both versions are well correlated (corr ~ 0.99). The largest differences between version 2.0 and version 2.1 occur at the higher COT (> 20) and CWP (> 100) values. These values, however, have a very low frequency of occurrence. The increased difference between the two versions with increasing COT and CWP is explained by the non-linear relationship between cloud reflectance and COT. This is confirmed by Roebeling et al. (2005), who showed that small errors in radiative transfer simulations at 0.6 and 1.6 μm can affect retrievals of cloud optical thickness and effective radius strongly. They found for optically thick clouds (COT> 60) that differences of 3% in the simulated reflectance result in differences of up to 30% in retrieved COT. For CPH the differences between version 2.0 and 2.1 are marginal. A comparison of CPH retrievals from both versions show a slight decrease in the percentage of ice clouds from 33.4% in version 2.0 to 32.8% in version 2.1.

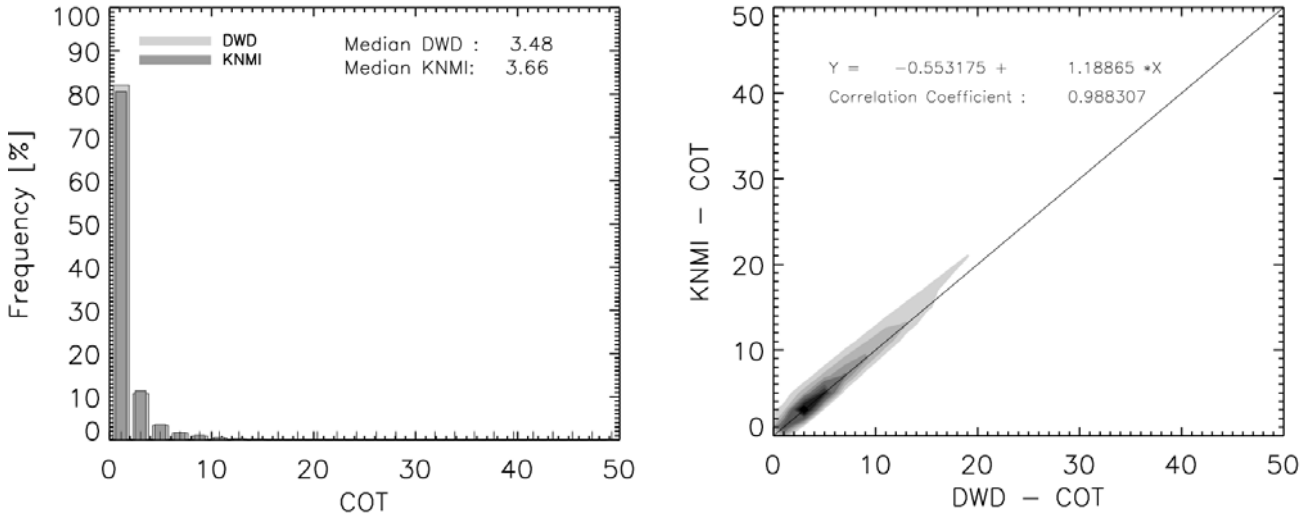

Figure 2.1 *Frequency distribution and density plot of COT values retrieved with CPP Version 2.0 at the DWD and CPP Version 2.1 at KNMI for 7 June 2004, 11:45 over Europe.*

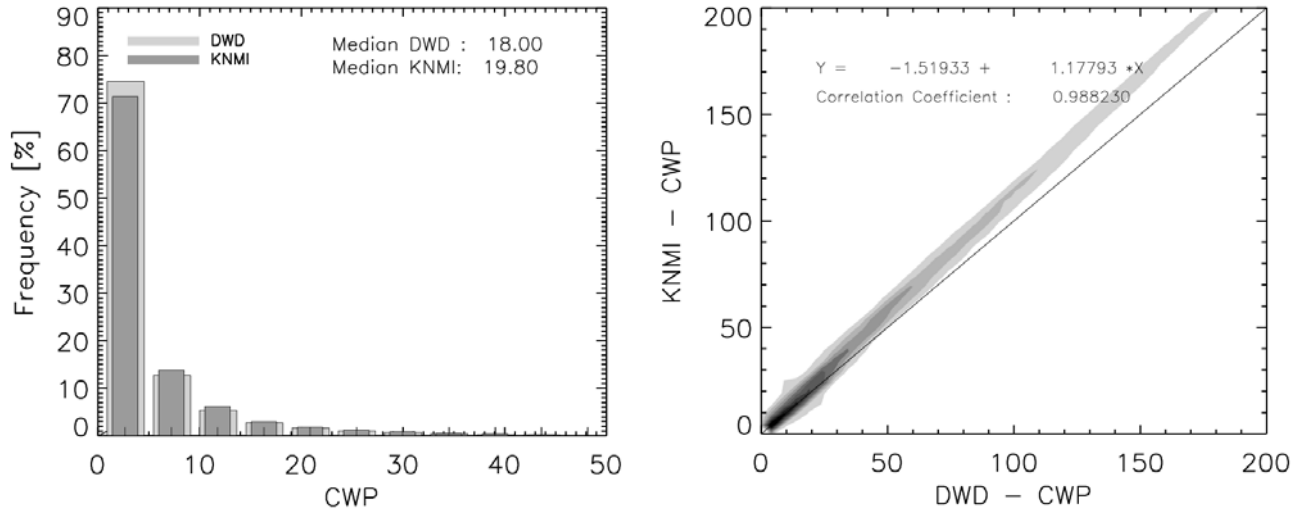

Figure 2.2 *Frequency distribution and density plot of CWP values retrieved with CPP Version 2.0 at the DWD and CPP Version 2.1 at KNMI for 7 June 2004, 11:45 over Europe.*

2.3.3 Comparison of MODIS and SEVIRI cloud properties

The comparison of CPP CPH, COT, and CWP with the corresponding MODIS Level 2 products was done over three 10^ox10^o areas, labelled tropical land (TL, 7^o-17^o N, 5^o W-5^o E), tropical ocean (TO, 5^o S-5^o N, 5^o W-5^o E), and subtropical ocean (STO, 20^o-30^0 S, 0^o-10^o E). Their position is indicated in Figure 2.3. The tropical land and ocean areas are characterized by several climate regions, i.e. the Inter-Tropical Convergence Zone (ITCZ) that shifts over this area during the boreal summer, the Western African monsoon convection near the coastline and semi-arid and desert-like in the Northern parts of the TL area. The subtropical ocean area is dominated by a persistent high-pressure area in which stratocumulus clouds prevail.

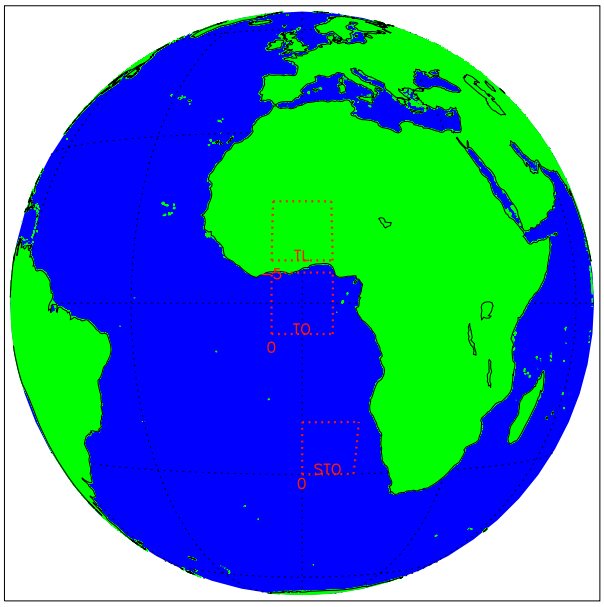

Figure 2.3 *The Earth as seen by METEOSAT-8. The three areas used for the SEVIRI-MODIS inter-comparison are denotes by dotted red boxes. TL=Tropical Land, TO=Tropical Ocean, and STO=Sub-Tropical Ocean. See text for the box boundaries.*

For ORR_V3, the comparison of MODIS and SEVIRI cloud properties was done for 21, 24, and 26 images during July 2006 for the TO, TL, and STO areas, respectively. The SEVIRI data were processed at DWD, using version 2.0 of the CM-SAF Cloud Physical Properties algorithm. The MODIS cloud phase, cloud optical thickness and cloud liquid water path fields were extracted from the Level-2 cloud products of the collection-5 dataset from Terra and Aqua. The MODIS granules that collocated best with the observation area were compared to the SEVIRI observations closest to the MODIS overpass time. Because only hourly SEVIRI images (at X:45 UTC) were available the observation times of SEVIRI and MODIS differed up to 30 minutes. The average time difference between SEVIRI and MODIS images was ~ 12 minutes. The SEVIRI and MODIS retrieved cloud properties were re-projected to a Mercator projection of similar grid size. To reduce the collocation errors, the MODIS images were shifted within a 15x15 pixel box to find the maximum correlation with the SEVIRI images. Finally, SEVIRI and MODIS pixels were selected with viewing zenith angles smaller than 50° and scattering angles 120° - 130° and 140° - 175°. For the STO area, pixels with viewing zenith angles smaller than 50° were included to obtain sufficient data, ceteris paribus. Scattering angles close to 180° and 137° were excluded to eliminate backscatter and rainbow effects, respectively. Logarithmic averaging was used to calculate the mean COT during the observation period and account for the quasi-logarithmic relationship between cloud albedo and COT, using the following equation:

$$\overline{\tau_{vis}} = \exp\left(\frac{\sum_1^n \log(\tau_{vis}(i))}{n}\right)$$

where $\overline{\tau_{vis}}$ is the logarithmically averaged COT, $\tau_{vis}(i)$ is the COT value of an individual observation, and *n* is the number of observations. Subsequently, a bin size of 2 was used to construct the frequency and frequency of differences distributions.

For LWP, linear averaging was used and all data values were accumulated at bin size 15. For comparison of SEVIRI CPH with MODIS cloud phase, pixels labelled “undefined” by the MODIS

algorithm were excluded. As a result, SEVIRI water or ice was compared to MODIS water, ice, and mixed phase.

2.4 Product list and resolution

The following table provides an overview about available cloud parameters derived from MSG/SEVIRI data.

Table 5.1 *Cloud products of CM-SAF (MSG/SEVIRI only) and corresponding temporal and spatial resolution, accuracy (bias) and precision (standard deviation), as laid down in AD 1.*

Product	Acronym	Resolution/accuracy/precision						
		Spatial	Temporal					
			Daily		Monthly		MMDC (Monthly Mean Diurnal Cycle)	
			accur. [%]	prec.	accur. [%]	prec.	accur. [%]	prec.
Fractional cloud cover	CFC	15 km	✓, 10		✓, 10		✓, 10	
Cloud type	CTY	15 km	✓, 20		✓, 20		✓, 10	
Cloud top temperature, height, and pressure	CTT, CTH, CTP	15 km	✓, 10		✓, 10		✓, 10	
Cloud optical thickness	COT	15 km	✓, 2		✓, 8		✓, 8	
Cloud phase (water)	CPH	15 km	✓, 5		✓		✓	
Cloud phase (ice)	CPH	15 km	✓, 30		✓		✓	
Cloud water path[1]	CLWP	15 km	✓, 10		✓, 15		✓, 15	
1 CLWP units in gm^{-2}								

Target accuracies of the cloud phase product were defined separately for water and ice clouds after the validation study. Precision remained undefined in AD 1 but we propose some appropriate numbers as result of the validation study in section 10.3.

2.5 Cloud observation instruments

Several instruments are available at European measurement sites Cabauw, The Netherlands (51.97 °N, 4.93 °E), Chilbolton, UK (51.14 °N, 1.44 °W) and Palaiseau, France (48.71 °N, 2.21 °E) to measure cloud parameters from ground. These are briefly described in the following subsections. Each site is equipped with radar, LIDAR and a suite of passive instrumentation. The use of active instruments (LIDAR and cloud radar) resulted in detailed vertical profiles of important cloud parameters, which cannot be derived from current satellite sensing techniques. At the CloudNET sites of Chilbolton and Palaiseau, dual-channel MicroWave Radiometers (MWRs) are operated. The radiometer at Chilbolton measures at 22.2 and 28.8-GHz, while the radiometer at Palaiseau measures at 24 and 37-GHz (DRAKKAR). More information on the CloudNET project can be found on http://www.met.rdg.ac.uk/radar/cloudnet/.

2.5.1 LIDAR

The LIDAR is a high power laser system that emits short pulses of light. The back-scattered radiation from atmospheric particles and molecules is recorded and the distance between the scattering event and the LIDAR instrument is measured. For cloud detection the LIDAR operates in a

wavelength region where the scattering originates predominantly from particles with radii between 1 and 3 µm which are typical for cloud water droplets.

2.5.2 Cloud radar

The cloud radar measures two cloud physical parameters: the distance between the instrument and the cloud particles and the velocity of the moving particles. The cloud radar is most sensitive to scattering of particles with high effective radii but the maximum sensitivity in this respect depends also on the operating frequency.

The two radars used here (at Cabauw and Chilbolton) operate at a frequency of 35 GHz, corresponding to a wavelength of 8.6 mm, which makes it sensible to cloud droplets in the range of 0 to about 200 µm (Rayleigh scattering regime). The cloud radar is especially suited to measure the particle volume and the moving velocity of ice crystals (Donovan et al., 1998). The travel time of the signal from the radar to a target and back causes a frequency difference between the transmitted and the received signal. This frequency difference can be obtained by multiplying the transmitted signal and the received signal (mixing) together with a low pass filter applied to the mixed signal. The resulting signal is called the “beat signal”. The frequency of the beat signal is then directly related to the distance of the target. Further, cloud radar measures the Doppler shift of moving particles. By measuring this shift, the velocity of the targets can be determined. The velocities are obtained by measuring the phase shifts of succeeding sweeps.

2.5.3 Microwave radiometer

Passive microwave radiometers provide brightness temperature measurements at different frequencies that are further used to infer the Integrated Water Vapour (IWV) and the Liquid Water Path (LWP). The measurements of the two-channel MWRs operated at the CloudNET sites are used to retrieve simultaneously LWP and IWV (Löhnert and Crewell, 2003). The microwave brightness temperatures at two frequencies have distinct atmospheric absorption characteristics. The 22 GHz brightness temperature provides mainly information about atmospheric water vapour, whereas the 30 GHz brightness temperature provides mainly information on the cloud liquid water path. The LWP retrieval algorithm is based on a statistical relationship between the observed brightness temperatures and LWP. This relationship is derived from simulated brightness temperatures (from radiative transfer model calculations) for different LWP values for given profiles of atmospheric temperature and humidity. Because of uncertainties in the instruments calibration and variations in the atmospheric profiles, the LWP retrievals during cloud free conditions can differ from zero, and may become both positive and negative. Marchand et al. (2000) have shown that using profile information from actual radio-soundings can significantly reduce the uncertainties due to the natural variability of atmospheric profiles. The correction method that is applied at the CloudNET stations determines the instrument calibration and atmospheric profile coefficients from the MWR brightness temperatures that are observed during clear sky periods. During these periods, which are identified from independent ceilometer observations, the LWP values are close to zero and the instrument calibration and atmospheric profile coefficients can be derived. During periods of cloud cover these coefficients are interpolated between consecutive clear sky observations (Gaussiat et al., 2006). The retrieval of LWP from MWR measurements is strongly disturbed by rainfall, since the instrument antenna or radiometer can be covered by water droplets or a thin water layer. Moreover, none of the MWRs are sensitive to ice clouds since ice crystals do not contribute to the MWR radiances at the probed frequencies.

According to Crewell and Löhnert (2003) the accuracy of LWP retrievals varies between 15 and 35 g m^{-2}. Note that these accuracies were derived from instrumental specifications and are of theoretical nature, reflecting only to a minor degree the normally distributed radiometric noise.

The two-channel MWRs that are operated at Chilbolton and Palaiseau have an expected accuracy of about 30 g m^{-2} (Crewell and Löhnert, 2003).

2.5.4 Pyranometer

The thermoelectric pyranometer is an instrument that can accurately measure broadband hemispherical irradiance in the solar spectral region. Pyranometers are mainly used at meteorological stations to measure the downwelling solar irradiance at the surface. It can either measure the global or the diffuse irradiance. The accuracy of standard pyranometers should be about 3% according WMO standards. However, mean solar irradiances may be underestimated by 3 – 8% due to pollution on the instrument dome, imperfect thermal cooling of the detectors and long maintenance intervals (Deneke, 2002).

2.6 Measurement campaigns

Ground-based observations that were gathered during two measurement campaigns were used to validate CM-SAF cloud parameters. The campaigns are described briefly in the following:

2.6.1 CloudNET

CloudNET was an EU-funded research project that aimed to develop and implement cloud remote sensing synergy algorithms by using data obtained quasi-continuously at three remote sensing stations. The project ran from 1st of April 2001 to 1st April 2005 and the above-mentioned three ground sites were involved. Each site is equipped with radar, LIDAR and a suite of passive instrumentation.

Active instruments (LIDAR and cloud radar) allow retrieving detailed vertical profiles of important cloud parameters, which cannot be derived of the same quality from current satellite sensing techniques. For these already existing cloud remote sensing stations (CRS-stations) a network was initially established and operated for a two years period (2003-2004) and a joint data archive was built up. The observations were used to evaluate four operational numerical models and to demonstrate the benefit of such operational network for various applications, among them validation of satellite measurements. At the time of writing (December 2006), CloudNET is still continuing operations on a voluntary basis at Cabauw.

2.6.2 CESAR

The Cabauw Experimental Site for Atmospheric Research (CESAR) in The Netherlands consists of a large set of instruments to study the atmosphere and exchange processes with the land surface. The CESAR objectives are the monitoring of long-term changes of atmospheric parameters, studies of atmospheric and land surface energetic processes for climate modelling, validation of space-borne observations and development and implementation of new measurement techniques. Monitoring operations started in 2000 and will continue till 2010 and even further. The site is equipped with remote sensing instruments, in situ tower instruments and in situ ground instruments. Among others, especially instruments like LIDAR, radar, ceilometer and pyranometer are operated which were partly used for the present study.

2.7 Other satellite observations

2.7.1 MODIS

The MODIS cloud product combines infrared and visible techniques to determine both physical and radiative cloud properties. Daily global Level 2 data are provided. Cloud-particle phase (ice

vs. water, clouds vs. snow), effective cloud-particle radius, and cloud optical thickness are derived using the MODIS visible and near-infrared channel radiances. Cloud-top temperature, height, effective emissivity, phase (ice vs. water, opaque vs. non-opaque), and cloud fraction are produced by the infrared retrieval methods both day and night at 5 × 5 km resolution. MODIS cloud products are initially retrieved at 1 km spatial resolution but are provided as 5 × 5 km average results. Finally, the MODIS cloud product includes the cirrus reflectance in the visible at the 1-km-pixel resolution, which is useful for removing cirrus scattering effects from the land-surface reflectance product. There are cloud data sets from both the Terra and the Aqua platform. (http://modis-atmos.gsfc.nasa.gov/index.html).

2.7.2 CLOUDSAT/CALIPSO

The Cloud-Aerosol LIDAR and Infrared Pathfinder Satellite Observation (CALIPSO) satellite mission was launched in April 2006 and first data became available in August 2006. CALIPSO provides detailed profile information about cloud and aerosol particles and corresponding physical parameters. CALIPSO's payload includes a polarization-sensitive active LIDAR (CALIOP), a passive Infrared Imaging Radiometer (IIR), and visible Wide Field Camera (WFC). CALIPSO is part of the A-train and follows a sun-synchronous orbit 700 km above the ground, the ascending node crossing the equator at 13:43 local time. There is a 16-day repetition cycle.

CALIOP measures the backscatter intensity at 1064 nm while two other channels measure the orthogonally polarized components of the backscattered signal at 532 nm.

The WFC is a fixed, nadir-viewing imager with a single spectral channel covering the 620-270 nm region, selected to match band 1 of the MODIS instrument on Aqua.

The IIR is a nadir-viewing, non-scanning imager having a 64 km by 64 km swath with a pixel size of 1 km. The CALIOP beam is nominally aligned with the centre of the IIR image. The instrument provides measurements at three channels in the thermal infrared window region at 8.7 mm, 10.6 mm, and 12.0 mm. These wavelengths were selected to optimize joint CALIOP/IIR retrievals of cirrus cloud emissivity and particle size.

More information about CALIPSO can be found at:

http://www.nasa.gov/mission_pages/calipso/main/index.html

3 MSG/SEVIRI cloud products: retrieval methods

This section is divided into the following paragraphs:

- Retrieval of cloud fractional cover
- Retrieval of cloud type
- Retrieval of cloud-top height parameters
- Retrieval of cloud physical properties

The calculation of the MSG/SEVIRI macrophysical cloud parameters (CFC, CT, CTx) of a single slot is performed by the MSG algorithm package MSGv1.2 developed by the NWC-SAF. A detailed description can be found in AD 11 and references therein. The computation of cloud products is sequential, i.e. the cloud fractional cover is derived first and is input to the cloud type and the cloud-top parameters retrieval.

3.1 Cloud Fractional Cover – CFC

The cloud mask retrieval algorithm is based a multi-spectral threshold technique where thresholds are scene-dependent dynamically adjusted (Derrien and LeGléau, 2005). The thresholds are based on pre-calculated radiative transfer simulations stored in look-up tables. Essential further input parameters are actual geographical data (e.g. land use, topography, etc.) and Numerical Weather Prediction (NWP) analyses. The latter are taken from the DWD GME model (see Majewski et al., 2002) with a temporal resolution of 3 hours and a spatial resolution of about 40 km. There are 40 atmospheric layers between ground and the topmost layer at 0.1 hPa.

3.2 Cloud Type – CT

The main objective of the NWC-SAF cloud type product (CT) is to provide a detailed cloud analysis. The NWC-SAF product distinguishes between 15 cloud classes. The CT product of CM-SAF is less detailed and clouds are grouped as follows: fractional clouds, semitransparent clouds, high, medium and low clouds (including fog) for all the pixels identified as cloudy in a scene. The CT algorithm is a sequential threshold algorithm applied to pixels. It uses the pre-computed cloud mask and spectral and textural features which are derived from the multispectral satellite images and scene-dependent (dynamic) thresholds.

3.3 Cloud-top height/pressure/temperature – CTx

The CTx product contains information on the cloud top pressure, temperature and height for all pixels identified as cloudy in the satellite scene.

A first step is to simulate SEVIRI radiances and brightness temperatures in channels 6.2µm, 7.3µm, 13.4µm, 10.8µm, and 12.0µm for clear and cloudy scenes using actual NWP temperature and humidity profiles which were interpolated in time. In order to reduce the computational effort, this is done for boxes of 32 × 32 SEVIRI pixels using the closest NWP grid column values in space. For very low, low, and medium to high thick clouds the cloud-top pressure is then retrieved at pixel resolution and corresponds to the best fit between the simulated and the measured 10.8µm brightness temperatures. High and semi-transparent clouds are treated as described in Schmetz et al. (1993) and Menzel et al. (1982), respectively, if the first correction fails. Cloud-top temperature and height are then retrieved from the known cloud-top pressure and atmospheric profiles from NWP data sets.

3.4 Cloud Physical Properties – COT/CWP/CPH

The principle of methods to retrieve cloud physical properties is that the reflectance of clouds at a non-absorbing wavelength in the visible region (0.6 or 0.8 μm) is strongly related to the optical thickness and has very little dependence on particle size, whereas the reflectance of clouds at an absorbing wavelength in the near-infrared region (1.6 or 3.8 μm) is primarily related to particle size. Note that the retrieval of particle size from near-infrared reflectances is weighted towards the upper part of the cloud (Platnick, 2001). The average penetration depth of reflected photons is affected by the amount of absorption, which depends on wavelength, particle type and size. The reflectance at 1.6 μm is found to be mainly a function of particle size for clouds with an optical thickness higher than about 8, whereas the reflectance at 3.8 μm is more suited for the retrieval of cloud particle size for thin clouds (COT > ~2) (Rosenfeld, 2004; Watts *et al.*, 1998). However, the 3.8 μm channel has a number of disadvantages that may lead to significant errors: (1) the radiance observed at 3.8 μm consists of both reflected solar radiance and thermal emitted radiance, (2) the signal to noise ratio is lower due to the approximately 4 times lower solar irradiance at 3.8 μm than at 1.6 μm, and finally (3) because the 3.8 μm retrievals represent the particle size of the upper part of the cloud these retrievals will be less representative for radiative transfer in optically thick clouds (Feijt *et al.*, 2004).

The Doubling Adding KNMI (DAK) radiative transfer model is used to generate the Look Up Tables (LUTs) of simulated cloud reflectances. DAK is developed for line-by-line or monochromatic multiple scattering calculations at UV, visible and near infrared wavelengths in a horizontally homogeneous cloudy atmosphere using the doubling-adding method (De Haan *et al.*, 1987; Stammes, 2001). The clouds are assumed to be plane-parallel and embedded in a multi-layered Rayleigh scattering atmosphere. The algorithm we utilize to retrieve cloud physical properties is based on reflectances at visible (0.6 μm) and near-infrared (1.6 μm) wavelengths. In the KNMI version of the CPP algorithm the pixel is assumed cloudy if the cloudmask identified the pixel as cloud contaminated, using a rudimentary cloud mask (Roebeling *et al.*, 2006c), and the observed reflectance at 0.6 μm is higher than the simulated clear sky reflectance over the observed surface. Note that in the in the DWD version of the CPP algorithm the NWCSAF algorithm is used for cloud detection. One-year of MODIS white-sky albedo data is used to generate the map of surface albedos. The white-sky albedo represents the bi-hemispherical reflectance in the absence of a direct component, which is a good estimate of the surface albedo below optically thick clouds. The COT and particle size are retrieved for cloudy pixels in an iterative manner, by simultaneously comparing satellite observed reflectances at visible (0.6 μm) and near-infrared (1.6 μm) wavelengths to LUTs of RTM simulated reflectances for given optical thicknesses and particle sizes (Watts *et al.*, 1998; Jolivet and Feijt, 2005). Table 3.1 summarizes the governing characteristics of the cloudy atmosphere, together with information about intervals of cloud properties and viewing geometries used for the DAK simulations. During the iteration the COT values that are retrieved at the 0.6 μm channel are used to update the retrieval of particle size at the 1.6 μm channel. This iteration process continues until the retrieved cloud physical properties converge to stable values. The interpolation between cloud physical properties in the LUTs is done with polynomial interpolation for COT values and linear interpolation for particle size. For optically thin clouds (COT < 8) the retrieved particle size values are unreliable. For these clouds an assumed climatological averaged effective radius is used that is 8 μm for water clouds and 35 μm for ice clouds, which is close to the values used by Rossow and Schiffer (1999). To obtain a smooth transition between assumed and retrieved effective radii a weighting function is applied to the effective radius retrievals of clouds with COT values between zero and eight. The retrieval of cloud thermodynamic phase is done simultaneously with the retrieval of COT and particle size. The phase “ice” is assigned to pixels with a Cloud Top Temperature (CTT) lower than 265 K for

which the 0.6 μm and 1.6 μm reflectances correspond to DAK simulated reflectances for ice clouds. The remaining cloudy pixels are considered water clouds.

SCIAMACHY spectra are used to calculate the conversion coefficients between the simulated line reflectances of DAK and the channel reflectances of SEVIRI at 0.6 and 1.6 μm. These spectra are convoluted with the SEVIRI spectral response functions to obtain SEVIRI channel reflectances, which are divided by the DAK reflectances to obtain the line-to-band conversion coefficients (Roebeling *et al.*, 2006a).

The droplet effective radius (r_e) is the adequate parameter to represent the radiative properties of a size distribution of water particles that is given by (Hansen and Hovenier, 1974):

$$r_e = \frac{\int_0^\infty r^3 n(r)\,dr}{\int_0^\infty r^2 n(r)\,dr}$$

where *n(r)* is the particle size distribution and *r* is the particle radius. This definition is used to retrieve the effective radius for water clouds between 1 and 24 μm. For ice clouds we assume a homogeneous distribution of C0, C1, C2 and C3 type imperfect hexagonal ice crystals from the COP data library of optical properties of hexagonal ice crystals (Hess *et al.*, 1998). Knap *et al.* (2005) demonstrated that these crystals can be used to give adequate simulations of total and polarized reflectances of ice clouds.

The CTT is calculated from 10.8 μm brightness temperatures and the emissivity of the cloud (ε_λ). The ε_λ is calculated from the cloud optical thickness at wavelength λ (τ_λ) with the following equation (Minnis *et al.*, 1993):

$$\varepsilon_\lambda = 1 - \exp\left(\frac{-\tau_\lambda}{\cos\theta}\right)$$

where $\cos\theta$ is the cosine of the viewing zenith angle. The (absorbing) cloud optical thickness in the thermal infrared (τ_{tir}) is related to the (scattering) cloud optical thickness in the visible (τ_{vis}). This relationship depends on particle size and particle thermodynamic phase. For large water and ice particles τ_{tir} is about $0.5\,\tau_{vis}$.

The CWP is computed from the retrieved cloud optical thickness at wavelength at 0.6 μm (denoted as τ_{vis}) and droplet effective radius (r_e) as follows (Stephens, 1978):

$$CWP = \frac{2}{3}\tau_{vis} r_e \rho_l$$

where ρ_l is the density of liquid water. For ice clouds, the CWP is retrieved for imperfect hexagonal ice crystals with an assumed effective radius of 6, 12, 26 and 51 μm.

Table 3.1 *Properties of the cloudy atmosphere and the surface that are used for the radiative transfer calculations to generate the LUTs.*

Parameter	*Settings*	
Atmospheric vertical profiles of pressure temperature and ozone	Midlatitude summer [a]	
Aerosol model	none	
Cloud height	1000 - 2000 m	
Solar zenith angle (θ_0)	0 - 75°	
Viewing zenith angle (θ)	0 - 75°	
Relative azimuth angle (ϕ)	0 - 180°	
Cloud Optical Thickness	0 - 128	
Surface albedo (ocean)	MODIS white sky reflectance	
Surface albedo (land)	MODIS white sky reflectance	
	water clouds	*ice clouds*
Cloud particle type	Spherical water droplet	Imperfect hexagonal ice crystal [b]
Cloud particle size	1 –24 μm	C0: L= 10, D= 8 μm [c]
		C1: L= 30, D=20 μm [c]
		C2: L= 60, D=44 μm [c]
		C3: L=130, D=82 μm [c]
Size distribution	Modified gamma	-
Effective variance (v_e)	0.15	-

[a] The midlatitude summer atmosphere model was taken from Anderson et al. (1986).
[b] The imperfect hexagonal crystals are obtained from Hess et al. (1998) and have a distortion angle of 30°.
[c] L and D are the length and the diameter of the hexagon, respectively.

4 Validation of cloud fractional cover - CFC

This section is divided into the following paragraphs:

- Product description
- Validation task
- Validation against ground-based measurements
- Comparison against MODIS observations
- Comparison against CALIPSO observations
- Summary

4.1 Product description

The MSG/SEVIRI cloud mask distinguishes between cloud-free, cloudy, and partially cloud pixels of a scene (and also the additional categories snow/ice and unprocessed pixels). The cloud mask is used in subsequent processing steps as input. The cloud mask for individual scenes is used to calculate the fractional cloud cover (CFC) in 15 km grid resolution (the resolution for official CM-SAF products). The CFC product is primarily available as daily and monthly mean products. Furthermore, the monthly mean diurnal cycle and instantaneous results are available, the latter on special request.

4.2 Validation task

This validation report focuses on the performance of the MSG cloud mask of the full visible MSG disk. We compared data of four months in 2004 (May, July, October, December) and four in 2006 (July to October) against available synoptic observations. We used only manually operated land stations and ship observations but excluded buoy measurements due to the known limited performance of such automated measurements. The geographical distribution of stations on the MSG disk is depicted in Figure 4.1 together with the overlaying cloud mask (white) for a single slot (14th July 2004, 14:15 UTC). As expected, the majority of stations is located in Northern mid-latitudes over land while there is a lack of ground-based measurements in large parts of Africa but also in the visible parts of eastern South-America. The unbalanced distribution of stations is quantitatively gathered in Figure 4.2.

Note, that we have not used the operational daily and monthly mean average CFC product in this validation study. The temporal resolution of synoptic observations is comparably low (typically 6 obs/day against 24 obs/day) and a comparison of daily averages as is would therefore be of limited use. Instead we calculated daily averages from those MSG slots closest to synoptic observation times. These were subsequently used to generate monthly mean products. These products are available at the same spatial resolution as the CM-SAF standard product which is 15 × 15 km^2. Also the ground-based measurements needed to be subsumed first as daily and monthly mean products. We used synoptic data only from those stations that reported at least 6 measurements per day at 20 days per month. As a consequence, ship measurements were mostly filtered out.

4.3 Validation against ground-based measurements

When synoptic observations are compared to satellite measurements the strongly differing observation techniques have to be properly considered. While the satellite instrument is downward looking with, in the case of SEVIRI, a latitude-dependent footprint size between 3 and 15 km, the ground-based observer is upward-looking, reporting one observation representative for the visible

part of the sky at the position of the observer. Additionally, different quantities are reported: SEVIRI observations are divided into 4 classes (clear sky, cloudy, fractional clouds, snow/ice on the ground) the synoptic observation provides the cloud fraction in octa. Here, 0 octa stands for clear sky, while 8 octa stands for completely overcast sky. As soon as a cloud is visible, 1 octa has to be reported while as soon as the sky is visible between clouds, at most 7 octa are reported. Between 2 and 6 octa, the observation should represent the actual cloud fraction above the synoptic station.

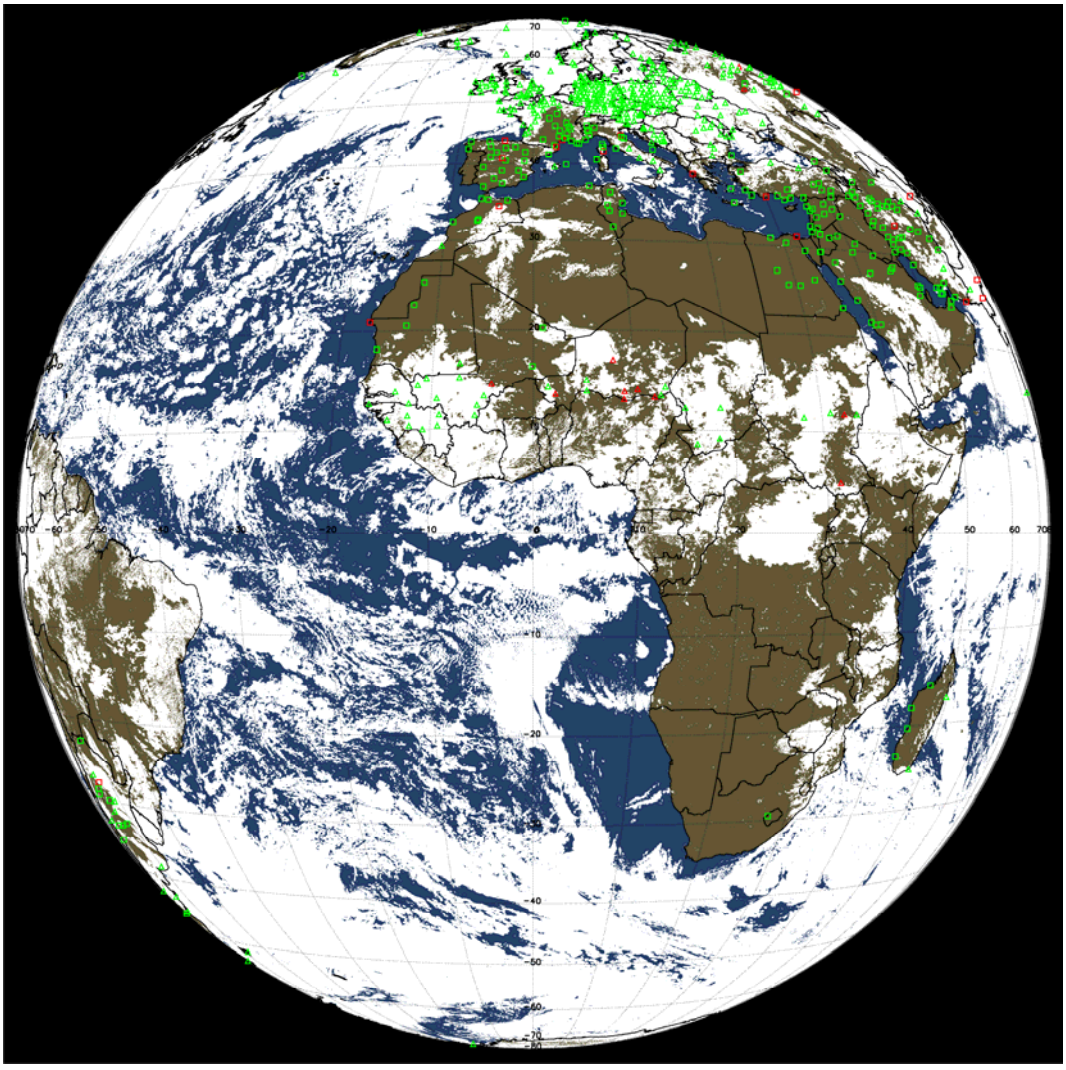

Figure 4.1 *Geographical location of synoptic stations and retrieved cloud mask (white-shaded areas) on 14th July 2004, 14:15 UTC. Green symbols in the left panel denote agreement between satellite and synoptic observations and red symbols represent stations where results disagree.Triangles denote cases of cloudy synoptic observations while squares denote cloud-free observations.*

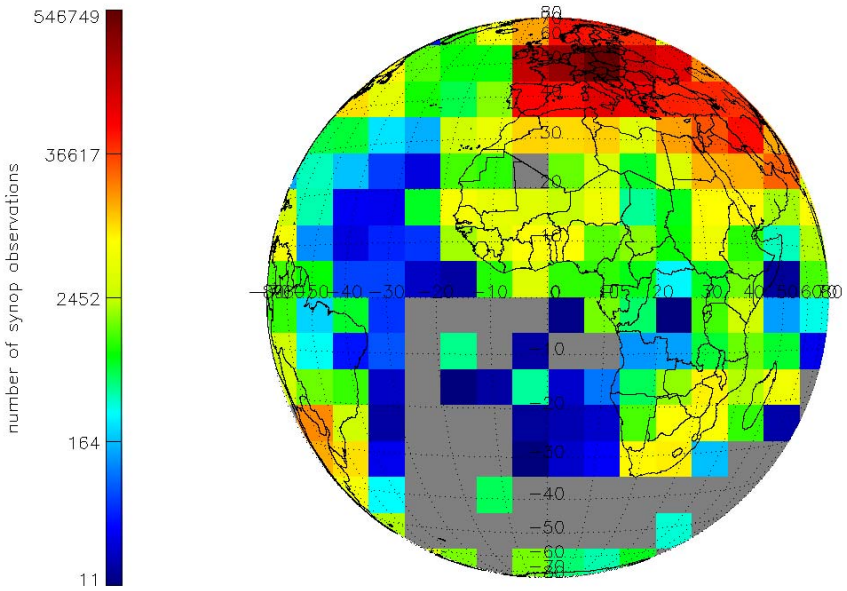

Figure 4.2 *Number of synoptic stations per square used for the SEVIRI cloud mask validation. Squares represent 232 × 232 SEVIRI pixels. The figure shows quantitatively the known geographically unbalanced distribution of synoptic stations on the visible MSG full disk.*

We performed the validation of instantaneous results, the daily and monthly means and the mean monthly diurnal cycle of the cloud fractional cover. While instantaneous results can relatively easy be compared by a temporal (and spatial) collocation, more effort is required for the other quantities. The SEVIRI standard products cannot be taken because corresponding independent measurements are not available (see also section 4.2). We therefore first needed to calculate daily and monthly means for those pixels and ground-based measurements being close together in time and space.

4.3.1 Mathematical methods

The Kuiper skill score (Hanssen and Kuipers, 1965) is a statistical measure that allows to determine the probability that a predicted event occurs, relative to its casual occurrence. Here, we apply it to the satellite measurements (the predicted value) relative to ground-based observations (the synoptic data). We determine a so-called contingency matrix which contains the estimates that observations synop/sat are cloud-free/cloud-free, cloud-free/cloudy, cloudy/cloud-free, and cloudy/cloudy.

Table 4.1: *Contingency matrix of satellite and synoptic observations.*

		Satellite	
	scenario	cloud-free (cf)	cloudy (cc)
Synop	Cloud-free (cf)	a	b
	Cloudy (cc)	c	d

Note, that the contingency matrix contains only results from unambiguous synoptic observations which are 0,1,7, and 8 octa. Then, the Kuiper skill score (*Kss*) is calculated as follows:

$$Kss = \frac{a \cdot d - c \cdot b}{(a+b) \cdot (c+d)} \text{ where } -1 \leq Kss \leq 1$$

while the *hit rate* which is a measure for the proportion of correct measurements (predictions) reads:

$$hitrate = \frac{a+d}{a+b+c+d} \text{ where } 0 \leq hitrate \leq 1$$

In contrast to the *hit rate* the *kss* is well-suited to estimate the skill of quantities which are not symmetrically distributed, as e.g. the cloud fractional cover (cloudy scenes occur more frequently than cloud-free scenes).

We define entries of Table 4.2 and Table 4.3 following in section 4.3.2 using entries of the contingency matrix as follows:

- Count cf/cc = number of all synoptic reports with 0, 1, 7, or 8 octas
- sat_hits_cf_syn = a / (a + b) sat clear sky confirmed by synop
- sat_hits_cc_syn = d / (c + d) sat cloudy confirmed by synop
- syn_conf_cf_sat = a / (a + c) cloud-free synop confirms cloud-free sat
- syn_conf_cc_sat = d / (b + d) cloudy synop confirms cloudy sat
- mean synop = average CFC of all synops
- bias = mean sat CFC – mean syn CFC using all synoptic records (also 2,3,4,5,6 octas)

Thus, the bias was determined using all records while the Kuiper skill score only relies on totally cloudy and cloud-free pixels.

Moreover, the comparison was not only done for just all synoptic stations. Instead, we further distinguish land and sea stations, stations along the coast lines and in mountainous terrain but also daytime and night-time measurements and observations during twilight conditions. This allows a more detailed view on the performance of the satellite measurements. Results for CFC are summarized in Table4.2 and Table 4.3.

The categories day, twilight, and night are distinguished by ranges of the sun zenith angle (sza): sza (day) < 85°, 85° < sza (twilight) < 90°, 90° < sza (night). “Mountain” stations are higher than 2000 m a.s.l. (to group stations which can easily be surrounded by snow-covered surfaces). The class “Coast” is defined for synoptic stations and pixels closer than 3 km to the coast line (≈ 1 SEVIRI pixel).

The previous validation study (AD 7) was performed only for the baseline area (30°N to 80°N, 60°W to 60°E) and left open the question about the validity of a correction of the cloudiness of partially cloudy pixels (labelled fractionally cloudy or cloud-contaminated by the basic cloud mask method) outside the area of interest. We therefore repeated this approach (named bias correction) in the current study and show both uncorrected and bias-corrected results. As defined in AD 7 the cloud fraction above a synoptic station was calculated from the remaining MSG pixels using:

$$CFC_{MSG} = \frac{\sum cloudy + f_{frac} \sum fractionalclouds}{\sum allpixels}$$

The tuning factor or bias correction factor f_{frac} controls the quantitative contribution of fractional clouds to the total SEVIRI cloud cover of pixels. It is further discussed in the following section while detailed results are presented in subsequent sections.

4.3.2 Validation of instantaneous SEVIRI results

The different temporal sampling of the measurements has to be considered. Synoptic observations are either made hourly or in 3 to 6 hours intervals, respectively. SEVIRI measurements are routinely available every 15 minutes. Since all synoptic measurements are done at the full hour, the SEVIRI slot starting 15 minutes before was taken for comparison. Scanning starts at the South Pole thus effectively reducing the time difference between satellite and synoptic measurements to less than 10 minutes. For all synoptic observations, the SEVIRI pixel from the closest time slot and closest to the synoptic station was considered. Note that for the instantaneous results we compare results of single SEVIRI pixels with synoptic records rather than the standard CFC product which is sampled in 15 × 15 km^2 grid boxes.

Analysis of partially cloudy pixels

The tuning factor (fractional cloud cover of partially cloud pixels) was introduced in the validation report of the ORR-V2 (AD 7). It turned out that the bias between satellite and ground-based measurements could be minimised assuming a cloud fraction of 75 % (0.75) for such pixels with sub-pixel cloudiness. This result was basically confirmed in the initial validation study of the SIVVRR-V3 (AD 9). The following Figure 4.3 shows the impact of the tuning factor with respect to the satellite-based cloud coverage:

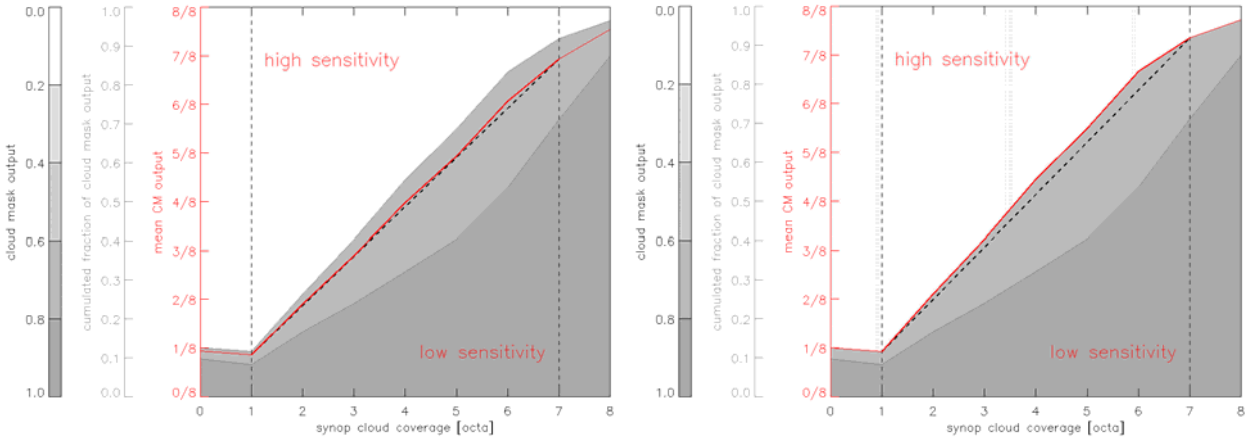

Figure 4.3 *Intercomparison of cloud coverage results with (left panel) and without a bias correction factor (right panel) of 0.75. Synoptic observations are on the x-axis and satellite observations are on the y-axis. Observe that presented results are only based on those cases when collocated SEVIRI pixels had synop-reported cloudiness of values 0,1,7 and 8, octa. Thus, the estimated mean fractional cloud cover from satellite (denoted mean CM output in the figure) and the corresponding mean synop cloud coverage represents a subset of all available data. See text for further explanation.*

The dotted line below the red curve represents the ideal curve (i.e. a linear dependency where we have a perfect matching between synop and satellite estimations) of the cloud mask intercomparison while the red line above is the result of the comparison. Clearly, the ideal (dotted) line and the red line are closer together in the left panel of Figure 4.3 where the tuning factor was set to 0.75 rather than 1. The medium grey-shaded area represents all the cases with SEVIRI fractional clouds present. This area is largest for the synoptic cloud coverage being between 4 and 7 octa which indicates that the “true” cloudiness of partially cloudy pixels is not 1. In other words, the sub-pixel cloudiness is overestimated if a partially cloudy pixel is simply considered as being fully cloud covered.

Results without bias correction

We first performed our analysis without further tuning of the SEVIRI cloud mask. Detailed results for e.g., different underlying ground and daytime and night-time measurements can be found in Table 4.2. Special emphasis is with entries "weighted" in Table 4.2 and later in Table 4.3 which contain results of an analysis that tried to harmonise the unbalanced distribution of synoptic stations

- over land and sea,
- over different regions on the visible full disk (many stations in Europe, few elsewhere)
- and with regard to the local mean time of observations (more measurements during daytime)

Simply spoken the method addresses a higher weight to underrepresented stations over sea and outside Europe (and for night-time measurements) and results of the satellite/synop comparison are therefore more representative for the full disk as it would be if stations are equally distributed.

Table 4.2 *Detailed results of comparison of CFC results for eight months (see section 4.2) without bias correction for fractional clouds from SEVIRI, and synop observations (count all = number of all synop reports, count cf/cc = number of only clear/cloudy synop reports, mean synop = mean CFC synop, sat hits cf syn = probability that cloud-free sat is confirmed by synop, sat hits cc syn = probability that cloudy sat is confirmed by synop, syn conf cf sat = probability that synop confirms cloud-free sat, syn conf cc sat = probability that synop confirms cloudy sat, hit = hitrate, Kss = kuiper skill score, bias = bias (mean sat – mean synop). Coloured results are discussed in the text below.*

scenario	count all	count cf/cc	mean synop	sat hits cf syn	sat hits cc syn	syn conf cf sat	syn conf cc sat	hit	Kss	bias
overall	3682740	2304745	0.535	0.878	0.942	0.920	0.910	0.914	0.820	0.044
weighted	1.E+06	575242	0.539	0.802	0.872	0.818	0.860	0.843	0.675	0.056
land	3047112	1932234	0.534	0.882	0.941	0.920	0.913	0.916	0.823	0.043
sea	150302	82670	0.610	0.748	0.961	0.884	0.904	0.900	0.709	0.109
coast	485326	289841	0.517	0.872	0.942	0.926	0.897	0.910	0.813	0.028
day	2008448	1171056	0.555	0.915	0.955	0.931	0.945	0.939	0.871	0.038
night	1489594	1018698	0.502	0.838	0.932	0.919	0.861	0.887	0.770	0.061
twilight	148698	114991	0.587	0.930	0.889	0.815	0.960	0.903	0.819	-0.025
mountain	39386	22937	0.375	0.835	0.912	0.958	0.695	0.858	0.747	0.087

From results in Table 4.2 we can draw the following conclusions:

1. The overall agreement of CFC derived from satellite and synoptic observations is comparably good. The bias is around 0.04 (scenario "overall") and the Kuiper skill score is almost 0.82.

2. The known different sensitivity of the satellite-based CFC retrieval over land and sea surfaces is confirmed by our analysis. The main reason can be seen in column "sat hits cf syn" where an entry of 0.748 over sea (relative to 0.882 over land) indicates that the satellite observations overestimate the synoptic CFC observations over sea. Consequently, the bias is more than two times higher than for observations over land surfaces (0.109 versus 0.043).

3. As expected, comparison of daytime measurements gives better results than night-time measurements which is presumably due to the additional information from solar channels and the lower frequency of temperature inversions that hamper the detection of clouds during night (probability that synoptic confirms cloudy satellite measurements = 0.861 only). Note however that also night-time synoptic measurements are of lower quality under certain conditions, for example if semi-transparent cirrus clouds are present which typically remain undetected from ground-based observations.

4. The performance under twilight conditions is comparably good (although the low negative bias of -0.025 indicates some increase in the amount of undetected clouds) while records of stations in mountainous terrain compare not very well with satellite measurements (bias = 0.087). Here, it is again the low probability that synoptic observations confirm a fully cloudy sky as seen by the satellite sensor which again is a hint that satellite measurements overestimate the “true” cloudiness. Satellite data overestimate the synoptic records which can partly be an effect of snow covered scenes and misclassified pixels.

5. Validation results for coastal stations and over mountainous terrain, although satisfying, may be biased by the geographical distribution, thus by the viewing geometry. The European area where the majority of such stations is located is always seen under moderately high observation zenith angles.

6. The observed bias of overall results (0.044) is clearly lower than corresponding user requirements (10%).

Weighted results are not as good as “overall” results (lowest *Kss* of only 0.675) which is a hint that the accumulation of reliable stations over land in the Northern midlatitudes is the driver of the overall good performance of our comparison. This can be seen also in Figure 4.4 where the bias of CFC retrievals is shown as a function of the geographic position. Best performance is found for Europe and the Mediterranean and North African regions (bias low, Kuiper skill scores high) while the positive bias increases with the satellite observation angle (increase towards the edges).
A negative bias is found for the tropical regions and over adjacent sea surfaces which does not necessarily mean that the SEVIRI retrieval is less accurate here. In the tropics, we believe that also the increasing differences in the geometrical viewing conditions for the satellite and surface observations contributes to the found differences. . In the equatorial part of Africa SEVIRI pixels represent smaller areas on Earth compared to the case over Europe and northern Africa while synoptic measurements are still representative for a large area which here will be more frequently cloud-contaminated. The near zero satellite viewing angle also increases the chances of finding cloud-free spots compared to cases with more slanted view. It is also well-known that the surface observation in cases of dominantly convective cloud cover tends to overestimate the sky cover of clouds (holes in between vertically extending cloud elements are not seen by the observer at a slanted view).

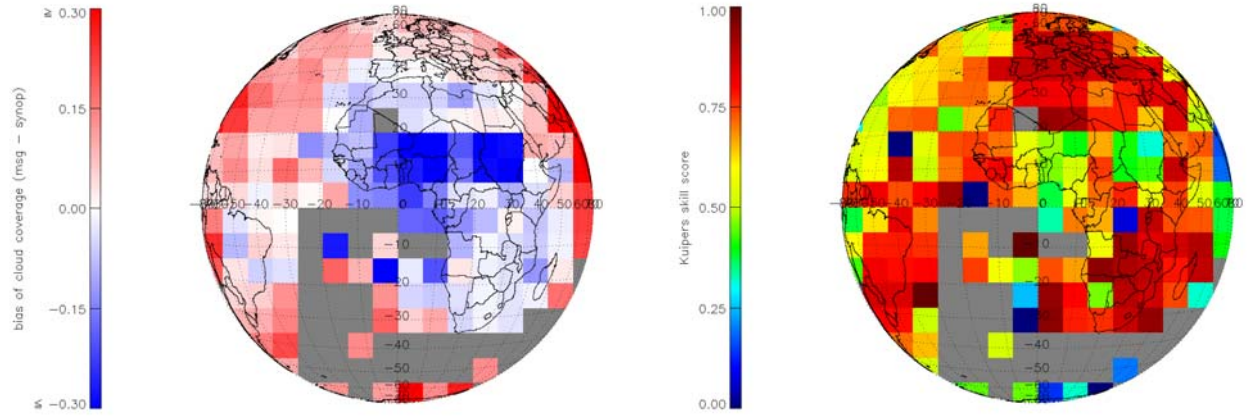

Figure 4.4 *Bias (left panel) and Kuiper skill score (right panel) of satellite measurements of CFC relative to synoptic measurements*

We performed also a more detailed analysis of the performance of results as a function of observation and solar zenith angles, respectively, geographic latitude, local time, season (day-of-year), and the cloud-base height.

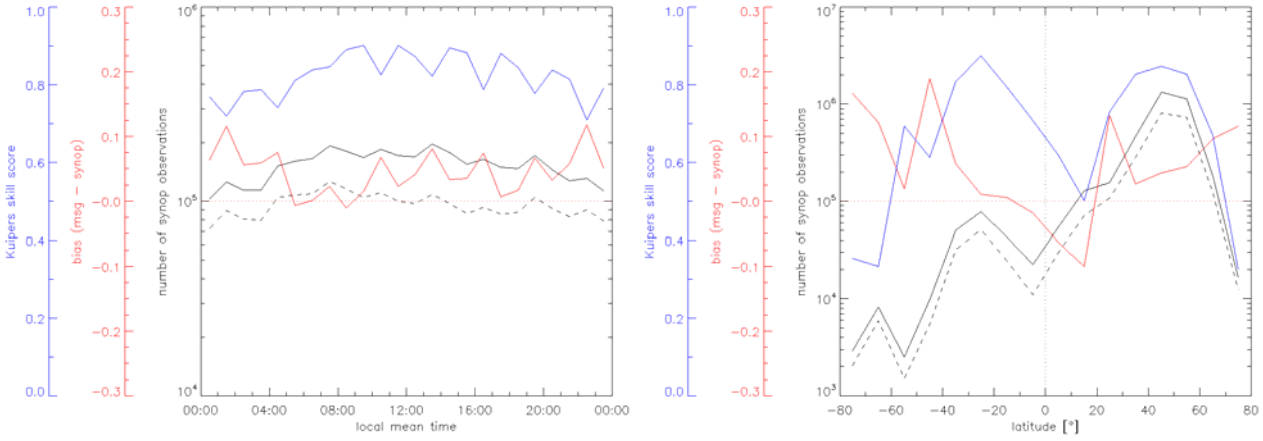

Figure 4.5 *Kss (blue), bias (red), number of synoptic reports (black) and number of synoptic reports (in octa) 0,1,7,8 (black, dashed) as function of local time (right panel) and latitude (left panel).*

Figure 4.5 shows the performance of satellite observations as function of the local mean time of observation (left panel) and latitude (right panel). The observed bias is positive outside the tropics but negative in the tropics (right panel). As expected the skill score is decreasing with increasing latitudes which is basically an effect of the increasing satellite zenith angle and the scenery effect. The SEVIRI pixel size increases with latitude and scenery effects may then cause an overestimation of cloudiness. There is a weaker dependency of both skill score and bias as function of the local mean time. Lower skill scores during night-time confirm the limitations of infrared-only satellite observations with regard to the detection of clouds. The black curve in Figure 4.5 (right panel) nicely shows the increasing number of observations (logarithmic scaling !) between 30°N and 50°N over Europe.

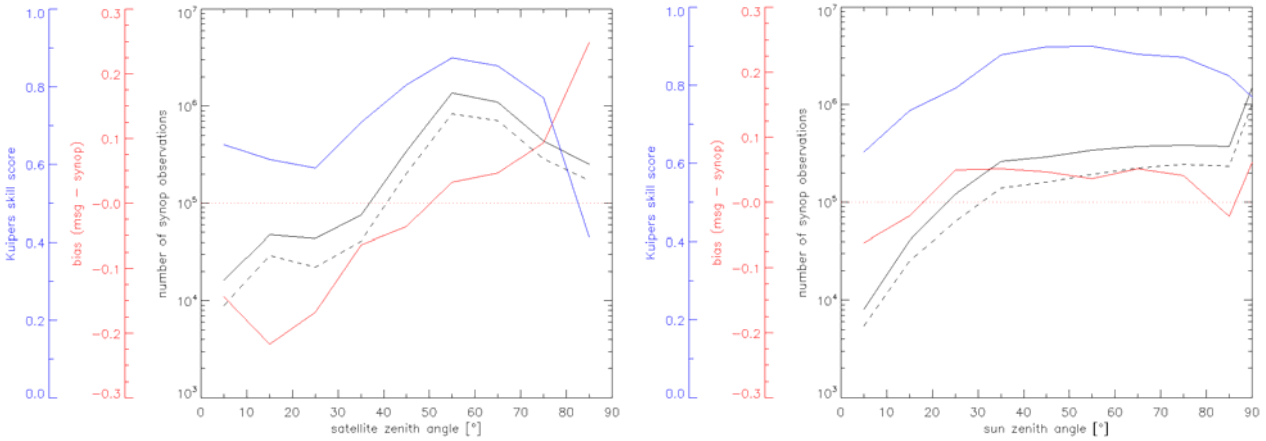

Figure 4.6 *Kss (blue), bias (red), number of synoptic reports (black) and number of synoptic reports (in octa) 0,1,7,8 (black, dashed) as function of the satellite (observation) zenith angle (left panel) and the solar zenith angle (right panel).*

The performance of satellite observations with respect to the satellite and sun zenith angles is depicted in Figure 4.6. The bias changes from negative to positive values if we go from low to high observation zenith angles (Figure 4.6, left panel), which translates into an underestimation of the satellite-derived cloud coverage close to nadir viewing conditions in the tropics and an overestimation of the cloud coverage towards the edges of the full disk. There is a lower dependency as function of the illumination conditions (right panel), except the known deviation over the tropics (for low solar zenith angles). The performance expressed by the skill score remains rather good, even for sunrise and sunset conditions.

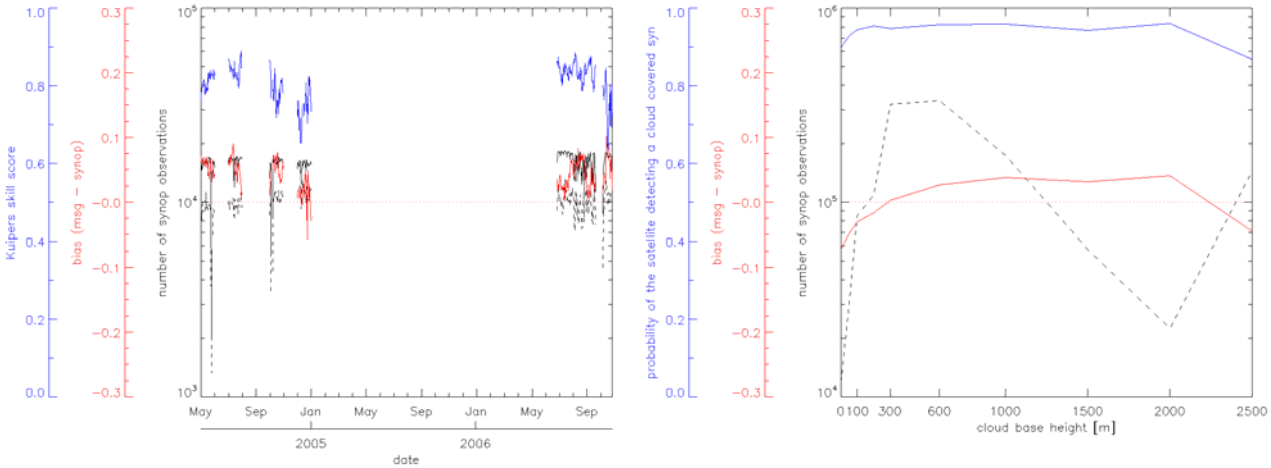

Figure 4.7 *Probability that the satellite observations correspond with the synoptic record "cloudy" (blue), bias (red), number of synoptic reports (black) and number of synoptic reports (in octa) 0,1,7,8 (black, dashed) as function of the day-of-year (left panel) and the cloud-base height (right panel). Note, that synoptic observations with cloud base height equal or higher than 2.5 km are summarised in the same class.*

Satellite and synoptic observations of the cloudiness may also differ as a function of the season (Figure 4.7, left panel) and the cloud-base height (Figure 4.7, right panel). There is a decreasing skill score from summer to winter conditions in the Northern hemisphere while the seasonal bias is closer to zero in December but remains constant otherwise. The bias shown on the right panel of Figure 4.7 is negative for both very low and higher cloud-base height which indicates an underestimation of satellite observations of very low clouds (low contrast between modelled clear-

sky brightness temperature and measured cloud-top brightness temperature). The negative bias towards the right edge (cloud-base height ≥ 2.5 km) is likely due to the fact that for off-equator conditions the problem of collocating high clouds and the surface position increases for higher clouds. For example, at 45 degrees viewing angle the projected dislocation of the cloud top is as large as the cloud altitude itself (by geometrical considerations) and this obviously introduces noise and uncertainties in our comparison. It shall however be noted that also ground-based measurements are known to miss thin cirrus clouds, especially during the night.

Results with bias correction

The same detailed analysis as presented above was performed with a bias correction applied, i.e. with a tuning factor of 0.75 (75%) that is the cloud coverage of partially cloudy SEVIRI pixels. Results are presented in Table 4.3 where notations are similar to those defined in Table 4.2.

As expected, the bias changes and is much lower than without the bias correction applied. We therefore conclude that if the satellite retrieval indicates a partially cloudy pixel the probability of having a (partially) cloud-covered pixel is higher than the “false alarm” of having a clear-sky pixel instead. Thus, the sensitivity of the satellite retrieval with respect to clouds is higher than with respect to clear-sky conditions. In that sense, we better understand the higher cloudiness over sea surfaces where cloud detection is generally easier due to the higher contrast, at least during daytime conditions if solar channels are present. It is however likely that also night-time temperature contrasts might be higher over sea (except for cases of temperature inversions and in windy conditions with a well mixed ocean surface layer) which is due to the smooth temporal change of the sea surface temperature relative to the air temperature.
The observed overall bias (0.008) is very low and much better than corresponding user requirements (10%) and this holds for “weighted” results as well (0.015). We omitted the presentation of detailed but rather similar graphics for the sake of clearness in this report.

Table 4.3 *Detailed results of the comparison of CFC results with a bias correction factor of 0.75 applied to SEVIRI partially cloud pixels and synop observations (count all = number of all synop reports, count cf/cc = number of only clear/cloudy synop reports, mean synop = mean CFC synop, sat hits cf syn = probability that cloud-free sat is confirmed by synop, sat hits cc syn = probability that cloudy sat is confirmed by synop, syn conf cf sat = probability that synop confirms cloud-free sat, syn conf cc sat = probability that synop confirms cloudy sat, hit = hitrate, Kss = kuiper skill score, bias = bias (mean sat – mean synop).*

scenario	count all	count cf/cc	mean syn	sat hits cf syn	sat hits cc syn	syn conf cf sat	syn conf cc sat	hit	Kss	bias
overall	3682740	2304745	0.535	0.878	0.942	0.920	0.910	0.914	0.820	0.008
weighted	1.E+06	574776	0.540	0.805	0.873	0.818	0.863	0.845	0.678	0.015
land	3047112	1932234	0.534	0.882	0.941	0.920	0.913	0.916	0.823	0.007
sea	150302	82670	0.610	0.748	0.961	0.884	0.904	0.900	0.709	0.067
coast	485326	289841	0.517	0.872	0.942	0.926	0.897	0.910	0.813	-0.005
day	2008448	1171056	0.555	0.915	0.955	0.931	0.945	0.939	0.871	-0.005
night	1489594	1018698	0.502	0.838	0.932	0.919	0.861	0.887	0.770	0.034
twilight	184698	114991	0.587	0.930	0.889	0.815	0.960	0.903	0.819	-0.053
mountain	39386	22937	0.375	0.835	0.912	0.958	0.695	0.858	0.747	0.056

4.3.3 Validation of daily and monthly means and the monthly mean diurnal cycle of CFC

We calculated daily means only for stations where at least 6 observations per day were available. Monthly means were then calculated from daily means for those stations where at least 20 days per month were available. This rule was also followed for the calculation of the monthly mean diurnal cycle. As for instantaneous values we applied a bias correction factor or tuning factor of 0.75 and 1 which is attributed to the cloud coverage of partially cloud SEVIRI pixels. In order to avoid a huge number of images we show only results where the tuning factor was 0.75. Note that we compare now average SEVIRI results of 15 × 15 km^2 boxes (the standard resolution of CM-SAF cloud products) with synoptic observations.

Daily mean CFC

Since we filtered ground-based stations with less regular temporal resolution of measurements the overall number of stations in use is lower than for the comparison of instantaneous CFC values. Daily averages were calculated both for ground-based and satellite-based observations and these were compared using the same methods as already presented in the previous section.

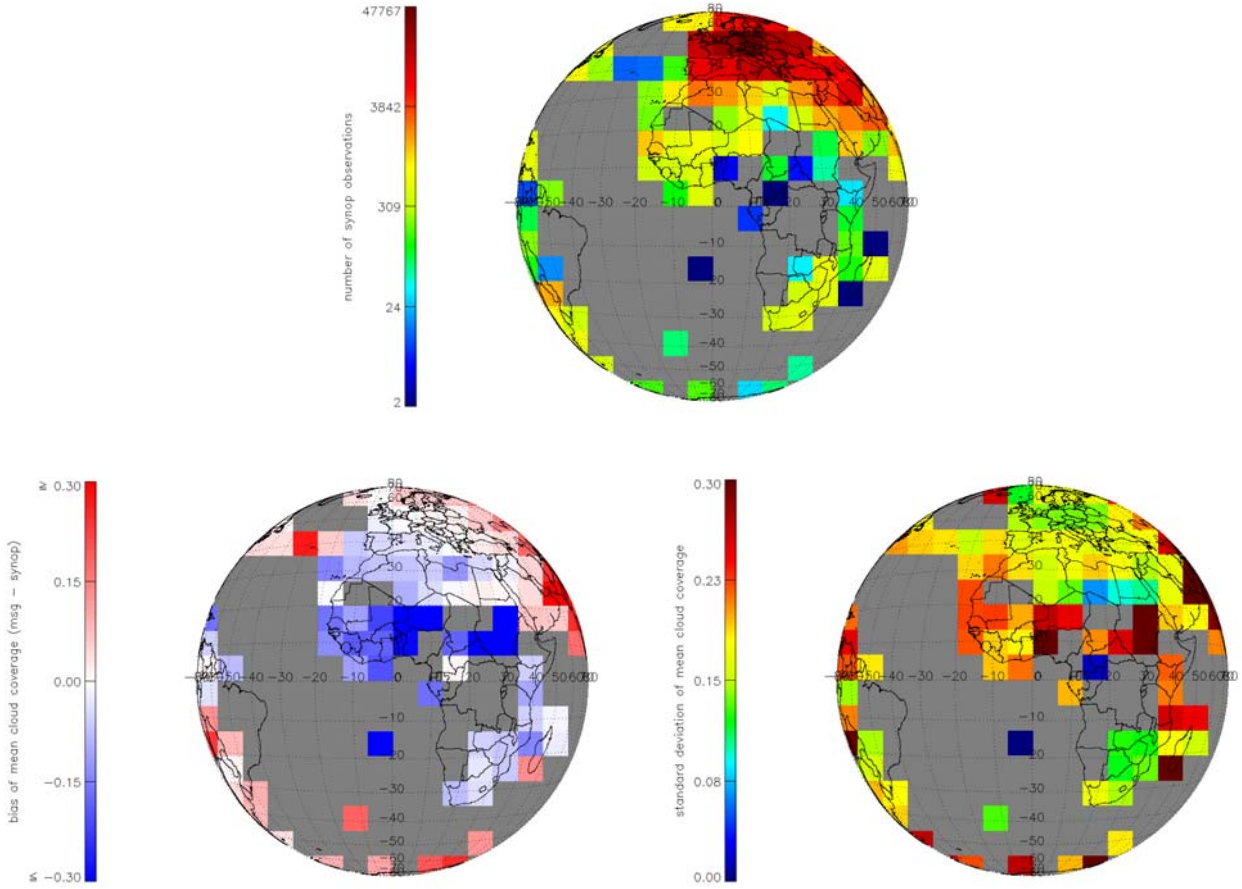

Figure 4.8 *Number of stations (upper panel) used and bias (lower left panel) and standard deviation (lower right panel) of the daily mean cloud fractional cover derived from SEVIRI and synoptic measurements.*

As for instantaneous results the spatial distribution of synoptic records is rather inhomogeneous (Figure 4.8, upper panel). The bias is negative in the tropics where the standard deviation is however high, due to the reduced number of reliable stations (Figure 4.8, lower panels). A systematic positive bias is again observed towards the edges of the visible disk. The daily mean results are closest to corresponding synoptic measurements over the Northern midlatitudes. Here, the bias is negligible and the standard deviation is in the order of one octa. The overall standard deviation

however is higher (~1.5 octa) while the spatial pattern is very similar (Table 4.4 and Figure 4.8). Note, that we omitted the category "weighted" because the spatial distribution of synoptic records is strongly irregular (almost no data over water) which may cause a corresponding irregularly weighted result by using very limited observations as representative values for all water surfaces.

Table 4.4 *Detailed results of the comparison of the daily mean CFC.*

scenario	count	avg syn	bias f=0.75	std f=0.75	bias f=1	std f=1
overall	266073	0.513	0.016	0.178	0.049	0.187
land	223923	0.514	0.018	0.176	0.051	0.185
sea	6717	0.614	0.068	0.198	0.104	0.208
coast	35433	0.488	-0.005	0.183	0.028	0.193
mountain	2637	0.316	0.044	0.192	0.071	0.206

The target accuracy (bias) of the daily mean CFC of 10% is easily reached by both the bias-corrected and uncorrected results.

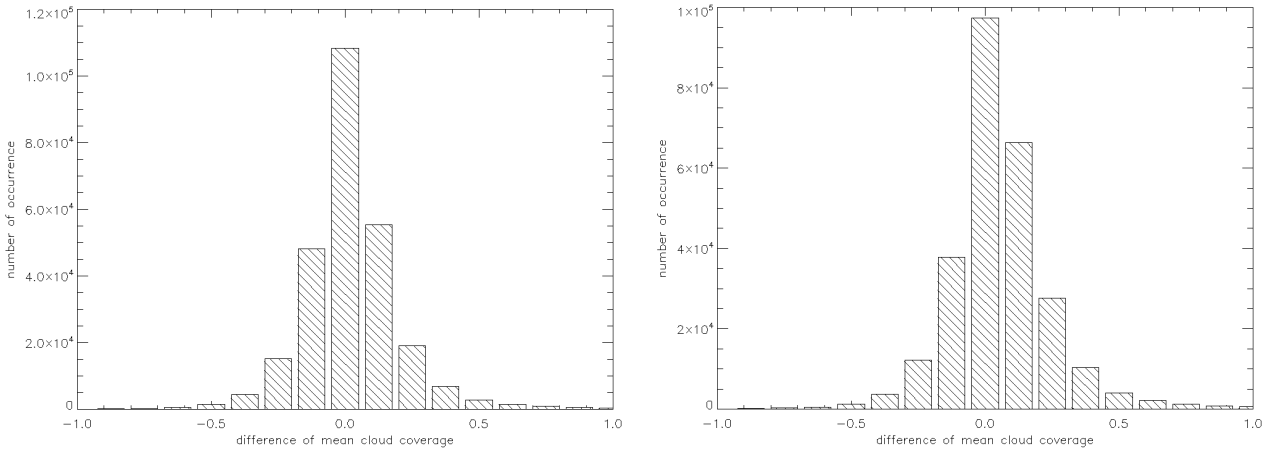

Figure 4.9 *Histogram of the difference of the daily mean cloud coverage with (left panel) and without bias correction (right panel). Differences were calculated as SEVIRI – SYNOP.*

The (positive) impact of the bias correction factor becomes obvious if we compare daily mean CFC results where such correction was applied. Figure 4.9 shows the histogram of the difference of the mean daily cloud coverage for the two cases. Clearly the distribution is balanced in the left panel of Figure 4.9 (almost no bias) while the right panel shows an overestimation of the SEVIRI CFC relative to the synoptic data.

Monthly Mean CFC

Monthly mean CFC values were generated for SEVIRI and ground-based records. We reduced again the amount of contributing synoptic stations in order to fulfil the selection criterion (20 daily means per month, 6 observations per day). Consequently, there are almost no measurements over water surfaces available anymore and results can be interpreted as representative for land stations only (Figure 4.10, top panel). We observe however very similar bias and standard deviation for the monthly mean CFC, with known deficiencies and features with respect to the spatial distribution (Figure 4.10, lower panels).

Table 4.5 *Detailed results of the comparison of the monthly mean CFC.*

situation	count	avg syn	bias f=0.75	std f=0.75	bias f=1	std f=1
overall	8564	0.517	0.016	0.114	0.049	0.117
Land	7231	0.520	0.017	0.112	0.050	0.115
Sea	201	0.608	0.067	0.121	0.105	0.128
Coast	1132	0.581	-0.001	0.121	0.031	0.125
mountain	58	0.302	0.024	0.069	0.049	0.077

While there is a negative bias in the tropics (for low and moderate satellite observation zenith angles) the opposite is true towards the edges of the full disk. The CFC performs best over Northern midlatitudes. The standard deviation is mostly below 1 octa (0.125) except for boxes at the edges of the visible disk and also along the Western African coast. Detailed results for different underlying ground and with and without bias correction are summarized in Table 4.5. As for instantaneous CFC results the target accuracy (bias) of 10% is reached for both bias-corrected and uncorrected results (0.016 and 0.049 at a mean CFC of 0.517, respectively) of the monthly mean CFC.

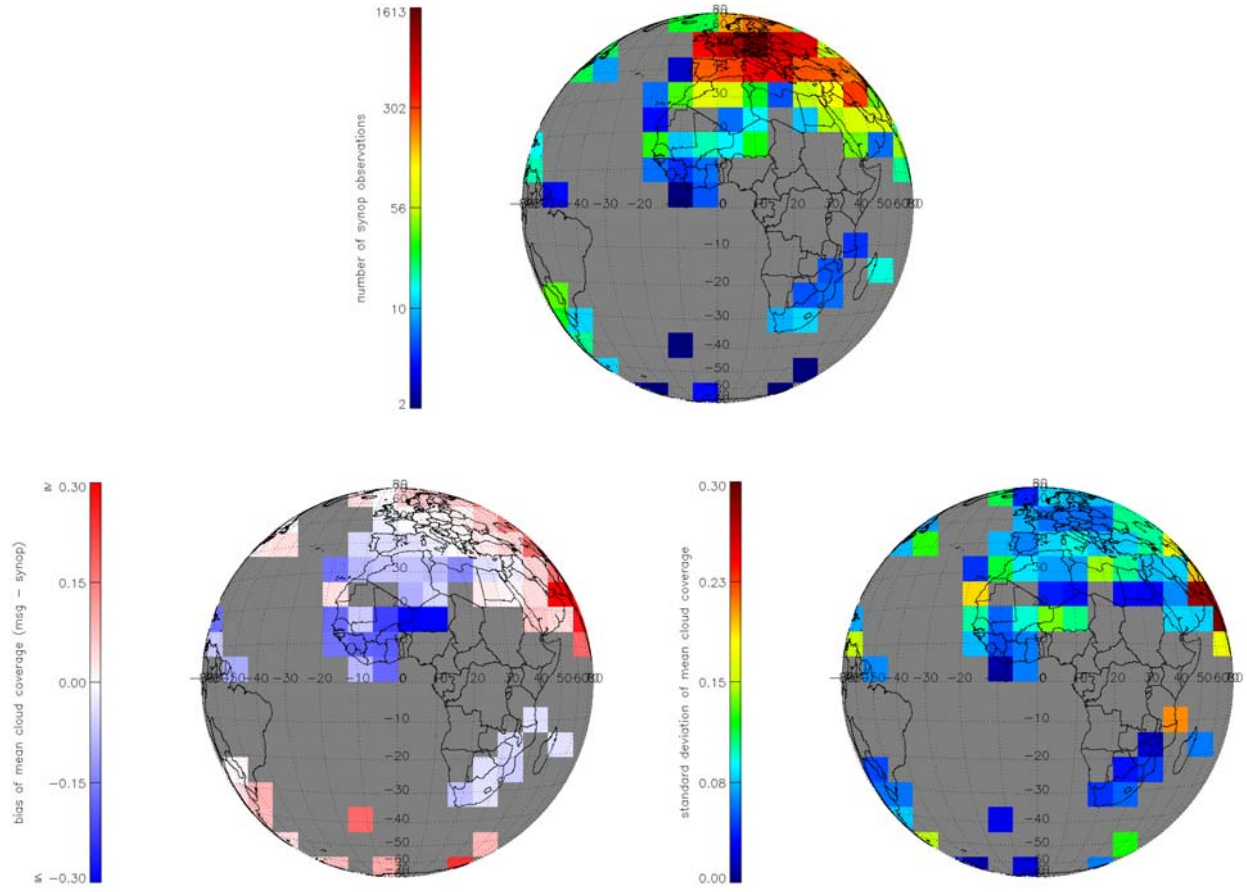

Figure 4.10 *Number of stations (upper panel) used and bias (lower left panel) and standard deviation lower (lower right panel) of the monthly mean cloud fractional cover derived from SEVIRI and synoptic measurements.*

Monthly Mean Diurnal Cycle of CF

Monthly mean diurnal cycles were computed for both ground-based and space-based records and the proven method for comparing the results was applied. A minimum of six synoptic measurements per day was required to generate the daily mean diurnal cycle and at least twenty days per month needed to be available for calculating the monthly mean diurnal cycle.

Table 4.6 *Detailed results of the comparison of the monthly mean diurnal cycle of CFC.*

scenario	Count	avg syn	bias f=0.75	std f=0.75	bias f=1	std f=1
overall	115823	0.537	0.011	0.143	0.046	0.147
land	96992	0.538	0.011	0.140	0.046	0.143
sea	3164	0.604	0.058	0.153	0.096	0.161
coast	15667	0.516	-0.001	0.158	0.032	0.165
mountain	838	0.349	0.055	0.167	0.085	0.176

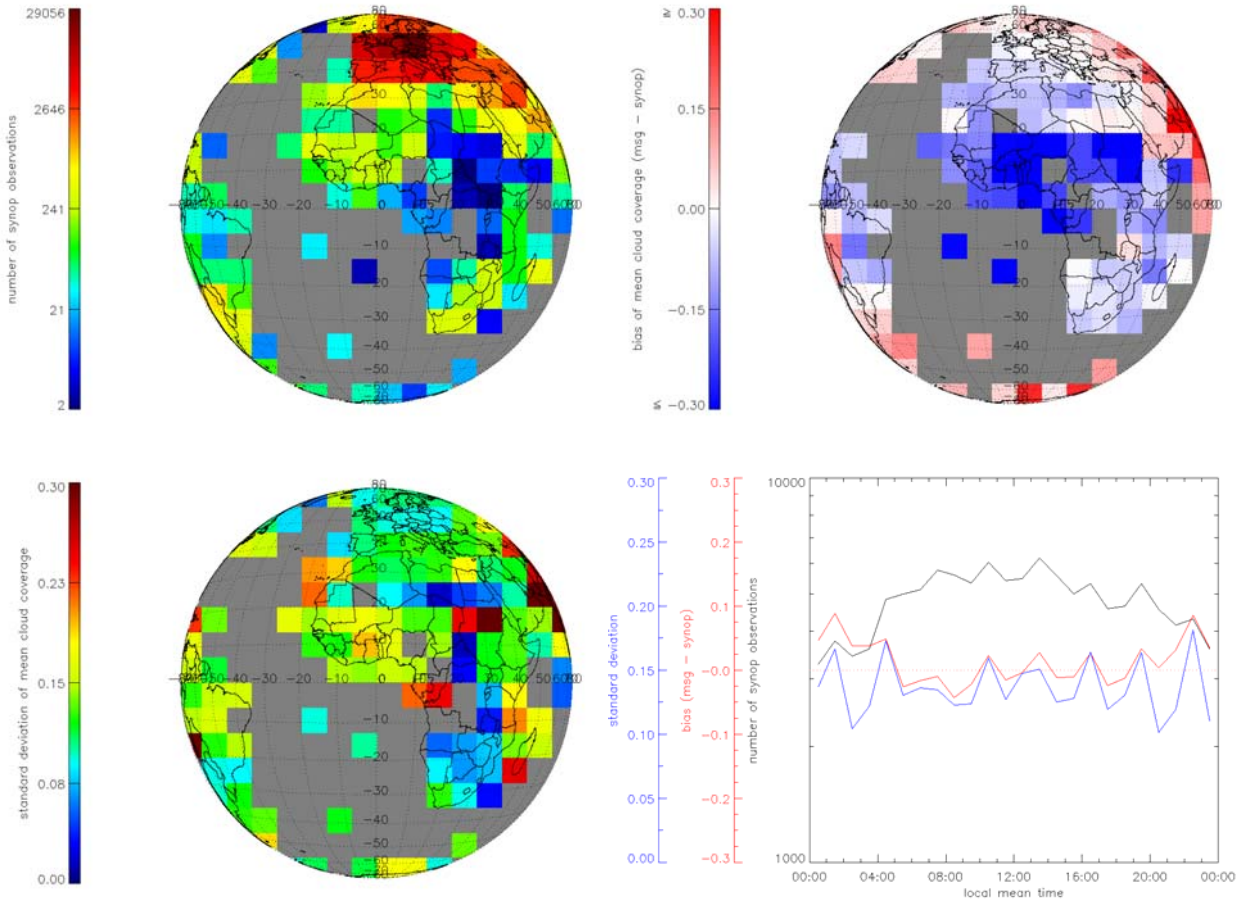

Figure 4.11 *Number of contributing stations (upper left) , bias (upper right), standard deviation (lower left) and bias and standard deviation of the monthly mean diurnal cycle of the cloud fractional cover as function of the local mean observation time (lower right). Results were derived from SEVIRI and synoptic measurements.*

As expected, results from the previous studies are basically reproduced. We observe a negative bias of CFC in the tropical belt and a positive bias becomes obvious towards the edges of the full disk (Figure 4.11, upper right). The overall bias is close to zero while the standard deviation is larger than one octa (0.125). Table 4.6 provides detailed numbers.

In Figure 4.11 (lower right) we provide the bias and the standard deviation of the CFC monthly mean diurnal cycle as function of the local mean time of observation. As can be seen from the figure the bias increases during the night but remains low during the day. The temporal pattern of the standard deviation is inconspicuous and follows mostly the pattern of the bias. We therefore conclude that the SEVIRI CFC picks up the daily course of the cloudiness although problems remain for night-time observations. It is however known that also night-time ground-based observations of the cloudiness are error-prone, especially if optically thin cirrus clouds are present. These clouds however can comparably easy be detected from space using the known split-window approach. It is therefore likely that observed differences are not only due to imperfect SEVIRI retrievals but also due to erroneous ground-based measurements.

4.4 Comparison against MODIS observations

We compared all SEVIRI and MODIS CFC results from 1st August 2006 on the full visible MSG disk. Data from twelve orbits from TERRA and AQUA were considered for this study. We used MODIS level 2 data (cloud fraction) with a reduced spatial resolution of about 5 km. As for synoptic data we calculated bias and skill scores.
The computation of skill scores included only MODIS pixels which were either almost cloud-free (≤ 1 octa) or almost fully cloud-covered (≥ 7 octa). There is typically a temporal delay between MODIS overpass times (every 100 minutes) and hourly SEVIRI observations and we try to compensate the possible movement of cloud structures by further sampling the MODIS data into 25 × 25 km^2 boxes applying a boxcar average filter. Since we restrict the comparison to almost homogeneous scenes (either cloudy or cloud-free) this running average is not fudging the results. We then compared this MODIS grid box result against instantaneous nearest-neighbour SEVIRI measurements which were not further sampled. Note that the calculation of bias values was done for all MODIS pixels (no restriction with respect to the cloud coverage).

As for the comparison against synoptic data we performed the analysis both with values 0.75 and 1 for the cloudiness of SEVIRI pixels which were classified as “fractional clouds”. Only bias-corrected results were visualized for the sake of clearness in the report.

Such satellite-satellite intercomparison may help to identify systematic errors and strengths and weaknesses of retrieval algorithms. It has also the advantage that observations are both done from satellites which avoids problems with multi-layered cloudy scenes when a comparison of ground-based and satellite-based cloud observations may not provide meaningful results.

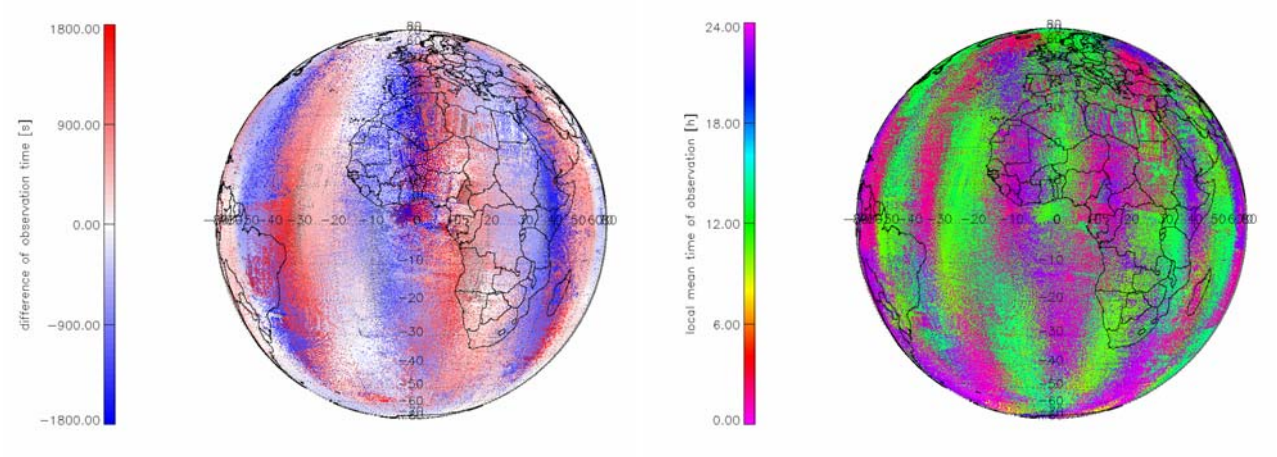

Figure 4.12 *Spatial distribution of the difference of observation times of SEVIRI and MODIS (left panel) and local mean time of observations (right panel) for 1st August 2006.*

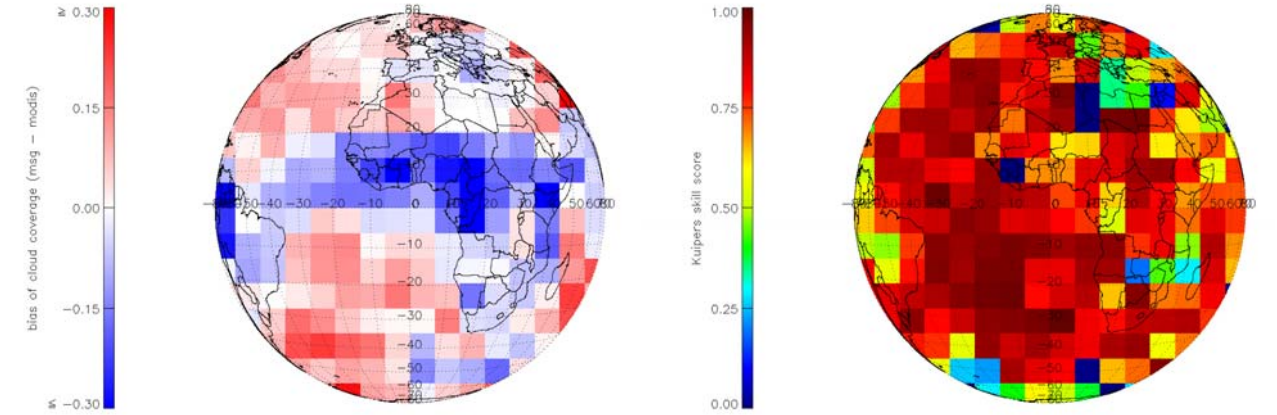

Figure 4.13 *Bias (left panel) and Kuiper skill score (right panel) for the differences of cloud fractional cover derived from MSG/SEVIRI and MODIS observations from 1st August 2006. Partially cloud-covered SEVIRI pixels contributed with fractional cloud coverage of 0.75.*

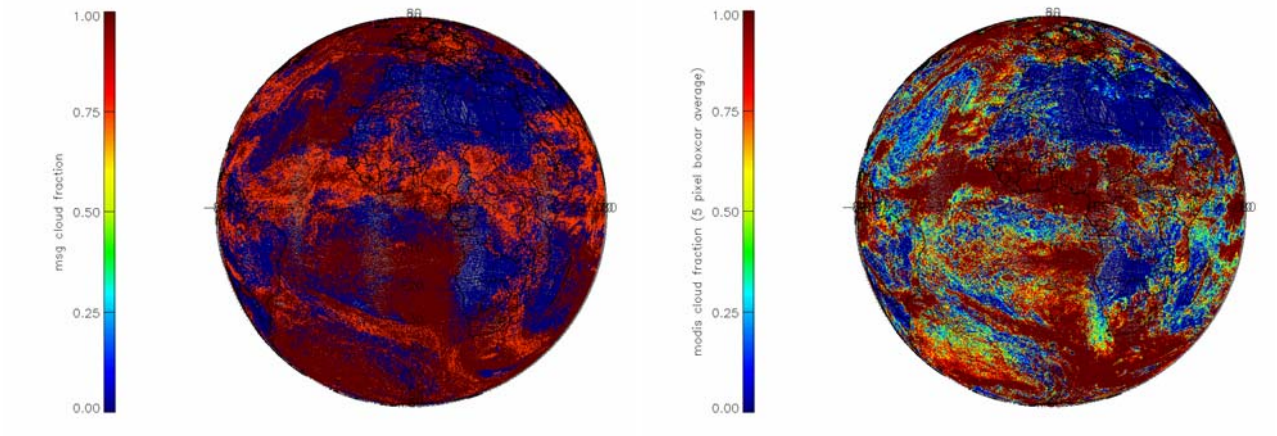

Figure 4.14 *Composite images of the cloud mask of SEVIRI (left panel) and MODIS (right panel) for 1st August 2006. Only those SEVIRI pixels closest in space and time to MODIS pixels are shown on the left panel.*

Figure 4.12 shows the spatial distribution of the difference of observation times of SEVIRI and MODIS (left panel) and the corresponding local mean time of observations (right panel). The spatial pattern of both the distributions is due to the sun-synchronous orbits of the TERRA and AQUA spacecrafts with inclinations of 98.2°. We allowed a maximum temporal shift of 30 minutes between the observations of SEVIRI and MODIS. Observations were made more frequently during midday and midnight hours than during morning (6:00) and evening (18:00) hours.

Clearly there is a negative bias of the cloud coverage in the tropical belt while SEVIRI results tend to be higher at larger observation zenith angle (Figure 4.13, left panel). There seems to be a slight trend to negative values over land surfaces in the midlatitudes while the bias is positive over water surfaces in the same latitudinal belts. The skill scores are generally high over water surfaces and a lower quality of results towards the edges of the full disk is obvious (Figure 4.13, right panel). Note that results become better over the tropical regions if partially cloudy pixels are assumed to be fully cloud-covered (no bias correction). However, the observed neutral or positive bias in other regions (North and South Atlantic, Europe) then becomes worse.

There is some evidence that the convectively formed cirrus cloud sheets (mainly caused by convective outflow/divergence at high regions over convective cells) in the tropical region are interpreted as being fractional (only cloud contaminated) by MSG while they are interpreted as opaque by MODIS. A possible explanation would be that the MSG algorithm has been mainly tuned for European conditions (for NWCSAF purposes). It seems as when the algorithm faces the extreme tropical conditions with a very moist atmosphere it fails in the proper distinction between fractional and semi-transparent clouds.

As for ground-based measurements we analysed the impact of the fractional cloud cover of partially cloud pixels. Figure 4.14 shows the cloud mask of SEVIRI and MODIS at pixel level and clearly there is an accumulation of SEVIRI cloud coverage values of 0.75 in the tropical belt but also partly over water surfaces in the midlatitudes (orange tones). These areas are known to be often covered by cirrus clouds which are detected by MODIS at a high resolution of 1 km. Such cirrus formation may remain undetected by SEVIRI at larger pixel scale. We therefore interpret the result as such that the tuning factor partly compensates for undetected optically thin cirrus clouds by SEVIRI. Detailed results applying the tuning parameter of 0.75 are summarized in Table 4.7 below. We believe however that a correction factor representing the cloudiness of SEVIRI fractional clouds would be a more complex function of the observation geometry rather than a fixed value for the entire area.

We have summarized the results (bias, skill scores) in the following Table 4.7:

Table 4.7 *Detailed results of the comparison of SEVIRI and MODIS CFC results with bias correction factor of 0.75 applied to SEVIRI partially cloud pixels (count bias = number of all data points used for bias calculation, count matrix = number of all data points used for kss calculation, mean MOD = mean MODIS cloud fractional cover, SEV hits cf MOD = probability that cloud-free SEVIRI is confirmed by MODIS, SEV hits cc MOD = probability that cloudy SEVIRI is confirmed by MODIS, MOD conf cf SEV = probability that MODIS confirms cloud-free SEVIRI, MOD conf cc SEV = probability that MODIS confirms cloudy SEVIRI, hit = hit-rate, Kss = kuiper skill score, bias = bias (mean SEVIRI – mean MODIS). Highlighted numbers emphasize good (green) and less good (red) results.*

scenario	count bias	count matrix	mean MOD	SEV hits cf MOD	SEV hits cc MOD	MOD conf cf SEV	MOD conf cc SEV	hit	Kss	bias
overall	6953455	4375768	0.587	0.917	0.933	0.886	0.952	0.927	0.850	0.0
weighted	1.E+06	631452	0.570	0.906	0.931	0.893	0.940	0.921	0.837	0.004
land	2427339	1681420	0.477	0.932	0.846	0.876	0.912	0.892	0.778	-0.08
sea	4485744	2671300	0.647	0.899	0.967	0.903	0.966	0.950	0.866	0.043
coast	40372	23048	0.534	0.838	0.931	0.904	0.881	0.890	0.769	-0.04
day	3516589	2138270	0.543	0.919	0.959	0.943	0.942	0.942	0.878	0.032
night	3397361	2208843	0.629	0.917	0.911	0.821	0.961	0.913	0.828	-0.034
twilight	39505	28655	0.790	0.536	0.965	0.655	0.945	0.918	0.501	0.086
mountain	58919	37227	0.491	0.852	0.910	0.908	0.855	0.881	0.763	0.020

Although the bias is almost zero the skill score is only moderately high which is a hint that MODIS and SEVIRI CFC results differ frequently. A relatively high negative bias (and a low skill score) is found over land (SEVIRI underestimates the cloudiness relative to MODIS) while the opposite is true for twilight conditions which however occur seldom. Here, it is the low probability that SEVIRI and MODIS results agree for cloud-free conditions (0.536) that is responsible for the low skill sores.

Without providing all the details it shall be mentioned that the overall bias is in the order of 0.04 if partially cloudy SEVIRI pixels are counted as fully cloudy pixels (i.e. the tuning factor is 1). We therefore conclude that application of the tuning factor is also advantageous if SEVIRI CFC results are compared with other satellite-borne cloud coverage data of better spatial resolution.

We further analysed the bias and the skill scores also as functions of the latitude and the satellite observation zenith angle of SEVIRI (Figure 4.15).

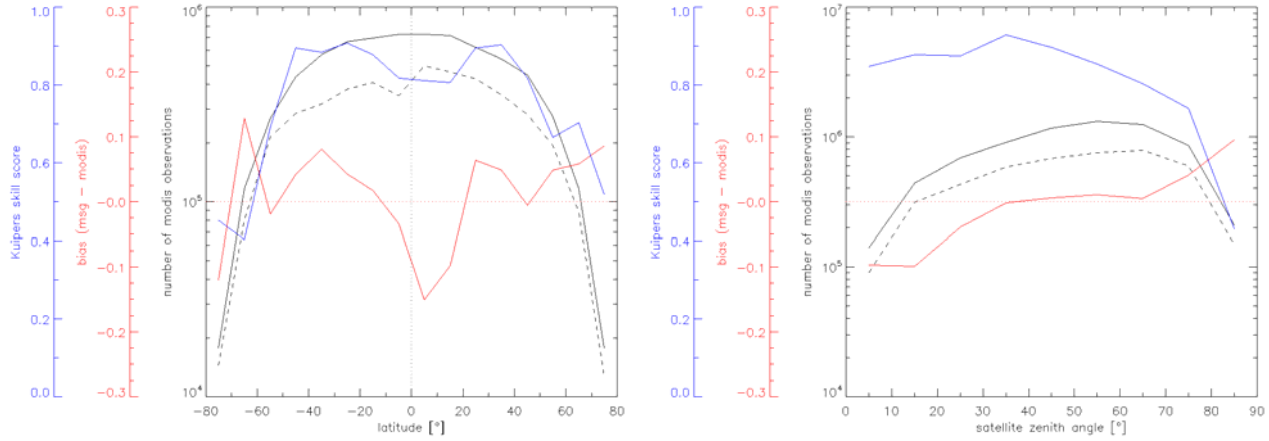

Figure 4.15 *Kss (blue), bias (red) and the number of collocation pixels (black) as function of latitude (left panel) and satellite observation zenith angle (right panel). Partially cloud-covered SEVIRI pixels contributed with fractional cloud coverage of 0.75.*

Kss values are highest in both the subtropical regions between 20 and 40 degrees latitude but decrease towards the equator and higher latitudes. The bias however is highest in the tropical belt but performs best at midlatitudes. The number of collocation observations decreases with increasing latitude. On the other hand, the kss rapidly decreases for satellite zenith angles larger than 30°. The lower performance is also seen in the increasing bias which however becomes obvious not until the satellite zenith angle is larger than 65 degrees.

Further analysis of the kss and the bias as a function of the sun zenith angle and the local mean time of observation is shown in Figure 4.16.

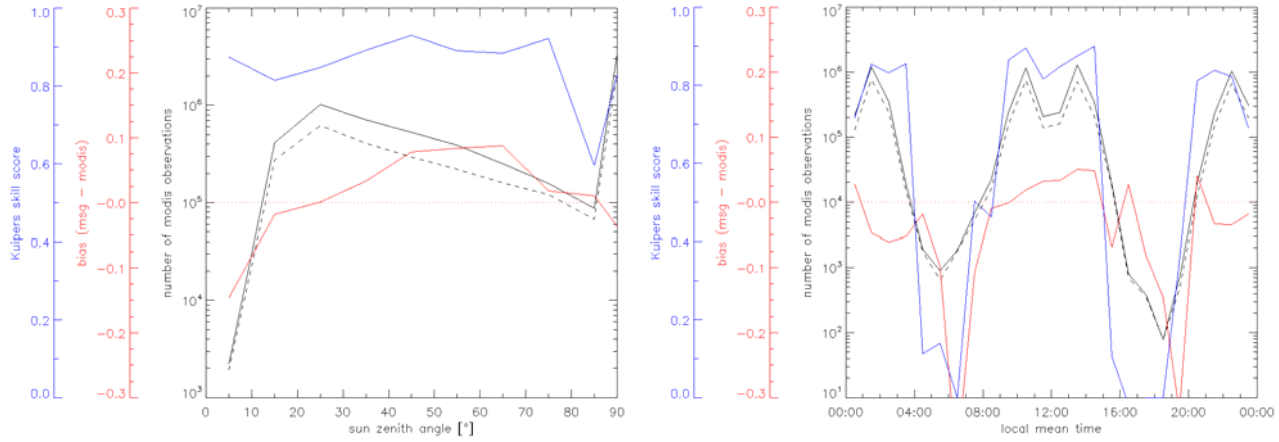

Figure 4.16 *Kss (blue), bias (red) and the number of collocation pixels (black) as function of sun zenith angle (left panel) and the local mean time of observations (right panel). Partially cloud-covered SEVIRI pixels contributed with fractional cloud coverage of 0.75.*

The skill score remains high for a large range of solar zenith angles of up to 70 degree. The bias however is high (negative) for low solar zenith angles below 15 degrees and again for solar zenith angles between 25 and 65 degrees (positive). This is basically in line with the results shown in Figure 4.13.
The temporal pattern of kss, bias and the number of observations is influenced by the orbit geometry of TERRA and AQUA on 1st August 2006 (Figure 4.16, right panel). Strong negative biases go in line with low skill scores but also a small number of collocation observations during

local morning and evening hours (i.e. close to twilight conditions). From Figure 4.12 (right panel) we see however that these observations were made around Antarctica at high observation zenith angles of SEVIRI where Figure 4.15 (left panel) showed already low Kss values and a negative bias for this area.

4.5 Comparison against CALIPSO observations

We compared all SEVIRI and CALIPSO (or better CALIOP) CFC results from 1st August 2006 on the full visible MSG disk. As for the comparison against synoptic data we performed the analysis both with values 0.75 and 1 for the cloudiness of pixels which were classified as “fractional clouds” by the SEVIRI retrieval. Only bias-corrected results were visualized for the sake of clearness in this report.

CALIOP footprints are small stripes of ≈ 100 m. The nominal horizontal resolution is about 330 m. These footprints are overlaid by the swath of the IR imager which is 64 × 64 km^2. The pixel size of IR data is about 1 km. To avoid ambiguous results we used the information from the IR imager at 10.6 µm to select homogeneous scenes for the comparison with SEVIRI results. Therefore we defined a stripe of 21 pixels (10 at each side of the LIDAR track plus a centre pixel matching the track) around and calculated the standard deviation of the radiances. A low standard deviation is then a signal for homogeneous scenes. We defined a threshold value of 50 W m^{-2} sr^{-1} μm^{-1} (for the standard deviation) to remove stripes with large scatter from the collocation study. The threshold value is a compromise between the number of remaining pixels and observed inhomogeneities. The CALIOP level 2 product provides also information about the number of atmospheric layers that can be distinguished with respect to their physical properties. Only those pixels were considered where just one layer was found which avoids difficulties with multi-layered cloud scenes.

It shall be emphasized that SEVIRI and CALIOP basically measure different quantities: There is the SEVIRI emitted and reflected radiance from the earth-atmosphere system while CALIOP measurements are based on the analysis of backscattered (polarized and unpolarized) radiation from an active remote sensing system in the VIS and NIR spectral bands. Thus, CALIOP measures physical quantities of cloud particles (water droplets, ice crystals) and aerosol particles. The CFC product is then derived from this basic analysis after spatial sampling of the results of single LIDAR measurements. Although CALIOP operates in the VIS and NIR spectral bands, it is possible to measure under daylight conditions but with slightly reduced performance. The backscattered radiation can still be separated into the LIDAR part and the underlying atmospheric noise which is due to photons registered by CALIOP that originate from scattering processes of incoming solar radiation.

CALIOP footprints (enlarged in figure) and local overpass times or the local mean time of observations are shown in figure 4.17. It becomes obvious that only midday and midnight observations can be compared with SEVIRI which is a consequence of the suns-synchronous polar orbit of CALIPSO.

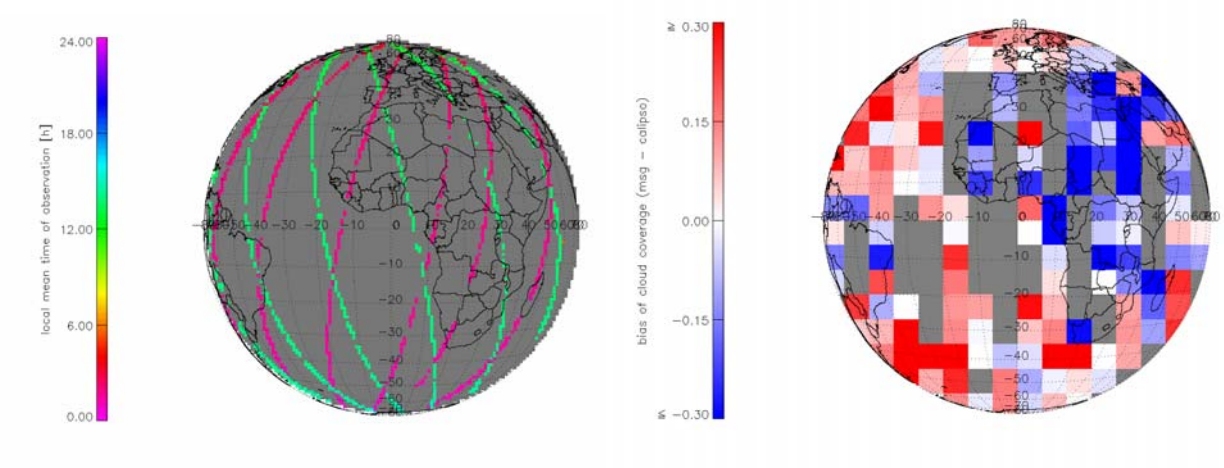

Figure 4.17 *CALIPSO local mean time of observation for orbits of 1st August 2006 (left panel) and geographic distribution of the bias of the cloud fractional cover (right panel).*

As for the comparison with synoptic data and MODIS results we calculated the bias and the Kuiper skill score of CFC. SEVIRI fractional clouds contributed with a cloudiness of 0.75 to the SEVIRI CFC result. The geographic distribution of the bias between SEVIRI and CALIOP CFC values is shown in Figure 4.17. We provide the bias also as function of the observation zenith angle and the cloud-top height (see also section 6.5).

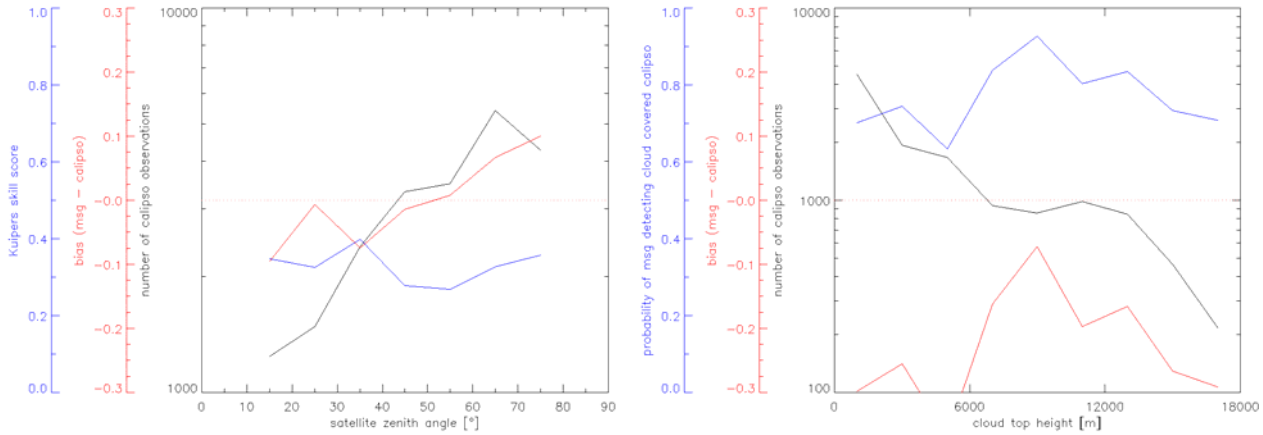

Figure 4.18 *Kuiper skill score (blue), bias (red) and the number of observations (black) as function of the satellite zenith angle (left panel) and the cloud-top height (right panel).*

We found a strong contrast in the bias over land and over sea surfaces. While the bias is mostly negative over land surfaces (SEVIRI underestimates the cloud fractional cover), we see the opposite effect over water surfaces. This result is in line with previous results shown for MODIS and synoptic data but more pronounced. We need however to take into account that CALIOP observations of a single day are not homogeneously distributed over land and sea surfaces (see Figure 4.17). The bias increases also towards the edges of the visible MSG disk which is again in line with MODIS and synoptic results (Figure 4.18, left panel). The bias increases for very low and very high clouds which coincides with the lower probability that SEVIRI measurements confirm a cloud-covered CALIOP pixel (Figure 4.18, right panel). Here, the best performance is found for mid-level and high clouds between 6000 m and 10000 m height a.s.l.

Detailed results are summarized in Table 4.8. Classes “mountain”, “coast” and “twilight” occurred rarely and were not further analysed.

Table 4.8 *Detailed results of the comparison of SEVIRI and CALIOP CFC results with bias correction factor of 0.75 applied to SEVIRI partially cloud pixels (count bias = number of all data points used for bias calculation, count matrix = number of all data points used for kss calculation, mean CAL = mean CALIOP cloud fractional cover, SEV hits cf CAL = probability that cloud-free SEVIRI is confirmed by CALIOP, SEV hits cc CAL = probability that cloudy SEVIRI is confirmed by CALIOP, CALIOP conf cf SEV = probability that CALIOP confirms cloud-free SEVIRI, CAL conf cc SEV = probability that CALIOP confirms cloudy SEVIRI, hit = hitrate, Kss = kuiper skill score, bias = bias (mean SEVIRI – mean CALIOP). Highlighted numbers emphasize good (green) and less good (red) results.*

scenario	count bias	count matrix	mean CAL	SEV hits cf CAL	SEV hits cc VAL	CAL conf cf SEV	CAL conf cc SEV	hit	Kss	bias
overall	21626	21626	0.574	0.604	0.743	0.636	0.717	0.684	0.347	0.021
weighted	1.E+06	1.E+06	0.560	0.551	0.784	0.667	0.690	0.682	0.335	0.077
land	7102	7102	0.545	0.787	0.551	0.594	0.756	0.658	0.338	-0.148
sea	14405	14405	0.588	0.503	0.832	0.677	0.705	0.696	0.335	0.106
day	10089	10089	0.565	0.550	0.830	0.714	0.705	0.708	0.380	0.100
night	11326	11326	0.578	0.654	0.664	0.586	0.725	0.660	0.318	-0.049

The skill scores are generally low which is a consequence of low probabilities that cloud-free SEVIRI pixels are confirmed by corresponding CALIOP observations (column SEV hits cf CAL in Table 4.8) and again that CALIOP cloud-free pixels coincide with corresponding SEVIRI observations (column CAL conf cf SEV in Table 4.8). Obviously, there are many SEVIRI footprints that were identified as being cloud-free where CALIOP still detects clouds at a lower spatial scale. CALIOP and SEVIRI observations agree generally better for fully cloudy pixels. There are also clearly different results for daytime and night-time observations: While SEVIRI overestimates the cloudiness during the day the opposite is true for night-time (SEVIRI IR-only) observations.

4.6 Comparing CM-SAF polar and geostationary CFC products

As a final illustration of how MSG/SEVIRI CFC results depend on the MSG viewing angle we will here show results based on mutual comparisons of results from the CM-SAF polar satellite product based on NOAA AVHRR data (processed with the NWCSAF PPS software) with corresponding results (matched in time) from the currently validated MSG cloud scheme.

Results from April 2006 have been chosen covering the CM-SAF Baseline area (NOAA AVHRR results are not produced outside of this region). Figure 4.19 visualises the difference between polar and geostationary CFC results over the central portion of the Baseline area (i.e., where we have good coverage of NOAA AVHRR data received by the HRPT station in Offenbach). The differences are formed using individual CFC values with a resolution of 15 km. Results are based on comparisons of more than 300 NOAA AVHRR overpasses.

We clearly see that differences between SEVIRI and AVHRR retrievals increase with latitude (increasing MSG viewing angle). The difference is largest for the NOAA-17 satellite. The smaller deviations and less pronounced correlation with MSG viewing angle for the NOAA-15 satellite is probably explained by less favourable observation conditions for both AVHRR and SEVIRI (predominantly twilight conditions) compared to the case of NOAA-17 (mid-day and mid-night condi-

tions). Results were derived using a bias-correction for cloud-contaminated pixels using a correction factor of 0.75 (75%) for both PPS and MSG retrievals. Interesting to note here is that results do not change drastically if using the factor 1.0 (100 %) instead (not shown here) but a noticeable reduction of the difference at high latitudes is seen. This appear to be related to having a higher frequency of cloud-contaminated pixels in the MSG SEVIRI cloud mask at high latitudes than for NOAA AVHRR.

Some interesting changes of results are seen when changing PPS versions from version 1.0 (currently used in the CM-SAF operational chain) to the new version 1.1 (to be implemented later the CM-SAF). The large negative difference over the northern part of Africa is largely removed which verifies an improved capacity of cloud detection over desert areas for PPS version 1.1. Furthermore, the MSG-PPS difference increases at higher latitudes as an effect of e.g. revised PPS treatment of cloud-ice discrimination and the labelling of thin cirrus clouds.

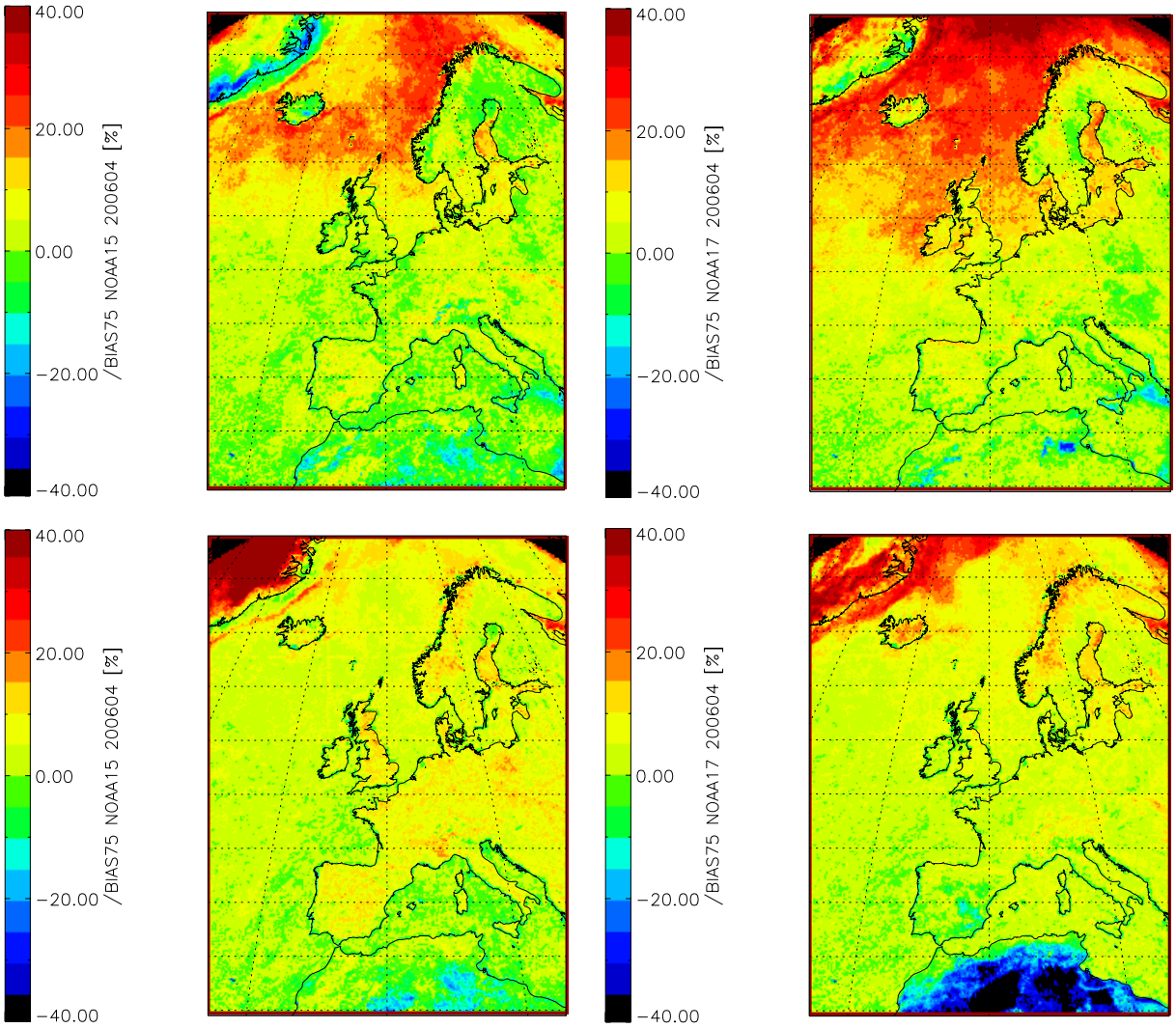

Figure 4.19: *The CFC bias (MSG CFC minus AVHRR CFC) in April 2006 using different versions of the NOAA AVHRR cloud software PPS. Top panels: PPS version 1.1 NOAA15 (left) and NOAA17 (right). Bottom panels: Corresponding results for PPS version 1.0. A bias correction factor for cloud contamined pixels was set to 0.75 (same for both MSG and NOAA) weighted factor is set to 75%.*

4.7 Summary

We performed an analysis of the cloud coverage retrieved from both synoptic records and satellite observations using instantaneous data from SEVIRI of the full visible earth disk. Furthermore, we analysed the quality of average quantities such as the daily and monthly mean values and the monthly mean diurnal cycle of the cloud fractional cover. In general, the results of this study and the initial validation study (see AD 9) which was performed for one summer month only agree quite well.

Taking into account all the differences between ground-based measurements and satellite observations (viewing geometry, subjective man observations, rules for ground-based observations) the overall performance of instantaneous CFC results is rather good which is confirmed by a high Kuiper skill score of around 0.82 and a bias of only 0.044 or 0.008 if tuned for partially cloudy sat-

ellite footprints. Daily and monthly mean values and the monthly mean diurnal cycle were further calculated for both ground-based and space-based records and we could show that the given target accuracy (bias) of CFC products is reached for all product types.

The detailed results however clearly show that satellite measurements overestimate the cloud coverage over sea surfaces (entries “sat hits cf syn” ≈ 0.75, column 5 in Table 4.2 and Table 4.3). Lowest Kuiper skill score and the lowest probability that the synoptic record “cloud-free” is confirmed by the satellite retrieval underline this result. We found also a negative bias of CFC in the tropical belt which is compensated by a positive bias towards the edges of the earth disk. Best performance of all products is found over Northern Midlatitudes, mainly over European land surfaces. Furthermore, we could show that night-time infrared-only measurements from satellites are of lower quality than corresponding daytime measurements which include the information from SEVIRI solar channels (skill scores lower, bias higher).

The unequally distributed ground observations with the bulk of measurements over land surfaces in mainly Europe pretend a better result as it would be for a spatially uniform distribution of synoptic observations. In fact these latter results draw a more realistic picture of the overall performance of the cloud coverage results derived from SEVIRI. The skill score is then 0.67 only (Table 4.2) and the high “false alarm” rate of detecting clouds from space where synoptic measurements indicate clear-sky conditions becomes evident.

We could answer the open question about the validity of the bias correction factor outside the baseline area. Our analysis confirmed this assumption since the bias of the full disk results is minimised if the tuning parameter is set to 75% (0.75). Since the impact of fractional clouds changes with the observation angle (scenery effect) one could also think about a more sophisticated correction factor that is a function of the observation angle. A more detailed study may confirm this assumption.

Comparison of SEVIRI CFC with MODIS CFC gave similar results: The SEVIRI cloud coverage is lower at low latitudes (low solar zenith angle, low observation angle) but increases towards the edges of the full disk and overestimates the cloudiness there. The overall bias is almost zero but some systematic differences became obvious over land surfaces (-0.08) and under twilight conditions (0.086). We applied the bias correction factor also for the comparison of SEVIRI and MODIS results and our analysis confirms the results of the comparison against ground-based measurements. SEVIRI CFC tends to overestimate the cloudiness outside the tropics which is reduced by applying a cloudiness of 0.75 for partially cloudy pixels.

The MODIS-SEVIRI intercomparison provides some evidence that SEVIRI often misses optically thin cirrus clouds, which may be due to the low spatial resolution and a corresponding comparably low impact of such clouds on the SEVIRI radiances. The tuning factor for the cloudiness of partially cloudy pixels of 0.75 seems to be useful to partly compensate this weakness of SEVIRI. It became however obvious that such tuning factor should then be a function of the observation zenith angle or the latitude. A more sophisticated study using more data would be required to determine such quantity which may be also a function of the season (moving ITCZ).

Comparison of SEVIRI CFC with CALIOP CFC basically confirms the known deficiencies of the CFC product: There is a negative bias over land which is even more pronounced than for MODIS while SEVIRI overestimates the cloudiness over water surfaces. We see a similar result for the daytime/night-time comparison: The bias is higher during the day but becomes negative for night-time observations. We observe also an increasing bias towards the edges of the visible MSG disk.

We reached the CFC target accuracy of 10% under almost all conditions. The only exception is the limited comparison of SEVIRI and CALIOP CFC over land where we found a negative bias of 0.148 over land surfaces. More data would need to be analyzed to confirm or discard this preliminary result. Our study also indicates (i.e., the quite low Kuipers skill scores) that more work has to be done on the utilisation of the new validation data sources from CloudSat and CALIPSO before we can feel confidence in how to use them in an optimal way.

5 Validation of cloud type - CTY

This section is divided into the following paragraphs:

- Product description
- Validation task
- Validation against ground-based measurements
- Summary

5.1 Product description

We basically reuse the validation method developed by SMHI that is described in the previous CM-SAF validation report of the ORR-V2 (AD 7) and the SIVVRR-V3 (AD 9).

CM-SAF provides daily and monthly averaged cloud type frequencies derived from MSG/SEVIRI data. The more detailed cloud type assignments of the NWC-SAF standard product are grouped in five categories (low level clouds, mid level clouds, high opaque clouds, high semi-transparent clouds and fractional clouds), and for each of this categories we first retrieve instantaneous results per slot which can be further averaged to daily and monthly cloud type frequencies.

The validation of cloud type categorizations using ordinary synoptic observations is more difficult than for the CFC product. The reason is the different observation geometry from ground and from space which hampers a simple comparison. For example, in cases of multi-layered cloudiness the surface may report exclusively low-level clouds while the satellite reports exclusively mid-level or high-level clouds. In such case it is impossible to evaluate the skill of the observation/interpretation since both the surface observation and the satellite interpretation fail in giving a correct description of the true cloudiness.

The validation study comprises data from May, July, October and December 2004 and compares cloud radar data from the Cabauw and Chilbolton measurement sites and radiosonde data with corresponding condensed cloud type categories derived from the MSG/SEVIRI product. Since other reliable cloud radar data outside Europe weren't available we concentrate on results over Europe showing mainly the performance of the new NWCSAF MSGv1.2 software version.

Cloud radar measurements are in contrast to synoptic observations seen as a reliable source to describe the occurrence and the exact vertical location of cloud layers, although being limited if optically thin and high clouds are present. We have simulated satellite viewing conditions (i.e., viewing from space) by assigning cloud types to the uppermost cloud layer detected by the cloud radar using vertical temperature and pressure information from radiosonde measurements.

5.2 Validation task

Cloud-top height data retrieved from radar profiles at Cabauw were compared to radiosonde measurements at De Bilt (The Netherlands; 52.10° N, 5.18° E) and the corresponding cloud pressure and cloud temperature data were retrieved. In July and May the Cabauw radar data set was available nearly continuously from 7:30 am to 5 pm. However, in December and October major data gaps occurred. In total, eighteen days with acceptable radar coverage were available in October and 21 days in December.

The cloud radar data from the Cabauw ground-station were pre-processed using a software routine developed by KNMI. For each 30 minutes time window a mean cloud radar cloud top height (CTH) value was provided which is temporally centred on the MSG observation time (slot time).

To account for non-homogeneous cloud layers and multi-layered clouds a minimum and maximum CTH value was computed.

CTY assignments (low-level clouds, mid-level clouds, high-level clouds) are provided for the corresponding mean, minimum and maximum CTH values. In contrast to the CTY definition driven by surface observations, cloud-top altitudes rather than cloud base altitudes were used to simulate satellite viewing conditions. Radar retrieved CTH values were converted to CTY assignments using corresponding radiosonde temperature [K] and pressure measurements [hPa] as follows:

- Low level cloud: $T > 0.8 \times T_{850hPa} + 0.2 \times T_{700hPa} - 8$
- Mid level cloud: $0.8 \times T_{850hPa} + 0.2 \times T_{700hPa} - 8 > T > 0.5 \times T_{500hPa} - 0.2 \times T_{700hPa} + 178$
- High level cloud: $T < 0.5 \times T_{500hPa} - 0.2 \times T_{700hPa} + 178$

The radar classes are then in line with the CTY class definition used in the MSG algorithm.

Daily radar-derived cloud type frequencies were computed by calculating the percentage of low, mid and high level clouds from 30 minutes cloud type assignments at the available MSG observation times for each measurement day. Single cloud radar samples with high temporal resolution (intervals of a few seconds) were not available and this limits the accuracy of the radar-retrieved CTY frequencies. In case of broken cloud fields or in the presence of non-continuous multi-layer cloud layers, one CTY assignment based on a 30 minutes averaged cloud-top height does not necessarily reflect the actual cloud situation and will presumably cause erroneous CTY frequencies. In order to study the effect of CTY misclassifications caused by the time window of 30 minutes, we have performed a direct MSG CTY/radar CTY comparison for two different data sets:

The first data set includes all MSG and radar CTY assignments of an investigated month, while the second data set only included cases with similar CTY assignments for maximum, minimum and mean CTY in a 30 min time window.

This method allows determining how often an averaged 30-minutes CTH value erroneously results in a mid-level cloud type when in reality only low level and high level clouds are present simultaneously. It therefore helps to better understand how results are affected under multi-layer cloud conditions. A further limitation of the computed radar CTY frequencies is the fact that the radar data was not filtered for rain clouds. Rain droplets can significantly reduce the radar signal and hence lead to incorrect CTH measurements.

MSG CTY frequencies were calculated for corresponding cloud radar observations. In a 3 x 3 pixel window (approximately covering a geographical region of 9 x 15 km) centered on the cloud radar station, MSG CTY values were retrieved and re-classified into the categories ‘low-level clouds’, ‘mid-level clouds’ and ‘high-level’ clouds. In addition, quality flags are derived for the investigated pixel boxes based on the quality flag information associated with the basic cloud type algorithm. Similar to previous validation studies (AD 7, AD 9) the radar/MSG inter-comparison was performed for a pixel box instead of a single MSG pixel to account for geometric dislocations of high cloud layers (the so-called parallax effect). A detailed study has shown that the definition of a pixel window of 3 x 3 pixels does not affect the validation results significantly.
The cloud radar does not allow distinguishing between semi-transparent and opaque cloud types. Consequently, these two groups were merged into one high-level cloud type category. Furthermore, the MSG category “fractional clouds” could not be validated since a corresponding cloud class cannot be retrieved from radar data.

5.3 Validation against ground-based measurements

We first compared daily averages of the cloud type frequencies derived from SEVIRI and radar observations at Cabauw site in the following scatter diagram (Figure 5.1) for the three cloud type classes. Note, that we compare only daytime measurements which limits also the SEVIRI daily averages to daytime measurements.

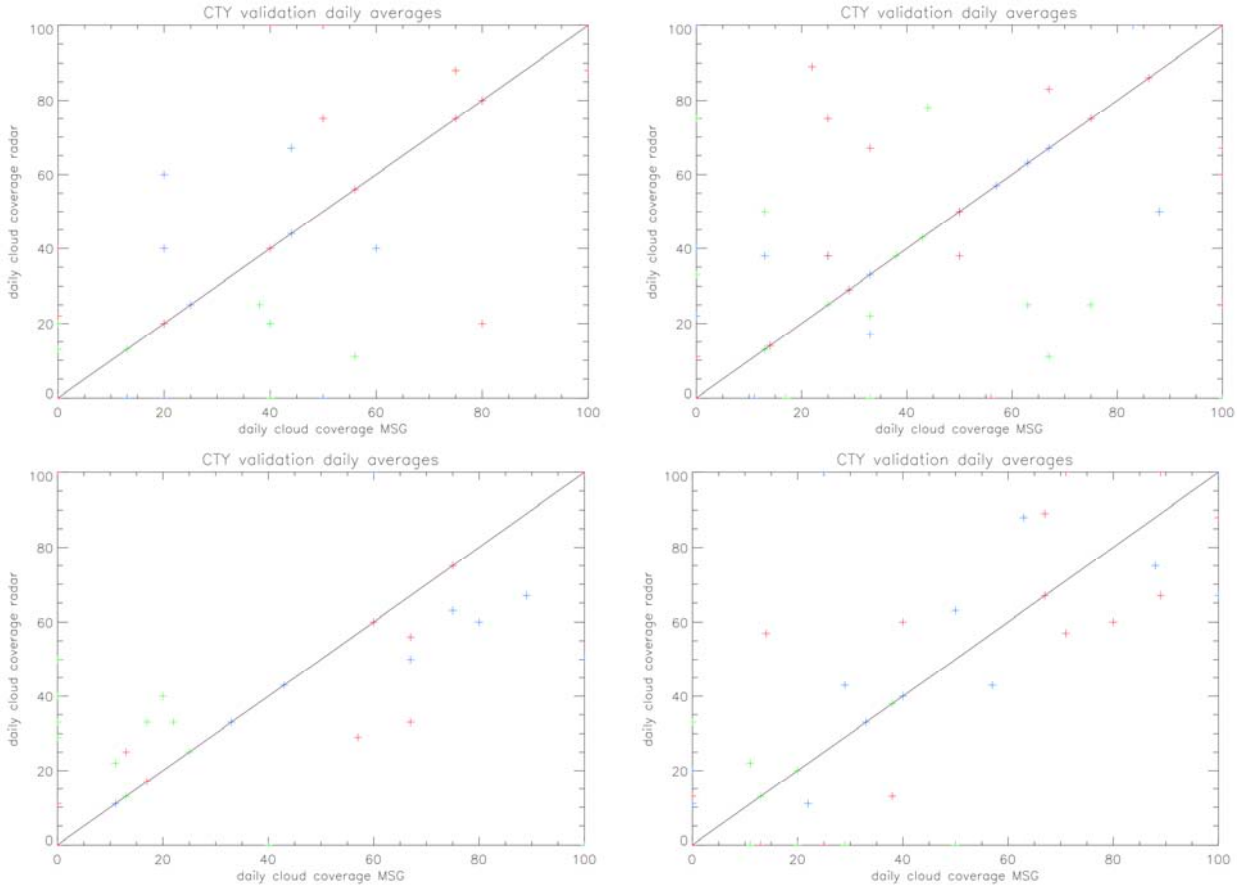

Figure 5.1 *Scatter diagram of the cloud coverage (in %) derived from SEVIRI and cloud radar measurements (blue = low-level, green = mid-level, red = high-level clouds) for May (upper left), July (upper right), October (lower left) and December (lower right) 2004.*

As expected best performance is seen for low-level clouds (blue marks) while the scatter increases for high-level clouds (red marks) and mid-level clouds (green marks). For the latter two classes there seems to be a slight trend towards a higher cloudiness observed by the cloud radar.

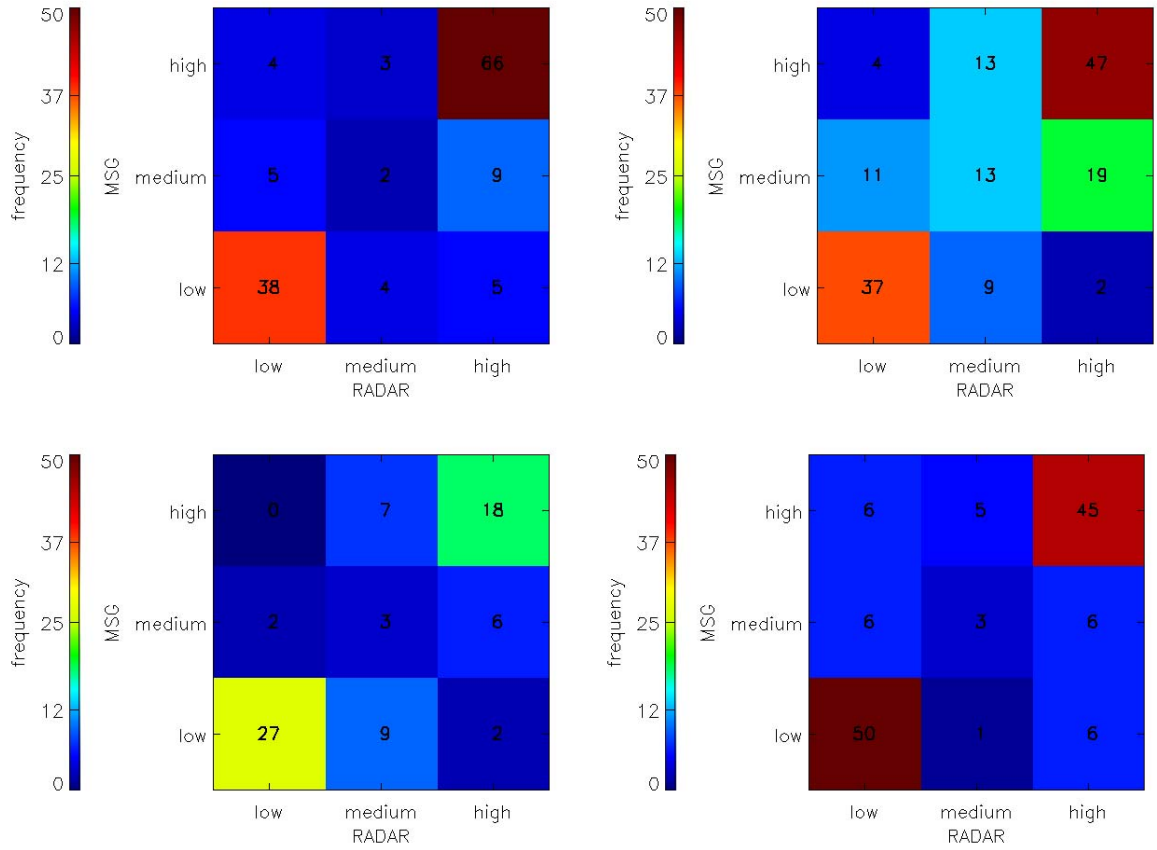

Figure 5.2 *Scatter diagram of daily cloud type occurrences derived from SEVIRI and cloud radar observations for the Cabauw measurement site in May (upper left), July (upper right), October (lower left) and December (lower right) 2004. Numbers in boxes denote the frequency of occurrences.*

A more condensed scatter diagram is given in Figure 5.2 where we show the occurrence of the three cloud type classes for both the observation sources. The best performance is found for low clouds and even better for high clouds where most coincident results were found. SEVIRI measurements tend to underestimate the height of high-level clouds (e.g. green box in Figure 5.2, upper right) but one should take into account that the assignment of cloud type classes/height regimes is strongly influenced by the temperature profile. Here, we rely on radiosonde data (radar) of a station close by Cabauw and basically NWP model data (SEVIRI) which are of different accuracy and are representative for a much larger area. Furthermore, the (uncorrected) brightness temperature retrieved from SEVIRI is typically not the same as the level temperature taken from NWP data since the emissivity of especially optically thin clouds is not equal to 1. Nonetheless, we could show that other combinations (radar-high – SEVIRI-low and vice versa) occur rarely which basically confirms the validity of our validation approach.

Statistics of this comparison are summarized in Table 5.1. The overall agreement of cloud categories varies between 77.9% (May 2004) and 62.6 % (July 2004), with high coincidence for low-level and high-level clouds.

Table 5.1 *Percentage of matches between cloud radar and MSG derived CTY's for May, July, October and December 2004. We used the MSG product quality flag for results of the right column.*

	all data		single layer[1]		good quality and single layer	
	N	*percentage matches*	*N*	*percentage matches*	*N*	*percentage matches*
May 2004						
All CTY's	136	**77.9**	82	**84.2**	69	**85.5**
…. low level clouds	*… 47*	*80.9*	*22*	*81.8*	*21*	*81*
… mid level clouds	*… 9*	*22.2*	*0*	*0*	*0*	*0*
… high level clouds	*… 80*	*82.5*	*60*	*85*	*48*	*87.5*
July 2004						
All CTY's	155	**62.6**	56	**76.8**	48	**77.1**
… low level clouds	*… 52*	*71.2*	*18*	*83.3*	*18*	*83.3*
… mid level clouds	*… 35*	*37.1*	*3*	*33.3*	*2*	*50*
… high level clouds	*… 68*	*69.1*	*35*	*77.1*	*28*	*75*
Oct 2004						
All CTY's	74	**64.9**	41	**75.6**	36	**72.2**
… low level clouds	*… 29*	*93.1*	*18*	*88.9*	*15*	*86.7*
… mid level clouds	*… 19*	*15.8*	*4*	*0*	*4*	*0*
… high level clouds	*… 26*	*… 69.2*	*19*	*79*	*17*	*76.5*
Dec 2004						
All CTY's	128	**76.6**	77	**85.7**	60	**85**
… low level clouds	*… 62*	*80.7*	*34*	*94.1*	*25*	*92*
… mid level clouds	*… 9*	*… 33*	*-*	*-*	*-*	*-*
… high level clouds	*… 57*	*… 79*	*42*	*81*	*34*	*82.4*

As stated in (AD 7), it seems that many of the radar-retrieved mid level cloud type assignments are actually caused by temporally averaged CTH values from multiple cloud layers or broken clouds. The usage of the MSG quality flag in combination with single layer data sets does not significantly affect the results.

The correlation coefficient is highest for low-level clouds and the scatter of results is lowest for this class. The lowest bias is found for high-level clouds (Table 5.2) while results are generally poor for mid-level clouds.

Table 5.2 *Linear correlation coefficient (r^2), bias and rms error for daily MSG CTY cloud frequencies compared with daily cloud radar derived frequencies.*

Month	low-level			mid-level			high-level		
	r^2	bias	rms	r^2	bias	rms	r^2	bias	rms
May 2004	0.92	1.4	16.4	0.61	4.6	15.0	0.84	-6.0	23.5
July 2004	0.82	-5.4	23.1	0.41	4.5	31.5	0.76	0.9	26.9
October 2004	0.62	14.7	33.7	-0.33	-4.7	35.1	0.52	-10.0	38.6
December 2004	0.87	-4.1	20.4	0.20	4.4	17.2	0.89	-0.3	17.6

We finally compared monthly frequencies of cloud type derived from radar and satellite observations (Table 5.3).

Table 5.3 *Comparison of MSG- and cloud radar retrieved monthly cloud type frequencies. Note, that only daytime SEVIRI slots were considered.*

	low-level [%]		mid-level [%]		high-level [%]	
	radar	msg	radar	msg	radar	msg
May 2004	41.6	43.0	6.2	10.7	52.2	46.3
July 2004	39.8	34.4	20.7	25.2	39.5	40.4
October 2004	31.7	46.4	21.2	16.5	47.1	37.1
December 2004	56.7	52.6	6.6	11.0	36.7	36.4

Monthly CTY frequencies agree within ± 20%, with best performance for high-level clouds which are almost equally distributed, although results of October 2004 are slightly degraded. Despite the limitations of the applied validation method (no rain filter, daily CTY frequencies are computed from 30 min averaged radar CTH values etc.) monthly MSG and radar retrieved CTY frequencies show good correlations and we believe that this should give good confidence in the quality of the CM-SAF MSG derived CTY frequencies. We conclude that we confirm the results shown in (AD 7) and (AD 9).

Results of the correlation (Table 5.2) between satellite and radar observations are comparable to results of the previous validation study (AD 7) while the monthly cloud type frequencies changed (Table 5.3). Here, we limited the comparison to daytime SEVIRI measurements which influenced the results such that relatively more low-level clouds and relatively less mid-level clouds were found in both data sets (being also consistent with results of the intercomparison of MSGv1.1 and MSGv1.2).

5.4 Summary

The MSG CTY product has been validated using detailed information about cloud layer occurrence and cloud layer altitudes derived from cloud radar measurements made in May, July, October and December 2004. The initial cloud height information from cloud radar was converted into corresponding cloud-top temperature and cloud pressure values which allowed the definition of cloud-type classes comparable to the MSG/SEVIRI CTY product. In general, the results confirm previous SIVVRR V2 and ORR-V2 validation results (AD 7, AD 9). Especially low-level and high-level cloud assignments appear to work well but problems are evident for mid-level clouds. However, this problem is largely explained by limitations of the validation method rather than by wrong MSG CTY assignments. The cloud type product is subject of a major revision during the upcoming CDOP, mainly because of the fact that we are currently not utilizing the results provided by the complete set of cloud products (e.g. corrected cloud top information and the set of cloud physical products).

6 Validation of cloud top parameters – CTH/CTT/CTP

This section is divided into the following paragraphs:

- Product description
- Validation task
- Validation against ground-based measurements
- Comparison against MODIS observations
- Comparison against CALIOP observations
- Summary

6.1 Product description

The height of the cloud-top in the atmosphere is an important parameter that influences the energy budget of the atmosphere. Shortwave radiation is transmitted through (semitransparent) clouds while the emitted thermal radiation is again reflected back to the earth-atmosphere system. The efficiency of this heating effect depends strongly on the height of the reflecting layer, i.e. the height of a cloud layer in the atmosphere. On the other hand, optically thick (water) clouds reflect the incoming radiation at the cloud-top height which can be low in the atmosphere (stratus-like clouds) but also very high (cumulonimbus-like convective clouds) and again it is this height level that is important for the energy budget of the earth-atmosphere system.

For the comparison with ground-based measurements we focus on the cloud-top height product. Other cloud-top parameters can be seen as just being similar products but partly controlled by dedicated input data (actual NWP data, climatological profile data). Instantaneous data and daily mean products from SEVIRI were compared against corresponding cloud radar data sets.

Furthermore, cloud-top pressure results were compared against other satellite-based (MODIS) results for one day on the full visible MSG disk. In addition, we performed a comparison of cloud-top height results derived from SEVIRI and CALIOP observations of the same day. Such comparison may help in the first line to identify algorithm weaknesses but can also be used to understand sensor-dependent advantages and drawbacks, for example due to the different observation geometry of geostationary and polar-orbiting satellites or measurement principles (active versus passive instruments, different spectral bands). This further helps to interpret e.g. time series of such quantities that were derived from different sensors.

Hourly cloud-top estimations of MSG/SEVIRI were compared with corresponding radar-derived CTH values. The results were validated against CTH retrievals derived from cloud radar measurements of the two observation sites Chilbolton (UK) and Cabauw (The Netherlands) in Europe. Other radar data outside Europe was not available which limits the significance of these results for the full MSG disk.

Also ground-based LIDAR data can be used to retrieve the cloud-top height, as it is presented in Trolez et al. (2005) for cloud products derived from AVHRR measurements. The basic software modules and the method of comparing cloud radar data and satellite observations have been already developed by SMHI for ORR-V2 (AD 7) and we re-applied the method successfully for the initial validation of SEVIRI cloud products (AD 9).

6.2 Validation task

The cloud radar data was collected at Cabauw (51.97 North 4.93 East) and Chilbolton (51.14 North 1.44 West) during the CloudNet campaign (see http://www.met.reading.ac.uk/radar/ cloud-net/index.html). At the time of writing we had access to Cloudnet data sets of 2004 which limits our validation results to this period.

During this campaign cloud radars at Cabauw and Chilbolton were typically operated from 7:30 am to 5:00 pm but data gaps still occurred frequently. Cloudy and clear sky conditions can actually unambiguously detected by the cloud radar through the appearance of corresponding sharp signals (cloud/no cloud). However, very high and optically thin cloud layers may occasionally not be detected by the cloud radar and hence CTH obtained by the cloud radar does not always reflect the height of the uppermost cloud layer. In order to compare the radar time series with the nearly instantaneous MSG/SEVIRI measurements, the radar CTH retrievals were averaged over 30 minutes centred on a quarter to and a quarter past the hour (MSG/SEVIRI observations are always made a quarter to the hour). Additional information about the CTH variation in the time window is provided by means of minimum and maximum CTH, which helps for correct interpretation since cloud layers do not necessarily persist for 30 minutes.

We compared daily mean results and instantaneous results of the cloud-top height. MSG/SEVIRI CTH pixels were selected in a 3 × 3 pixel box (approximately covering a geographical region of 9 × 15 km) centred at the position of the radar station to account for broken cloud fields and MSG navigation ambiguities. Maximum, minimum and average CTH values were generated using the CTH product flag including results from semi-transparent or opaque cloud categories.

In addition, we compared SEVIRI cloud-top pressure results with corresponding MODIS results from 1st August 2006. As for the cloud fractional cover we performed a collocation study for SEVIRI and MODIS “pixels” where the latter are available as 5 × 5 box average values. Here, we compared results over the full visible MSG disk.

Another collocation study was performed for SEVIRI and CALIOP cloud-top height results but based on single SEVIRI pixels. The temporal and spatial sampling of CALIOP data is much higher and we selected pixels closest in space and time for this comparison. Again, we could compare SEVIRI and CALIOP results over the full disk.

6.3 Validation against ground-based measurements

As already shown in the previous validation report (AD 7) there is a large scatter of cloud-top height results derived from the cloud radar and MSG/SEVIRI observations. In general, results from the Cabauw station agree better with MSG/SEVIRI results than corresponding results from the Chilbolton radar (see Figure 6.1 and Figure 6.2). This holds for instantaneous data pairs as well as the daily mean values.

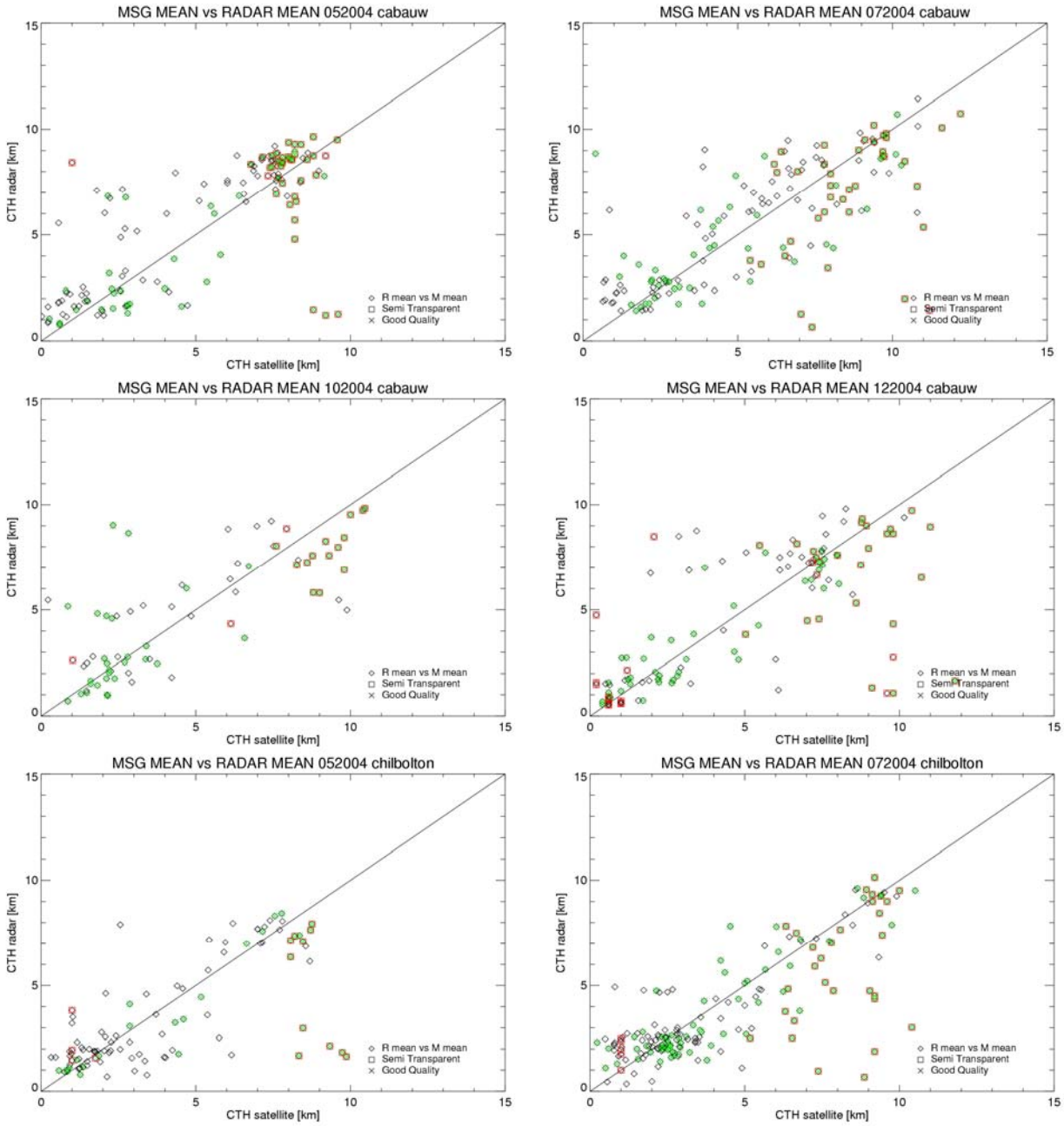

Figure 6.1 *Scatter diagram of CTH retrievals from SEVIRI (x-axis) and cloud radar (y-axis) using data of Cabauw of May (upper left), July (middle right), October (middle left), December (middle right) 2004 and of Chilbolton of May (lower left) and July (lower right) 2004. Green symbols represent results with good SEVIRI product quality while red symbols denote cases where semitransparent clouds were detected by SEVIRI.*

There is a general tendency of higher cloud-top values derived from satellite data which is more pronounced for Chilbolton. One reason could be the presence of semitransparent clouds which obviously hampers the comparison of cloud radar data and satellite observations. Outliers below the 1-1 line (red symbols below the 1-1 line in Figure 6.1, upper right panel) are likely to be due to multi-layer cloud scenarios where the cloud radar easily detects cloud layers close to the ground but misses the higher, optically thin semitransparent clouds. If we concentrate on results deemed

"good" by the satellite retrieval the overall agreement of results is not improving in a statistically significant manner.

Daily (Figure 6.2) and monthly mean averages of the cloud-top height were calculated from available measurements. Note, that we computed appropriate daily and monthly data only from those SEVIRI observations that were considered for the validation of instantaneous products.

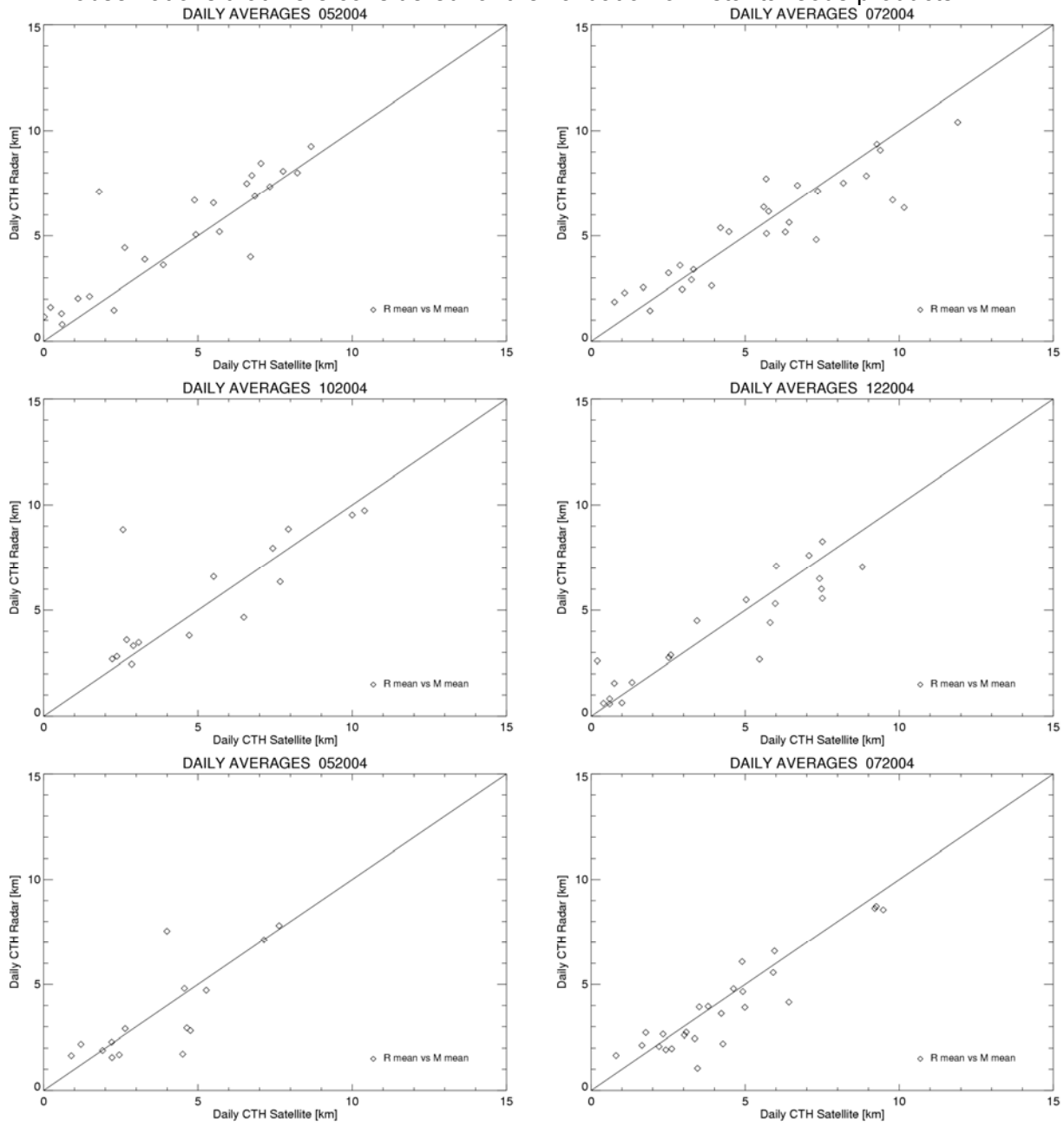

Figure 6.2 *Scatter diagram of daily averages of the cloud-top height derived from SEVIRI (x-axis) and radar observations (y-axis) of Cabauw of May (upper left), July (upper right), October (middle left), December (middle right) 2004 and of Chilbolton of May (lower left) and July (lower right) 2004.*

Statistical quantities are summarized in the following Table 6.1. The target accuracy of CTH (10%) is mostly reached for both data sets (except for Cabauw in May 2004). Results for Chilbolton are only available for the summer period.

Table 6.1 *Summary of statistical measures of the comparison of the cloud-top height derived from satellite and radar observations for the two selected stations and for four months. Chilbolton observations were only available for May and July 2004.*

Site	Monthly Mean MSG CTH [m]	Monthly Mean RADAR CTH [m]	Daily Mean RMS [m]	Daily Bias [m]	Daily Abs Mean Error [m]	Total Obs.	Quality [%]	Non Opaque [%]	Multi-ple Layers [%]	Correla-tion
May 2004										
Chilbolton	**3736**	**3575**	1414	161	965	102	36	83	10	0.78
Cabauw	**4368**	**5023**	1503	-655	1029	138	46	68	31	0.88
July 2004										
Chilbolton	**4328**	**3975**	983	353	767	176	49	77	19	0.92
Cabauw	**5622**	**5351**	1342	271	1030	154	51	72	30	0.90
October 2004										
Cabauw	**5257**	**5647**	1823	-391	1134	72	57	76	12	0.79
December 2004										
Cabauw	**4166**	**4027**	1200	139	937	127	48	57	0	0.91

In general, the observed bias between SEVIRI and ground-based observations is low and is typically below 10 %.

6.4 Comparison against MODIS observations

We compared SEVIRI and MODIS cloud-top pressure results on the full visible MSG disk for observations of 1st August 2006. In contrast to SEVIRI retrievals the MODIS algorithm retrieves the cloud-top pressure directly by a modified CO_2 slicing approach (Menzel et al., 1983), without the need of analyzing actual NWP profiles. A possible bias of NWP data may therefore translate into a corresponding bias of SEVIRI results. It shall however be noted that also the CO_2 slicing approach needs at least climatological profile information (temperature, humidity) as first guess. The MODIS temperature and height products are further derived from the cloud-top pressure using NCEP analyses data (Menzel et al., 2006).
To avoid further ambiguities the comparison was only performed for cloudy pixels that are classified in both cloud masks as being fully cloud-covered. As for the cloud fractional cover we included only data with a maximum difference of observation the time of 30 minutes. The spatial distribution of collocation data with respect to the observation time and the time difference between SEVIRI and MODIS scans can be seen in Figure 4..

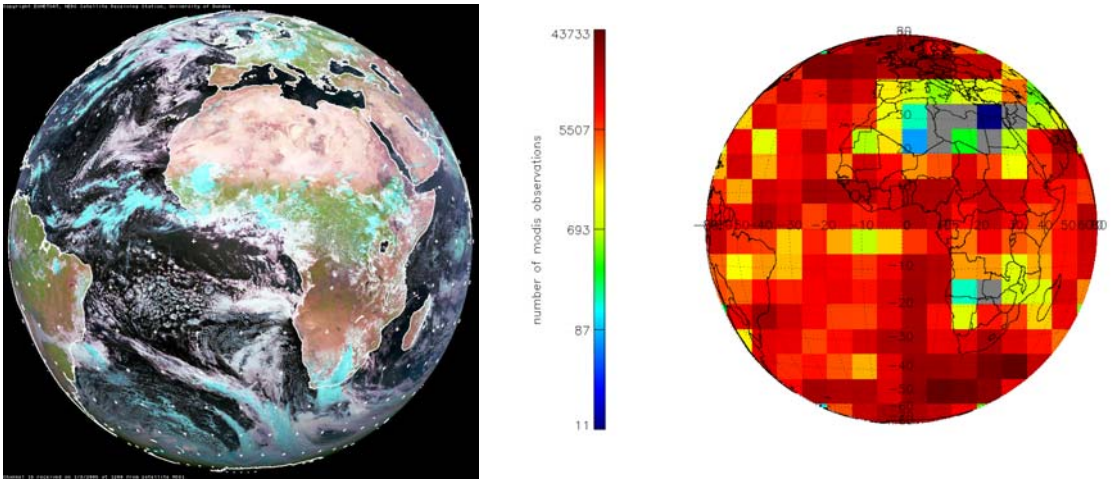

Figure 6.3 *False colour rgb image using SEVIRI channels 1-3 (provided by NERC satellite receiving station, Univ. of Dundee, UK) of SEVIRI from 1st August 2006, 12:00 UTC (left panel) and spatial distribution of the number of SEVIRI pixels used for CTP retrieval (right panel).*

Figure 6.3 (left panel) provides a false-colour impression of the geographic distribution of cloudy areas (high clouds in cyan, low clouds in white). The distribution of cloudy pixels on the disk (Figure 6.3, right panel) follows the typical pattern with low cloudiness over the known desert areas (Saharan, Namib) and more clouds towards the polar regions.

Special attention is with those areas where a positive bias of CTP is observed (Figure 6.4). These are in both hemispheres covered by scattered cumuli cloud fields at low altitude (as seen by SEVIRI).

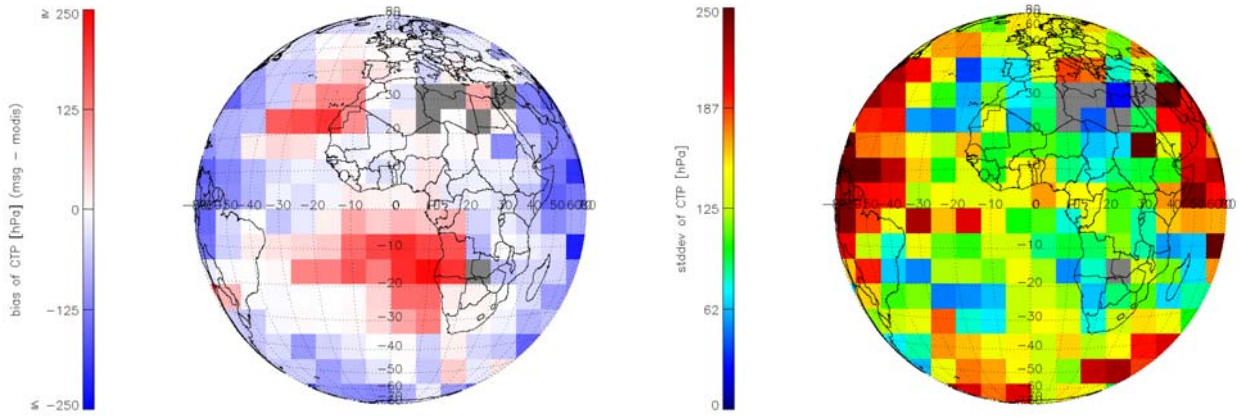

Figure 6.4 *Bias of the cloud-top pressure derived from SEVIRI and MODIS observations (left panel) and corresponding standard deviation (right panel). Observations are from 1st August 2006.*

As can be seen from Figure 6.4 these areas in the subtropical dry air zones stand out due to large differences between the SEVIRI and MODIS cloud-top pressures, respectively. Here, the SEVIRI result is much higher or the cloud-top (in km) is much lower in the atmosphere than for MODIS. On the other hand, there is an obvious trend towards the edges of the MSG disk in the sense that SEVIRI underestimates the cloud-top pressure and overestimates the corresponding cloud-top height. While the latter effect can be addressed to the increasing scenery effect for large observation zenith angles it is more difficult to explain the deviations along and in the trade wind zones.

As explained before we have to consider the impact of NWP data but it is likely that also large parts of the differences are due to missing cirrus clouds in the SEVIRI records. MODIS capabilities to detect thin cirrus clouds are advantageous (by the use of several CO_2 bands) while the capabilities of SEVIRI to detect optically thin cirrus clouds above optically thick water clouds are limited. Such cloud structures occur frequently in the subtropical and tropical regions. The cloud-top heights of convective clouds and a possible cirrus veil above are however likely almost the same and hence we do not observe large deviations of the cloud-top pressure in the tropical belt.

Detailed results are summarized in Table 6.2. We distinguish between low (> 700 hPa), mid-level (300 hPa < level < 700 hPa) and high clouds (< 300 hPa). We present also results for a common cloud mask, i.e. both the satellite sensors identified a scene as being fully cloud covered. As can be seen from Table 6.2 this holds only for two thirds of all SEVIRI observations and less than half of all MODIS observations. The overall bias using all pixels is about -114 hPa which is a rather large difference. If we however analyse only those pixels where both the cloud masks indicate a fully cloudy pixel the bias is remarkably low (-9 hPa), although the importance and the impact of pressure differences depend strongly on the height level where they occur. The overall standard deviation is 167 hPa (common pixels only) which indicates the large scatter of results. The standard deviation is higher over water surfaces (175 hPa versus 143 hPa) but there is no difference between daytime and night-time results. We further notice that the standard deviation is low where the bias is high and positive but high where the bias is strongly negative. As expected, the scatter increases towards the edges of the MSG disk where the bias becomes negative. The bias of high and mid-level clouds is surprisingly good (5 and 4 hPa, respectively) while SEVIRI low clouds are seen at higher cloud-top pressure (low cloud-top height). “Weighted” results which are

Table 6.2 *Detailed results of cloud-top pressure derived from SEVIRI and MODIS retrievals for different geophysical scenarios. Values are in hPa. SEVIRI = all SEVIRI observations, MODIS = all MODIS observations, common = cloudy pixels from both retrievals, avg Mo = average MODIS all pixels, avg Mo c = average MODIS common pixels, bias = SEVIRI – MODIS all pixels, bias c = SEVIRI – MODIS common pixels, std c = standard deviation common pixels, bias c h = bias common pixels high clouds, std c h = standard deviation common pixels high clouds, bias c m = bias common pixels mid-level clouds, std c m = standard deviation common pixels mid-level clouds, bias c l = bias common pixels low clouds, std c l = standard deviation common pixels low clouds.*

scenario	SEVIRI	MODIS	common	Avg Mo	Avg Mo c	Bias	Bias c	Std c	Bias c h	Std c h	Bias c m	Std c m	Bias c l	Std c l
overall	3078442	4828196	2226788	684	535	-114	-9	167	5	88	4	161	-38	211
weighted	429256	675594	305701	682	532	-120	-19	170	5	89	-5	157	-53	219
land	835776	1380323	575219	564	394	-149	-24	143	8	78	-21	126	-128	249
sea	2226111	3420587	1640792	730	585	-101	-3	175	1	95	12	170	-27	203
coast	16555	27286	10777	671	486	-148	-19	139	11	91	-20	136	-49	175
day	1401622	2308721	1010024	700	549	-117	-8	169	13	99	8	150	-40	211
night	1643892	2485417	1193200	665	525	-100	-10	167	-1	78	3	169	-36	210
twilight	32928	34058	23564	553	481	-84	-75	159	51	107	-79	137	-132	230
Mountain	23335	35299	13010	447	384	-95	-53	138	6	82	-75	145	-321	220

taking into account the inhomogeneous spatial distribution of collocation pixels are similar. As expected, the coincidence of CTP results is less good over mountainous terrain and under twilight

conditions (large negative bias). Daytime and night-time measurements perform equally well which is not surprising since solar channels are no used for the cloud-top pressure retrieval.

The target accuracy of SEVIRI CTP (10%) is easily reached, although initially formulated for daily and monthly mean values.

Two-dimensional scatter diagrams allow a quick view on the main differences of CTP values (Figure 6.5). It becomes obvious that SEVIRI overestimates the cloud-top pressure of low clouds between 700 hPa and 1000 hPa which is also seen in the difference CTP as function of the MODIS CTP (Figure 6.5, right panel). On the other hand, high clouds at low CTP are underestimated by SEVIRI, i.e. the cloud-top height is even higher than that for corresponding MODIS observations (Figure 6.5, left panel, orange square).

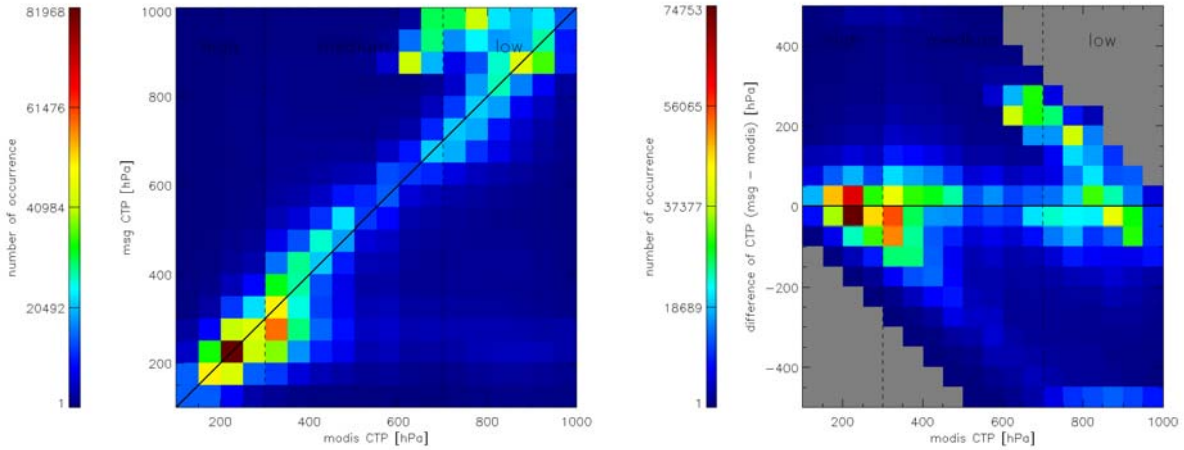

Figure 6.5 *The left panel shows a scatter diagram of the MODIS CTP (x-axis) versus the SEVIRI CTP (y-axis) and the right panel shows the difference of CTP (y-axis) as a function of the MODIS CTP. All observations from 1st August 2006. Boundaries low/medium and medium/high of CTP are at 700 hPa and 300 hPa respectively.*

A further analysis of the cloud-top pressure results as function of the latitude and the observation zenith angle of SEVIRI is shown in Figure 6.6. There is strong negative bias towards the Antarctic area and also with increasing solar zenith angle. A positive bias is observed in the Southern subtropical region (see also Figure 6.6, left panel) where the overestimation of the cloud-top pressure over the South Atlantic Ocean along the African coast (20” south) is not compensated by an underestimation of CTP in other regions at the same latitude. The bias at 20° - 30° north however is almost zero.

Note that such compensating effect is responsible for the rather good overall coincidence of results (low bias) while larger systematic differences occur over several regions.

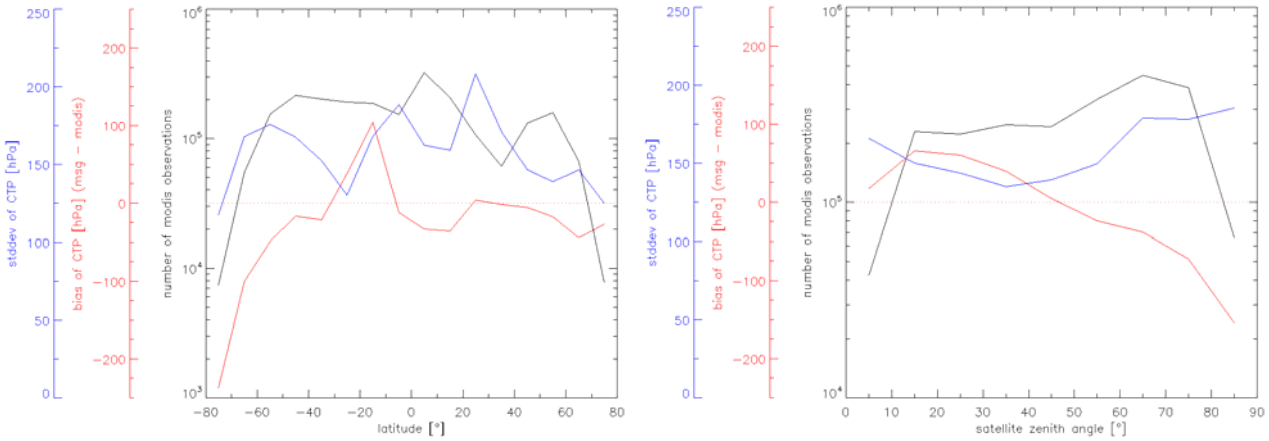

Figure 6.6 *Standard deviation (blue), bias (red) and number of observations (black) of the cloud-top pressure as function of latitude (left panel) and SEVIRI observation zenith angle (right panel).*

As for the cloud fractional cover there is a similar strong dependency between the bias and the local mean observation time (Figure 6.7, right panel). Values are worse for local sunset and sunrise conditions but the informational value is limited due to the limited number of available observations. The standard deviation is lowest for moderate sun zenith angles between 30 and 50 degrees whereas the course as function of the local mean observation time is more complex. Highest values are seen around 8:00 and 20.00 but only for the latter time the number of observations is low as well which may explain the large standard deviation. However, even for a huge number of collocation pixels between 10:00 and 14:00 the standard deviation remains around 150 hPa and indicates that CTP results from SEVIRI and MODIS measurements differ remarkably.

The bias depends weakly on the solar zenith angle (Figure 6.7, left panel) while the standard deviation is minimal for solar zenith angles between 35 and 55 degrees. The more complicated pattern of bias and standard deviation in Figure 6.7 (right panel) is a consequence of the sun-synchronous MODIS observations. Strong negative biases go in line with observations around 06:00 and 18:00 local mean time which were made in the polar regions under unfavourable observation conditions for SEVIRI.

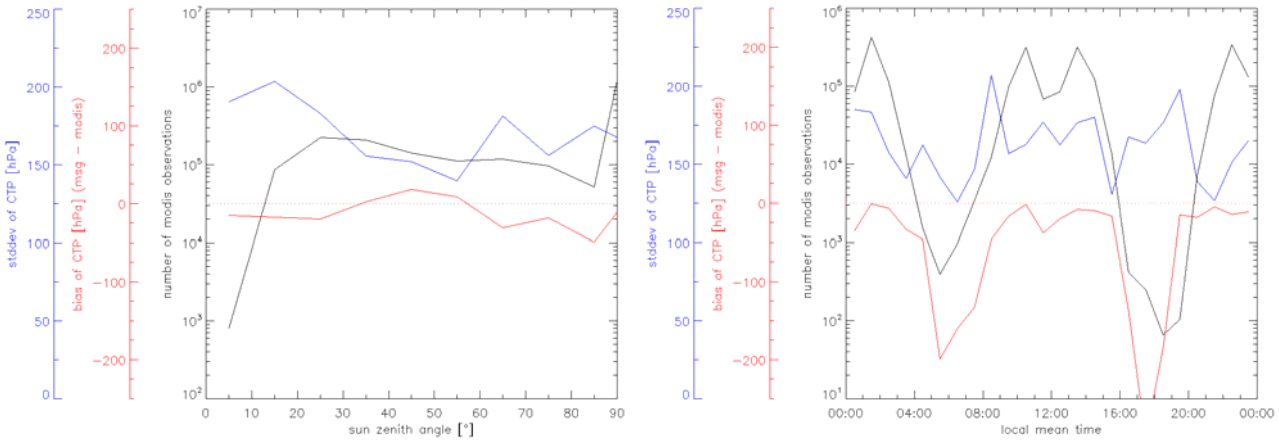

Figure 6.7 *Standard deviation (blue), bias (red) and number of observations (black) of the cloud-top pressure as function of the solar zenith angle (left panel) and the local mean time of observations (right panel).*

6.5 Comparison against CALIPSO observations

We compared SEVIRI and CALIOP cloud-top height results on the full visible MSG disk for observations of 1st August 2006. As for the comparison of CFC values we selected CALIOP pixels of homogeneous scenes (defined by the low standard deviation of surrounding IIR images at 10.6 µm) and one atmospheric layer. Under these boundary conditions the cloud-top height can easily be determined by CALIOP through the analysis of the time difference between the outgoing LIDAR signal and the registered backscatter signal. We took into account only fully cloudy pixels from both the sensors.

We first show the spatial distribution of the bias (Figure 6.8, left panel) and the corresponding standard deviation (Figure 6.8, right panel). There is a positive bias towards the edges of the MSG disk but SEVIRI tends to underestimate the cloud-top height elsewhere. The standard deviation increases towards the edges of the MSG disk. We observe also higher values over land than over water surfaces, although more observations would be required to consolidate this result.

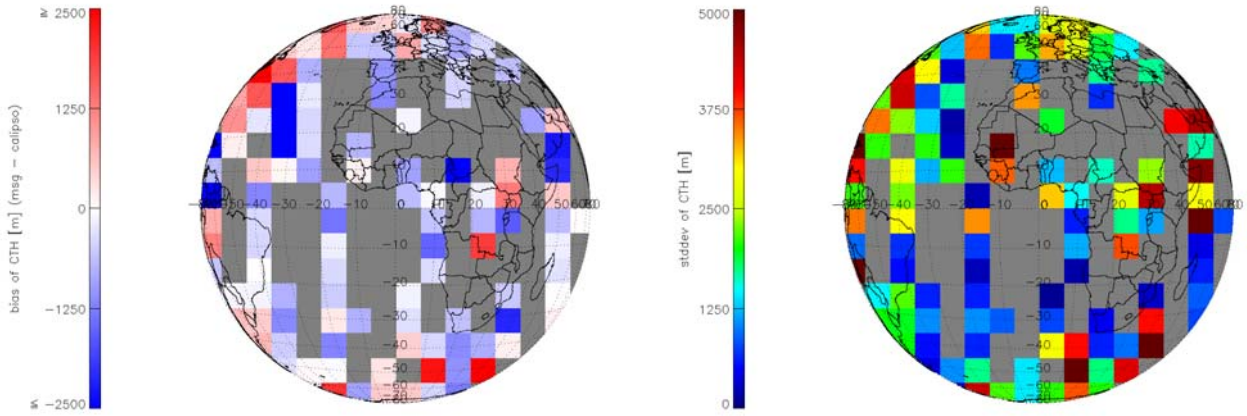

Figure 6.8 *Bias (left panel) and standard deviation (right panel) of CTH measurements derived from SEVIRI and CALIOP observations on 1st August 2006.*

We analysed the bias and the standard deviation as function of the satellite zenith angle (Figure 6.9, left panel) and we confirm the decreasing bias with increasing SEVIRI observation zenith angle. In fact, this shows that SEVIRI results underestimate the cloud-top height which is partially compensated by scenery and parallax effects at higher latitudes.

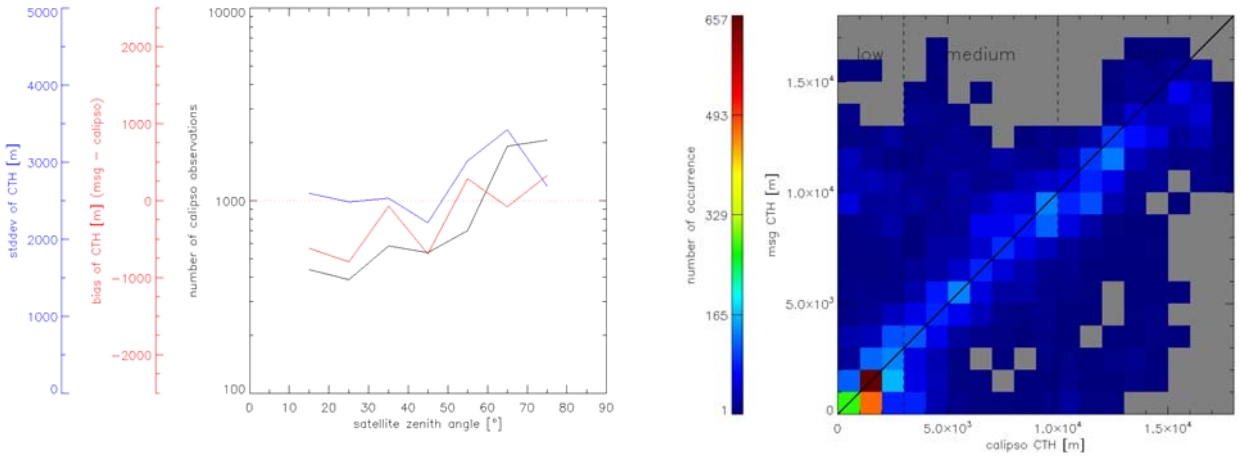

Figure 6.9 *Bias and standard deviation of retrieved CTH as function of the SEVIRI observation angle (left panel) and scatter diagram (right panel) of CTH (x-axis = CALIOP, y-axis = SEVIRI).*

We further distinguish between low clouds (CTH < 3000 m), mid-level clouds (3000 m < CTH < 10.000 m) and high clouds (CTH > 10.000 m). The scatter diagram shows that the height of low clouds is often overestimated by SEVIRI (Figure 6.9, right panel) but the opposite results is found for the height of high clouds which tend to be lower than if retrieved from CALIOP observations. Remember that we observed an increasing bias of CFC results for very low and very high clouds, as discussed in section 4.5.

Table 6.3 *Detailed results of cloud-top height derived from SEVIRI and CALIOP retrievals for different geophysical scenarios. Values are in m. SEVIRI = all SEVIRI observations, CALIOP = all CALIOP observations, common = cloudy pixels from both retrievals, Avg CAL = average CALIOP all pixels, Avg CAL c = average CALIOP common pixels only, bias = SEVIRI – CALIOP all pixels, bias c = SEVIRI – CALIOP common pixels, std c = standard deviation common pixels, bias c h = bias common pixels high clouds, std c h = standard deviation common pixels high clouds, bias c m = bias common pixels mid-level clouds, std c m = standard deviation common pixels mid-level clouds, bias c l = bias common pixels low clouds, std c l = standard deviation common pixels low clouds.*

scenario	SEVIRI	CALIOP	com-mon	Avg CAL [m]	Avg CAL c [m]	Bias [m]	Bias c [m]	Std c [m]	Bias c h [m]	Std c h [m]	Bias c m [m]	Std c m [m]	Bias c l [m]	Std c l [m]
overall	8126	12420	6618	5294	6001	293	–42	2917	–1862	2459	151	2500	887	3045
weighted	390530	559792	305620	5457	5934	275	31	2865	–1978	2611	223	2137	999	3136
land	2090	3871	1698	7119	8037	412	–135	3063	–1968	2244	806	2919	1181	3082
sea	5976	8468	4866	4453	5272	433	–13	2863	–1797	2606	–140	2221	836	3034
day	3963	5695	3285	4646	5052	207	149	2787	–1626	2195	–8	2264	931	2966
night	4012	6551	3192	5873	7016	458	–248	3064	–2007	2594	303	2732	787	3183

If we consider only those pixels where both the cloud masks indicated the presence of clouds (column “common” and all further column with “c” in Table 6.3), we find a rather low overall bias of – 42 m. Considering the unequally distributed collocations in space and time (over land and sea surfaces, during the day and under night-time conditions) we get a comparably low bias of +31 m. SEVIRI slightly underestimates the cloud-top height over land surfaces and under night-time conditions but a weak positive bias is found for daytime measurements. The standard deviation however is rather high for all scenarios (column Std c). We observe a systematic negative bias for

high clouds (column Bias c h) for all scenarios and a smaller positive bias for low clouds (Bias c l). The target accuracy of 10% is easily reached for CTH results based on the common pixels where the cloud masks agree.

Classes “coast”, “mountain” and “twilight” occurred rarely and these not meaningful results were omitted in the summary table above.

6.6 Summary

The validation of cloud-top height from instantaneous MSG/SEVIRI data using cloud radar observations showed that SEVIRI observations seem to overestimate the cloud-top height, especially if semitransparent clouds are present. However, for such cases and especially for multi-layer cloud scenarios the satellite observations and radar observations cannot be reasonably compared. A single-layer semi-transparent cirrus cloud will be comparably well detected by radar (due to the large effective sizes of ice crystals) but not necessarily by SEVIRI. In the case of multilayer clouds (thin cirrus above optically thick water cloud) the situation could be more problematic for the cloud radar but then also the satellite retrieval will encounter problems. An underestimation of the cloud-top height might occur when the semi-transparency correction fails (i.e. if the semi-transparent cloud is not detected by SEVIRI) and the radiance of the optically thick cloud below is dominating the measurement.
The performance improves for daily averages but known problems basically remain. The optional use of both the MSG/SEVIRI product quality flag and the usage of radar multiple layer information did not improve the results in this respect. Results for Cabauw compare well to previous results presented in the recent validation report (AD 7) while improved results are available for Chilbolton.

For other cloud top parameters we compared radar observations from Chilbolton (UK) and Cabauw (The Netherlands) with corresponding SEVIRI observations for four months in 2004. In general, we confirm the main results of the previous validation study with slightly different cloud type results since we concentrated on instantaneous daytime measurements rather than on the standard daily average products. While scenarios with high-level and low-level clouds agree well problems increase for mid-level clouds which are difficult to retrieve from radar data in the sense that results are finally comparable to satellite observations.

MODIS and SEVIRI cloud-top pressure results stem from different algorithm approaches and are in the case of SEVIRI partially influenced by NWP data. MODIS and SEVIRI results were filtered in the sense that only fully cloudy pixels for both the sensors were analyzed. Otherwise, the impact of the different cloud masks became very high. The overall bias is low (< 10 hPa) while the standard deviation of about 170 hPa is rather high. Depending on the height level of typical clouds this translates into a statistical variability of a cloud layer of more than 3 km. SEVIRI CTP is often higher for low clouds between 700 hPa and 1000 hPa but the height of high clouds is also often overestimated (SEVIRI CTP lower) if compared to MODIS results. There are large regional differences of CTP results:
SEVIRI CTP is systematically higher over water surfaces in the Midlatitudes (Northern and Southern hemisphere) and there is a clear trend towards a negative bias (SEVIRI CTP lower) for increasing satellite observation zenith angles and for high southern latitudes. As discussed above the first effect may be caused by thin cirrus clouds above water clouds which are not detected by SEVIRI although the SEVIRI cloud mask does not really support this explanation (see Figure 4., left panel). There, the majority of partially cloudy pixels is North of the equator in the ITCZ which is outside the regions where we observe a large positive bias. Another possible explanation would be the impact of NWP data. The “limb darkening” effect of the increasing negative CTP bias towards the edges of the disk may partly be due to scenery and parallax effects.

CALIOP and SEVIRI cloud-top height results were compared under the same premises: SEVIRI results are influenced by NWP data and we concentrated on fully cloudy pixels only. We found again a low overall bias of the cloud-top height of below 50 m that goes along with a high standard deviation of more than 2900 m. In contrast to the CTP results of the SEVIRI-MODIS intercomparison we observe an overestimation of the SEVIRI cloud-top height relative to CALIOP.

7 Validation of cloud phase – CPH

This section is divided into the following paragraphs:

- Introduction
- Validation task
- Validation against ground-based measurements
- Comparison against MODIS observations
- Summary
- Discussion

7.1 Introduction

This product provides information on the cloud thermodynamic phase (CPH). CPH is derived from 0.6 and 1.6 μm reflectance with an additional cloud top temperature check, derived from the 10.8 μm brightness temperature, for cloud flagged pixels initially labelled as "ice". More details on the retrieval method are provided in section 3.4.

7.2 Validation method

The CPH product was tested using one year (May 2004-April 2005) of data from SEVIRI for comparison with Cabauw cloud phase determinations. The validation of the CPH product was performed using the KNMI local processing environment. For validation at the Chilbolton site, a four month period (May-August 2004) was used.

SEVIRI data was processed at 15-minute resolution, giving 5841 and 984 cases for Cabauw and Chilbolton, respectively. For all months, retrievals between 7 and 17 UTC were included in the dataset. A validation area of 3x3 pixels (~250 km^2 at 50^o N) around Cabauw and Chilbolton was used for both the instantaneous and the daily and monthly mean validation.

Ground-based cloud phase retrievals were obtained from the Cloudnet target classification product (see Illingworth et al. 2006 for further details), which uses cloud radar vertical Doppler velocity, LIDAR attenuated backscatter coefficient, and NWP temperatures as input. The NWP wet bulb temperature 0^o C level is used as a first estimate for the level of the melting layer, which subsequently is refined using the cloud radar vertical Doppler velocity. The LIDAR attenuated backscatter profile is used to detect thin layers of super-cooled water within ice layers. The ground-based cloud phase is determined with a vertical resolution of ~90 m, and at a temporal resolution of 15 s for Cabauw and 30 s for Chilbolton. For each sampling period, the five upper gates (being ~450 m) of the ground-based cloud phase were examined. Further, in order to exclude mixed phase effects, only sample periods where at least 80% of the gates investigated had the same phase were included in the dataset. All data in a 20-minute time window centered at each satellite time slot was collected.

Instantaneous CPH retrievals were validated analogously to previous validation reports. That is, in order to be included in the comparison, at least 7 out of 9 pixels of the validation area were required to be cloud flagged. Furthermore, at least 80% of the cloud flagged pixels needed to be of identical phase. A similar approach was applied to the ground-based observations. The methodology above resulted in one cloud phase for satellite and ground-based observations for each 15-minute time slot. The resulting dataset comprised four categories: satellite water and surface water (a), satellite water and surface ice (b), satellite ice and surface water (c), and satellite ice and surface ice (d). These four categories were used as input for calculation of the monthly Kuipers skill score (Hanssen and Kuipers, 1965, see for its formula section 4.3).

To assess the suitability of the CPH product for longer time periods, the product output was evaluated at daily and monthly scale. This was done by calculation of daily and monthly liquid water phase ratios, φ_d and φ_m, respectively. For a 3x3 pixel area centred at the Cabauw and Chilbolton sites (~250 km^2), the number of cloudy pixels labelled as water at each day is divided by the number of cloudy pixels. Likewise for the ground-based measurements; the number of sampling periods with water at the cloud top was summed and divided by the total number of cloudy sample periods over each day.
The monthly liquid water phase ratio was obtained as follows. For each day a weighting factor was calculated using the number of collocated time slots:

$$w_d = \frac{n_d}{\sum n_d}$$

Subsequently, the daily liquid water phase ratios, φ_d, from CPH and ground-based observations were weighted by w_d and summed over the number of available days to obtain the monthly mean liquid water phase ratio, φ_m. Since the ice phase ratio is 1-φ, the absolute value of the difference between CPH and the ground-based derived ice phase ratio will be equal to the difference value for the water phase ratio. Consequently, only results for the liquid water phase ratio are presented.

7.3 Results ground-based comparison

7.3.1 Instantaneous CPH retrievals

Figure 7.1 shows the Kuipers skill scores found for the comparison of the CM-SAF CPH product and the CloudNET target classsification algorithm at Cabauw and Chilbolton. The skill of the CPH product is above zero throughout the year. Furthermore, it can be seen that after scores of ~0.70 during the summer months, which indicate good agreement (misclassification typically ~20%), skill scores gradually decline towards the winter months, having a minimum of 0.35 (misclassification ~50%) in January 2005. This minimum is followed by an increase to a Kuipers skill score of 0.85 (misclassification ~13%) in April 2005.

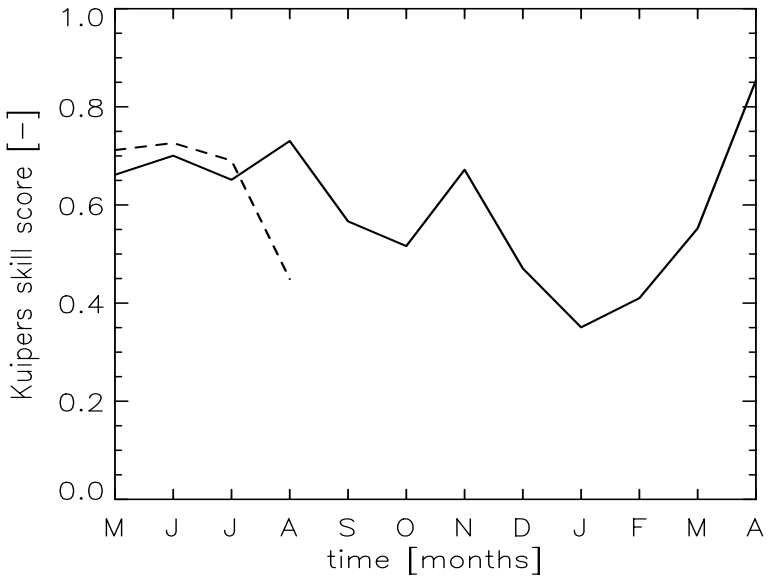

Figure 7.1 *Kuipers skill score of the CPH product for Cabauw (solid line) and Chilbolton (dashed line).*

The decline of skill of the algorithm in winter may be attributed to several factors: 1) a relative increase in the frequency of occurrence of backscatter geometries during the daylight period, 2) a

higher sensitivity of DAK LUT reflectance to changes in cloud properties at unfavorable viewing geometries during winter time, 3) possible deviations of the actual surface reflectance value from the MODIS surface albedo product value, such as sudden snowfall, and 4) a larger one-to-three dimensional reflectance difference at backscatter angles. Furthermore, in case of thin ice clouds over thick water clouds, the reflectance signal will be largely dominated by the water cloud. Since one-layer plane parallel clouds are assumed in the DAK radiative transfer calculations, this type of multi-layer cloud will very likely be labelled as water cloud. Although having only data available for four months, the results for Chilbolton are fairly similar to the Cabauw results, with August 2004 as an exception, having a Kuipers score of 0.15 lower than for Cabauw. It is not clear what the cause of this spurious decrease in skill is.

7.3.2 Daily and monthly CPH values

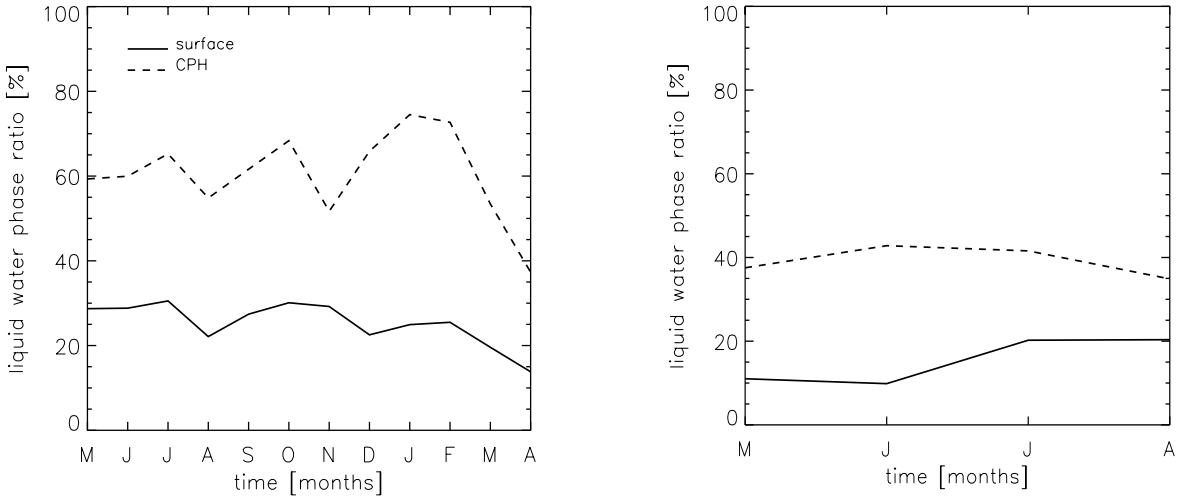

Figure 7.2 *Monthly mean liquid water phase ratio, φ_m, for Cabauw (May 2004-April 2005, left panel) and Chilbolton (May-August 2004, right panel). CPH values are denoted by dashed lines, surface values by solid lines.*

Figure 7.2 shows the weighted monthly averages of liquid water phase ratio (the number of clouds labelled as water to the total number of clouds) for Cabauw (left panel) and Chilbolton (right panel). Since for Cabauw an entire year was available, focus will be set to this dataset.

Throughout the year investigated, a large and systematic positive bias between the CPH and the surface observations exists. This bias increases from ~30% during the summer months towards 49% during January 2005. The increase in bias is partly accounted for by the reasons highlighted earlier. It was shown by Wolters et al. (2006) that (seasonally) averaged liquid water phase ratio derived from 0.6 and 1.6 μm reflectance has a reasonable bias in the summer months (~10%), but exhibits a strong increase in bias towards the winter (35-40%). The weighted averaged bias between CPH and ground-based derived liquid water phase ratio is 33% for Cabauw and 26% for Chilbolton.

The systematic positive bias between CPH and ground-based liquid water phase ratio is likely explained by differences in the abilities to detect thin cirrus clouds. SEVIRI is not able to detect cirrus with an optical thickness smaller than ~0.5, while cloud detection from the combined LI-DAR-radar approach is much more sensitive to this type of cloud. For Cabauw it appeared that the yearly averaged difference in detected thin cirrus with cloud geometrical thickness < 2000 m (which corresponds to τ ~ 0.5 for ice clouds, the ice cloud optical to geometrical thickness relation

was estimated from simultaneous measurements of these quantities by Matrosov et al., 2002) between satellite and surface is ~22% (Wolters et al., 2006).

Despite the large positive bias between CPH and ground-based liquid water phase ratio, the annual cycle in φ_m as detected by the surface instruments is re-produced reasonably by CPH. To quantify this further, the daily liquid water phase ratios derived from CPH and ground-based observations were normalized by the yearly mean liquid water phase ratio. Subsequently, monthly averages were calculated, which are shown for Cabauw in Figure 7.3. The CPH product agrees fairly well with the ground-based observations during the first six months (May-October 2004). For this period, a correlation of 0.83 between the normalized daily satellite and ground-based liquid water phase ratio is found. From November onwards, the agreement between satellite and surface decreases, which is indicated by a decline in correlation towards ~0.60.

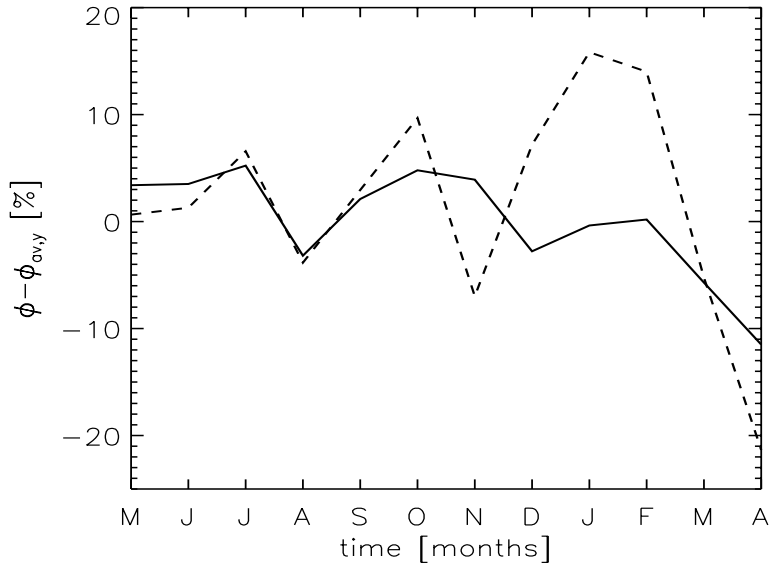

Figure 7.3 *Monthly averaged liquid water phase ratio normalized by the yearly average (φ-$\varphi_{av,y}$) for Cabauw as derived from CPH (dashed line) and surface observations (solid line), May 2004-April 2005.*

To further extend the previous analysis, the three-month normalized averages of satellite and ground-based liquid water phase ratio were also calculated; results are presented in Figure 7.4. As for the monthly normalized liquid water phase ratio, CPH has a fair ability to detect the seasonal changes in the amount of water during the periods MJJ and ASO. Although the above analyses are performed on a one-year dataset for Cabauw, it is emphasized that a longer dataset needs to be investigated to further analyze the abilities of the CPH product to detect trends in water or ice occurrence.

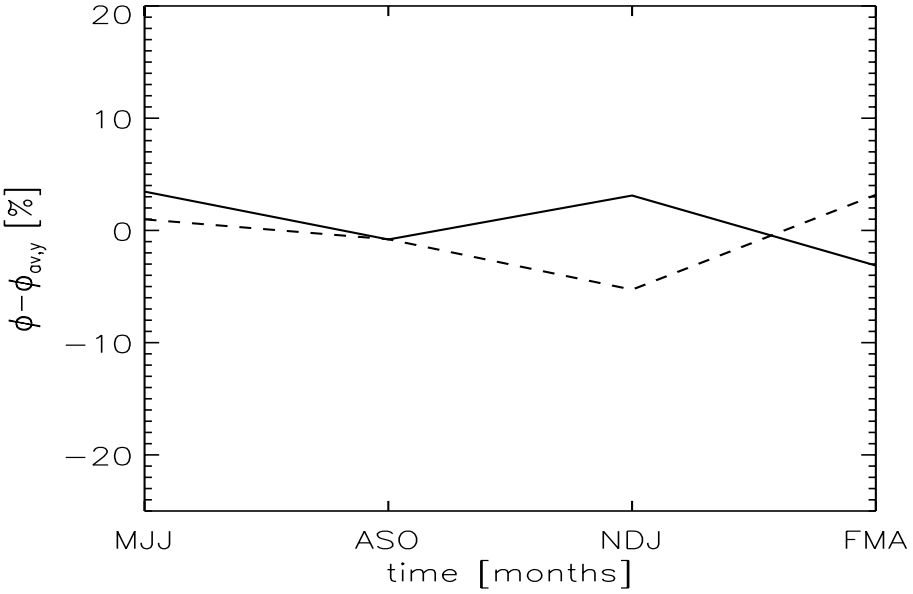

Figure 7.4 *Deviation of three-month average of φ to the yearly averaged φ derived from the CPH product (dashed line) and ground-based observations (solid line).*

7.4 Results comparison SEVIRI and MODIS

SEVIRI CPH retrievals were compared with cloud phase retrievals of the MODIS collection 5 infrared cloud phase determination algorithm, which uses 8.5-11 μm brightness temperature difference (the theoretical basis of this algorithm is discussed in Strabala *et al.*, 1994). The comparison in the tropical land and ocean areas can be regarded statistically firm, with ~2 million collocated cloud phase determinations over the period investigated. For the subtropical ocean area, collocation was more difficult. Due to less favourable viewing and scattering angle conditions only ~139000 collocated cloud phase retrievals were obtained. Nevertheless, the amount of data gives sufficient statistics.

Figure 7.6 shows the results of the comparison over the tropical land, tropical ocean, and subtropical ocean areas from top to bottom, respectively. The water, ice, and mixed phase ratios (only available for MODIS) are depicted on the left side, frequencies of differences found from collocated retrievals are shown on the right side.

Over the tropical land area, CPH labels ~14% more water clouds and ~10% less ice clouds than the MODIS algorithm. The remainder of the difference between the CPH and MODIS derived water cloud ratio is classified as mixed phase by MODIS. A comparison of the collocated retrievals in the right graph shows that the overall correspondence is ~77%, the misclassification of CPH water and ice retrievals is ~5% and ~15%, respectively. The large differences between CPH and MODIS in this area are likely caused by the spectral differences between the two algorithms. The MODIS cloud phase algorithm is much more sensitive to thin ice clouds, whereas ice clouds require a certain optical thickness to be determined from 0.6 and 1.6 μm reflectance. Furthermore, in case of optically thin clouds the measured reflectance or thermal infrared radiance is contaminated by the surface reflectance or emissivity, which is more variable over land than over ocean. In combination with the different spatial resolutions of the radiometers, this increases the probability for an erroneous cloud phase determination for both CPH and MODIS cloud phase.

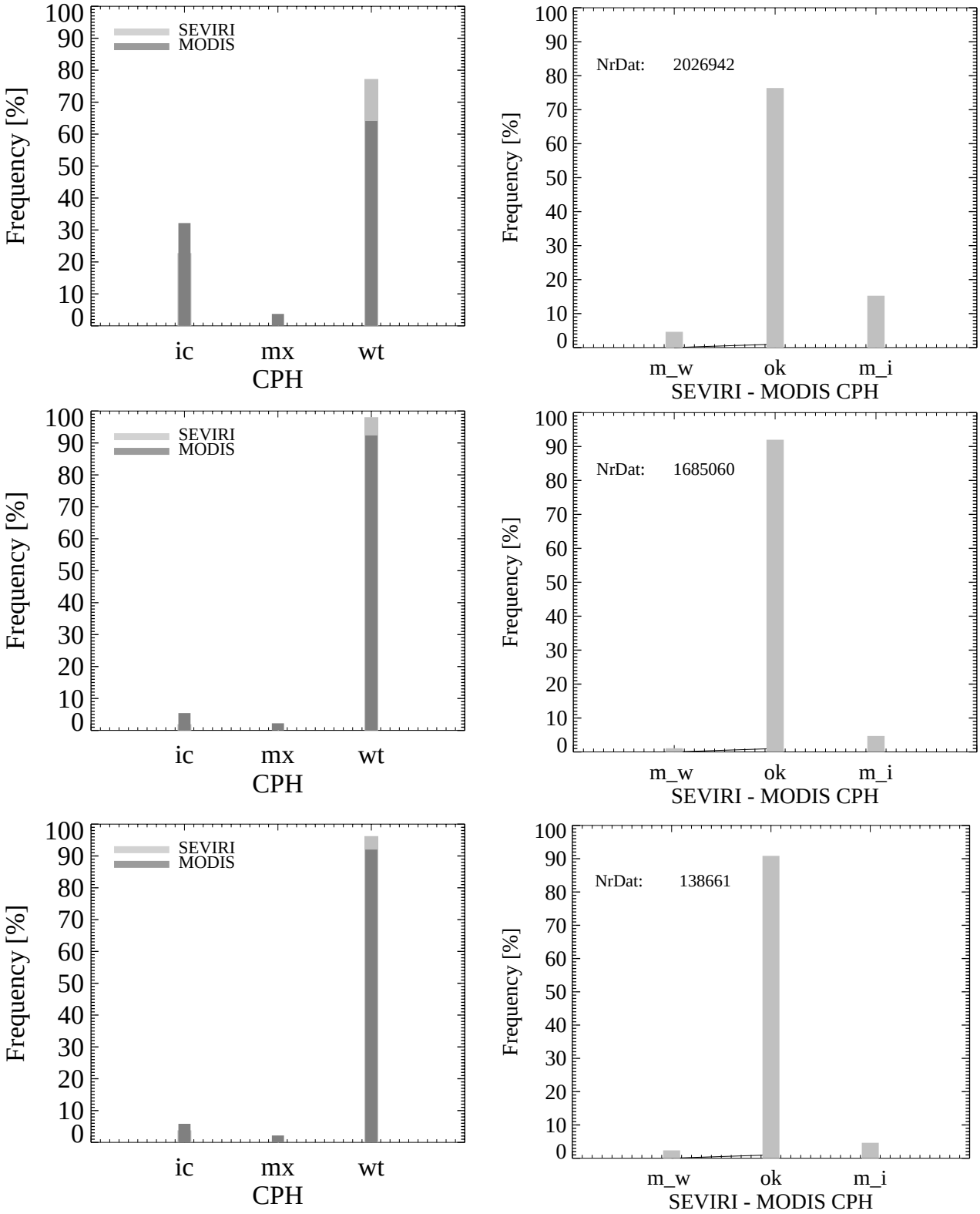

Figure 7.5 *Frequency distributions of SEVIRI CPH and MODIS cloud phase (left panels) and frequencies of differences (right panels) for collocated retrievals over the tropical land (upper panels), tropical ocean (middle panels) and subtropical ocean (lower panels) areas. Ic=ice, mx=mixed phase, wt=water; m_w=misclassified water by SEVIRI, ok=correspondence between SEVIRI and MODIS cloud phase retrievals, and m_i=misclassified ice by SEVIRI. See Figure 2.3 for the area boundaries.*

For the tropical ocean area, the derived percentages of water and ice clouds between SEVIRI and MODIS show a very good agreement. In total, ~92% of the collocated SEVIRI and MODIS retrievals are of identical phase. CPH classifies ~1% of the water and ~5% of the ice phase retrievals differently than MODIS. The latter cases are mainly labelled mixed phase by the MODIS algorithm (~3%).

Over the subtropical ocean area, the CPH algorithm detects ~4% more water than MODIS and ~2% less ice. The agreement between the collocated cloud phase retrievals is ~90%, misclassified water and ice by CPH are ~2% and ~5%, respectively. These results are similar to the results found for the tropical ocean area.

7.5 Summary

The cloud thermodynamic phase product (CPH) was validated for the European area with collocated cloud phase determinations from the Cloudnet sites of Cabauw (one year, May 2004-April 2005) and Chilbolton (four months, May-August 2004). The validation was performed on the instantaneous and the monthly mean product.

For the validation of the instantaneous product, only cases with cloud cover > 80% within a 3x3 pixel area around the surface validation stations were included. Furthermore, in order to minimize mixed phase effects, at least 80% of the cloud flagged pixels needed to be of identical phase. The CPH shows Kuipers scores of ~0.70 in comparison to the CloudNET data (indicating good skill, the corresponding misclassification by CPH is typically 20-25%) during the summer months, decreasing to ~0.40 (misclassification ~50%) during the winter months. The decline in skill is explained by 1) a relative increase in the frequency of occurrence of backscatter geometries during the daylight period, 2) a higher sensitivity of DAK LUT reflectance to changes in cloud properties at unfavorable viewing geometries during winter time, 3) possible deviations of the actual surface reflectance value from the MODIS surface albedo product value, such as sudden snowfall, and 4) a larger one-to-three dimensional reflectance difference at backscatter angles. Finally, the cloud phase determination from CPH deviates from the ground-based observations due to a different ability for detecting ice clouds of the respective instruments (SEVIRI radiometer versus cloud radar, 2^{nd} vs 6^{th} moment of particle size). While less data are available for Chilbolton (May-August 2004), they confirm the results for Cabauw.

The monthly mean liquid water phase ratio was determined for all cloud flagged pixels at each day; the respective days were weighted by the total number of time slots in a month. Results for Cabauw (May 2004-April 2005) showed a positive bias of ~30% in May 2004 versus the ground-based classification, increasing to 49% in January 2005, with a weighted yearly mean of 33%. For Chilbolton, the bias for the period May-August 2004 is 26%. The increase in bias during winter largely results from the different abilities for detecting thin cirrus clouds between SEVIRI and the ground-based instruments (cloud radar), which was estimated at ~22% (Wolters et al., 2006).

Comparison of CPH retrievals with cloud phase determination of the MODIS infrared cloud phase algorithm was done for three regions over Africa. Retrievals of both algorithms were remapped on a Mercator projection of similar pixel size for viewing angles of 0^{o}-50^{o}, scattering angles being 100^{o}-130^{o} and 140^{o}-175^{o}. In general, there is good agreement between collocated CPH and MODIS results. Over the ocean areas (comprising both ITCZ convection and quasi-persistent stratocumulus fields), agreement is more than 90%. This shows that CPH retrievals are reliable in the tropical region over homogeneous surface. The retrievals over tropical land show less agreement (~77%), which is probably caused by the different spectral regions used by the algorithms in combination with different surface characteristics (change in surface albedo versus surface emissivity).

7.6 Discussion

The Kuipers skill score of CPH is good for the instantaneous retrievals during the summer months. Towards the winter, skill decreases, as was also pointed out in the ORR V2 SR. The decrease in skill is mainly caused by the poor signal to noise ratio for the 0.6/1.6 μm reflectance test at low sun elevations. Furthermore, multi-layer cloud systems with thin ice over water clouds hamper the cloud phase determination algorithm used in CPH.

The monthly averaged liquid water phase ratio shows an increasing positive bias between CPH and ground-based derived values towards the winter months, with average values of 33% and 26% for Cabauw and Chilbolton, respectively. Large part of this bias is attributed to the difference in detection of ice clouds between SEVIRI and the ground-based instruments used. Depending on ice water content, particle size distribution, and ice crystal shape, a certain geometrical thickness of the ice cloud is required before it is detected by SEVIRI. Wolters et al. (2006) have estimated the difference in ice cloud detection between SEVIRI and surface at ~22% for the currently used Cabauw dataset.

From the comparison of CPH with the MODIS infrared cloud phase determination algorithm, over ocean areas a very good agreement was found. Over land, agreement is less with CPH detecting ~10% more water than MODIS. For optically thin clouds, the systematic positive bias in the detection of water clouds by CPH can partly be alleviated by adding thermal infrared spectral channels to the existing algorithm. From theoretical calculations, Minnis et al. (1998) showed that brightness temperature difference 11-12 μm is suited to derive cloud particle shape and phase for non-opaque clouds. Furthermore, Strabala et al. (1994) showed that ice and water clouds appear as distinct clusters when plotting 8.5-11 μm versus 11-12 μm brightness temperatures. However, validation of this method over long time periods with collocated surface data has not been performed so far.

Since multi-layer clouds are present in about half of the surface cloud observations (Tian and Curry, 1989), further improvement of the CPH product might be achieved by investigating the suitability to implement a multi-layer flag. Naisiri and Baum (2004) developed a technique to detect potentially overlapped clouds using a relationship between NIR reflectance and 11-μm brightness temperature from MODIS, which showed reasonable agreement with collocated airborne LIDAR observations. It is intended to spend a substantial part of the research on the issues highlighted above to improve the quality of the CPH product.

8 Validation of cloud optical thickness – COT

This section is divided into the following paragraphs:

- Introduction
- Validation task
- Validation against ground-based measurements
- Comparison against MODIS observations
- Summary
- Discussion

8.1 Introduction

The Cloud Optical Thickness (COT) is related to the vertically integrated cloud extinction at a reference wavelength, which is caused by cloud droplets for a given atmospheric column. The CM-SAF product provides information on the COT for pixels flagged cloudy by the cloud detection test (see chapter 4). As reference wavelength, 0.64 μm is used, which lies within the spectral region of the 0.6 μm SEVIRI channel. The COT is retrieved simultaneously with the effective radius of the cloud particles from the 0.6 and 1.6 μm SEVIRI channels based on the method of Nakajima and King (1990). More details about the retrieval are given in section 3.4.

8.2 Validation methods

The validation of the MSG/SEVIRI cloud optical thickness retrieval is carried out by a comparison with ground-based pyranometer measurements, and a comparison to MODIS-derived COT values. While the COT values of SEVIRI and MODIS are derived by the same methodology and can thus be directly compared, the validation with pyranometer data require further steps to link the COT and the output of the pyranometers, which measure the global downwelling solar irradiance.

Retrievals of COT based on global radiation measurements have been developed by several researchers (e.g. Leyonteva and Stamnes, 1994, Banard and Long, 2004) and have been used for the ORR-V2 validation (CM-SAF 2005). However, these retrievals require rigorous screening, because they only apply to completely overcast situations with homogeneous clouds (Boers et al., 2002). Otherwise, large biases and scatter will result, limiting the usefulness of ground-based COT retrievals for the validation of COT values from satellite, as the statistics are strongly affected by outliers. These effects are a consequence of the large temporal and spatial variability of clouds and the radiation field, of sampling effects due to differences in scales of the ground- and satellite-based measurements (Deneke et al., 2002, Li et al., 2005), and of the strongly non-linear retrieval function. Figure 8.1 clearly shows that the CM-SAF COT values, as well as any other COT product derived from satellite imagers based on reflected radiation, are highly sensitive to random and systematic errors, in particular at low and high values of COT. Large biases and scatter have to be expected for realistic estimates of the uncertainties in the retrieval inputs, e.g. for the sensor calibration, the radiative transfer calculations and the surface albedo estimates.

Due to these problems, the decision has been made for the ORR-V3 to compare the pyranometer data to results of the *Surface Insolation for Cloudy Conditions from SEVIRI* (SICCS) retrieval (Deneke et al., 2005). Using the CM-SAF cloud optical thickness and effective radius retrievals as input, as well as ancillary information on water vapour and surface albedo, the atmospheric transmission is estimated by the SICCS algorithm. For these steps, the DISORT radiative transfer code (Stamnes et al., 1988) is used to generate lookup tables of atmospheric transmission. This choice of model was made due to the availability of the Rapid Radiative Transfer Model (RRTM) correlated-k parameterization, which allows to account for gaseous absorption in broad-band

shortwave calculations (Mlawer et al. 1996). The atmospheric conditions and cloud properties are chosen identically to the conditions used for generation of the lookup-tables for the CPP retrieval. A climatological aerosol profile is added in the calculations, which is based on typical conditions found for the Netherlands.

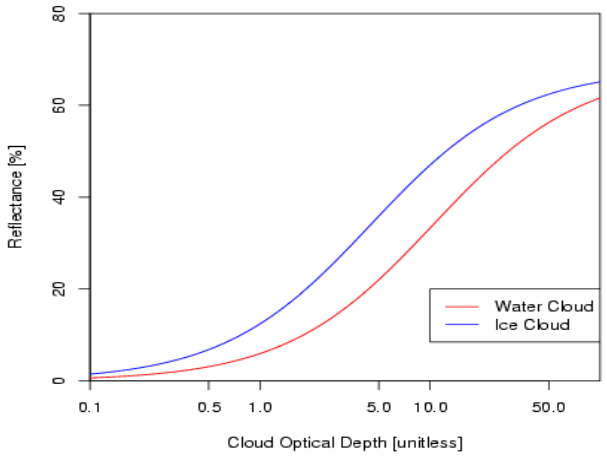

Figure 8.1 *Reflectance as function of COT, based on equation 23 of King(1987). Assumed are conservative scattering. Parameters have been chosen for a typical satellite geometry (for details, see section 7.3). The red and blue lines show the reflectance [in percent] for a water and ice cloud, respectively.*

As the atmospheric transmission is mainly determined by the COT (King, 1987), such a comparison also validates the satellite-retrieved COT, albeit more indirectly. However, we avoid the expected large differences and non-linearities in a direct inter-comparison of COT, as the narrow-band TOA reflectances and the broad-band atmospheric transmission are linearly related (Li and Leighton, 1993). Thus, the error statistics will be much more robust, but need to be translated to a corresponding uncertainty in COT a posteriori.

For the comparison, pyranometer measurements operated at the Cabauw, Chilbolton and Palaiseau CloudNet sites during the period May 2004 until April 2005 are used. For Cabauw, data are available only until the beginning of December 2004, due to an interruption of the measurements caused by the relocation to the new BSRN site, and a subsequent change in instrumentation and data formats. To account for the influence of water vapour absorption and surface albedo, a pre-operational dataset of CM-SAF's HTW product and the MODIS MOD43C 0.05 *degree* surface albedo product (Schaaf et al., 2002) are used as ancillary inputs.

For the comparison, the atmospheric transmission has been calculated from the global irradiances reported by the pyranometers, and is normalized by the incoming solar flux at the top-of-atmosphere. The values of transmission are then averaged to hourly values, and are compared with the satellite-retrieved values for a 2x4 pixel region around the pyranometer stations (corresponding to about 12x12 km^2). All four satellite image per hour are used to for the satellite retrieval to minimize temporal sampling errors (Nunez et al., 2005).

8.3 Results ground-based comparison

Results of the comparison between pyranometer data and the SICCS retrieval are shown in Figure 8.2. A high correlation of about 0.89 is found between SICCS results and measurements. The SICCS results underestimate the pyranometer-inferred transmission by 2.4%. On a case-by-case basis, a bias-corrected root mean square error of 9% is present. Figure 8.2(b) indicates the quantiles of transmission for pyranometers and SICCS retrieval agree well for thick clouds and the full range of solar zenith angles. For thin clouds and high values of transmission, however, an underestimate of the SICCS scheme by up to 8% is present. This behaviour is caused by an overestimation of COT for very thin clouds. Good agreement is found for clear-sky conditions (results not shown).

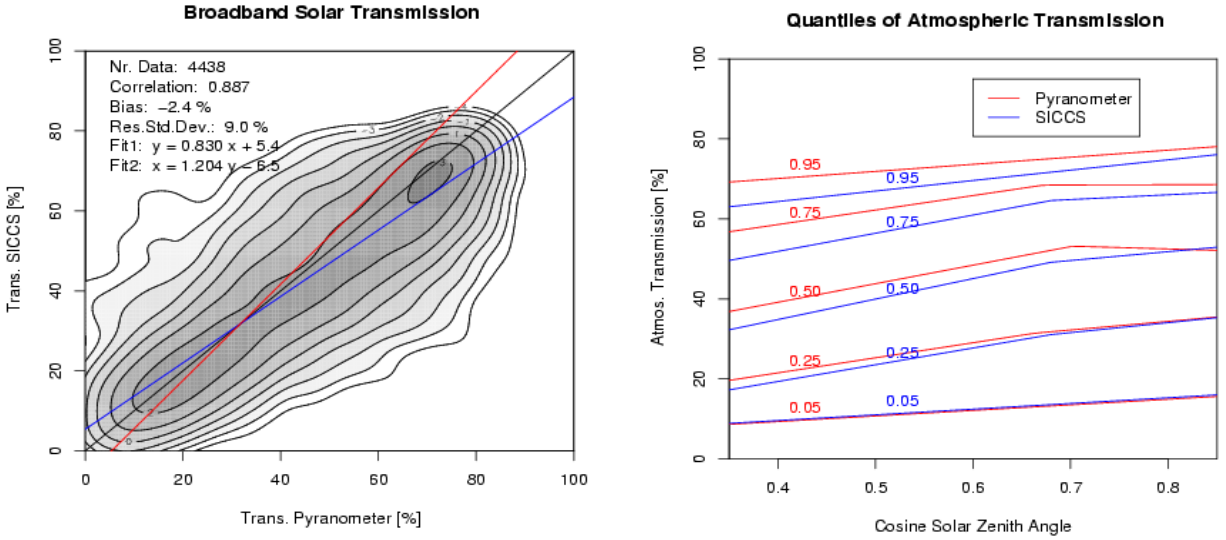

Figure 8.2 *(a) Density plot of hourly values of broad-band atmospheric transmission as measured by the pyranometer vs. 2x4 pixel average of SICCS-retrieved transmission. (b) Dependency of the quantiles of atmospheric transmission on the cosine of the solar zenith angle for pyranometer observations(red) and SICCS results(blue).*

The high correlation indicates that the COT values can provide accurate estimates of the broad-band solar atmospheric transmission and the variability induced by clouds. A bias of the transmission for thin clouds (COT less than about 8) of the SICCS results is responsible for the overall low bias of 2.4 % in transmission. As reason for this bias, an underestimate of TOA reflectance, as predicted by the combination of MODIS surface reflectance maps and the atmospheric correction provided by the DAK lookup tables, has been identified in comparison to observed SEVIRI reflectance for cloud-free situations. While no retrieval of cloud properties is carried out for cloud-free pixels due to the previously performed cloud detection, this leads to too large values of COT as soon as thin clouds are observed. The physical reason for this bias is currently the subject of further investigations, and is expected to be resolved by a future update to the processing. Possible causes are differences in calibration of the MODIS and SEVIRI sensors affecting low reflectance values, bi-directional effects of the surface reflectance, or shortcomings in the DAK radiative transfer calculations. It is expected that a fix will only require an update of either the MODIS surface albedo maps or the lookup-table input files, and no changes to the program itself. Until this problem has been fixed, users should be aware of the existence of a positive bias at COT values below about 8.

To link the results found for the broad-band atmospheric transmission to the uncertainty in COT, the form of the retrieval function, e.g. the relation between transmission and reflection, needs to

be considered. Here, the results of King (1987) are used and conservative cloud scattering is assumed. This analysis is based on asymptotic expressions for the reflection and transmission properties of the atmosphere, and derives an analytic expression for the error in cloud optical thickness. For our purposes, this should be sufficiently accurate. Furthermore, we assume that the scatter is of similar size both for transmission and reflection, and for the broad-band and narrow-band spectral regions. Choosing typical parameters[1] for equations 23 of King (1987), Figure 8.1 shows the reflectance as function of cloud optical depth.

The retrieval error is expected to be smallest in the linear region of the retrieval function at optical thicknesses of about 5 for the ice and 10 for the water cloud, were the slope is largest. The shift between water and ice clouds is caused by the difference in asymmetry parameters, and would disappear if the scaled optical thickness were used instead[2]. If 3 % and 5 % are assumed as typical errors for systematic and random errors in the reflectance, an accuracy of 10 % and 30 %, respectively, are found within a COT interval from 3 to 20 for ice and from 5 to 40 for water clouds. Misclassification of cloud thermodynamic phase can increase the retrieval error due to difference in asymmetry parameter for ice and water droplets (King, 1987).

8.4 Results comparison SEVIRI and MODIS

Results of the comparison of the CM-SAF and MODIS COT products are presented in Figure 8.3. The left panels show the frequency distribution of COT, while the right panels show the distributions of the absolute difference between retrieved COT from SEVIRI and MODIS. From top to bottom, results for the validation regions Tropical Land, Tropical Ocean and Subtropical Ocean (see section 2.3.3) are given.

The climatological distributions of retrieved COT compare favorably between the two satellites for all regions, with median values about 6 and 7. The SEVIRI retrieval seems to find systematically lower media values by about 10 percent, as compared to the MODIS results. The 66th quantile of the observed differences between SEVIRI and MODIS instantaneous COT are of similar magnitude as the median value, lying in a range between 6.0 and 8.0. This indicates the presence of a fairly large scatter which reduces the instantaneous precision of the retrieved COT values. Reasons for this scatter include resolution effects, temporal mismatch, parallax effects and geolocation errors. However, considering the sensitivity of the retrieval function to uncertainties in the input radiance (Fig 8.1), a scatter of this magnitude is to be expected. By averaging over longer time periods or larger regions, the precision increases significantly, as is evident from the good agreement of COT distributions found from both retrievals. Nevertheless, the overall accuracy of the retrieval results is limited by the accuracies of the sensor calibration and the underlying radiative transfer calculations.

8.5 Discussion and Summary

A comparison of the CM-SAF COT retrieval with ground-based pyranometer measurements and MODIS-retrieved COT values has been presented in this chapter. For both comparisons, good agreement of the CM-SAF COT product and the validation data is found. Based on the SICCS retrieval, the COT values can reproduce the atmospheric transmission with an accuracy of about 3%, and a random scatter of 9% for hourly averages. Little dependence of the accuracy of the retrieval on the solar zenith angle is found. As shown by Deneke et al. 2005, a strong reduction of

[1] Asymmetry parameter for ice/water clouds chosen as 0.65/0.85, respectively, reduced extrapolation length as 0.71, surface albedo as 0.1 at 0.6 µm, and product of the escape functions at solar and satellite zenith as 1.0.

[2] The scaled optical thickness is the optical thickness multiplied by a factor of (1-g), where g is the asymmetry parameter).

scatter occurs if results are averaged over longer times or larger regions. Both findings indicate that the retrieval accuracy will improve further for daily and monthly averages due to the reduction of sampling uncertainties and the cancellation of random error sources.

Directly comparing SEVIRI and MODIS-derived COT values, similar distributions are found, with a slight low bias of median COT of about 10% and climatological means of the COT having a value of about 7. The precision of individual COT values in comparison to the MODIS COT is reduced by significant random errors, which are of similar magnitude as the climatological mean (Q66 between 6 and 8). These errors can be reduced or even eliminated by averaging over larger regions or periods.

While part of this error might be attributable to the assumptions made in the retrieval input, a significant part is inherent in the error sensitivity of the retrieval method itself. This leads to large errors even for small temporal or spatial mismatches between SEVIRI-derived COT values and reference data. While these errors are random and are expected to cancel out to a large degree, the same sensitivity applies to systematic errors. The SEVIRI and MODIS calibration accuracies are expected to be no better than 3%, which can result in biases of up to 30 % (Roebeling et al., 2005).

A further important point to note for users of the COT data is the fact that the COT is an effective value for the whole pixel. Due to broken clouds and sub-pixel variability, the plane parallel bias likely affects the retrieval of COT (Marshak, 1994), as well as derived products such as cloud water path. Cahalan et al., 1994 propose a mean reduction of COT by a factor of 0.7 for stratocumulus clouds, which is used by the ECMWF general circulation model to account for this bias. However, the magnitude depends on cloud structure and sensor resolution, and is still the subject of active research. To enhance the accuracy, cases can be selected which are homogeneous, e.g. by constraining the standard deviation of cloud properties over a region of interest.

Despite the problems mentioned here, the COT product can provide an accurate estimate of the climatology of cloud optical properties and their impact on radiation, as is evident from the good agreement with both the MODIS retrieval and the ground-based measurements.

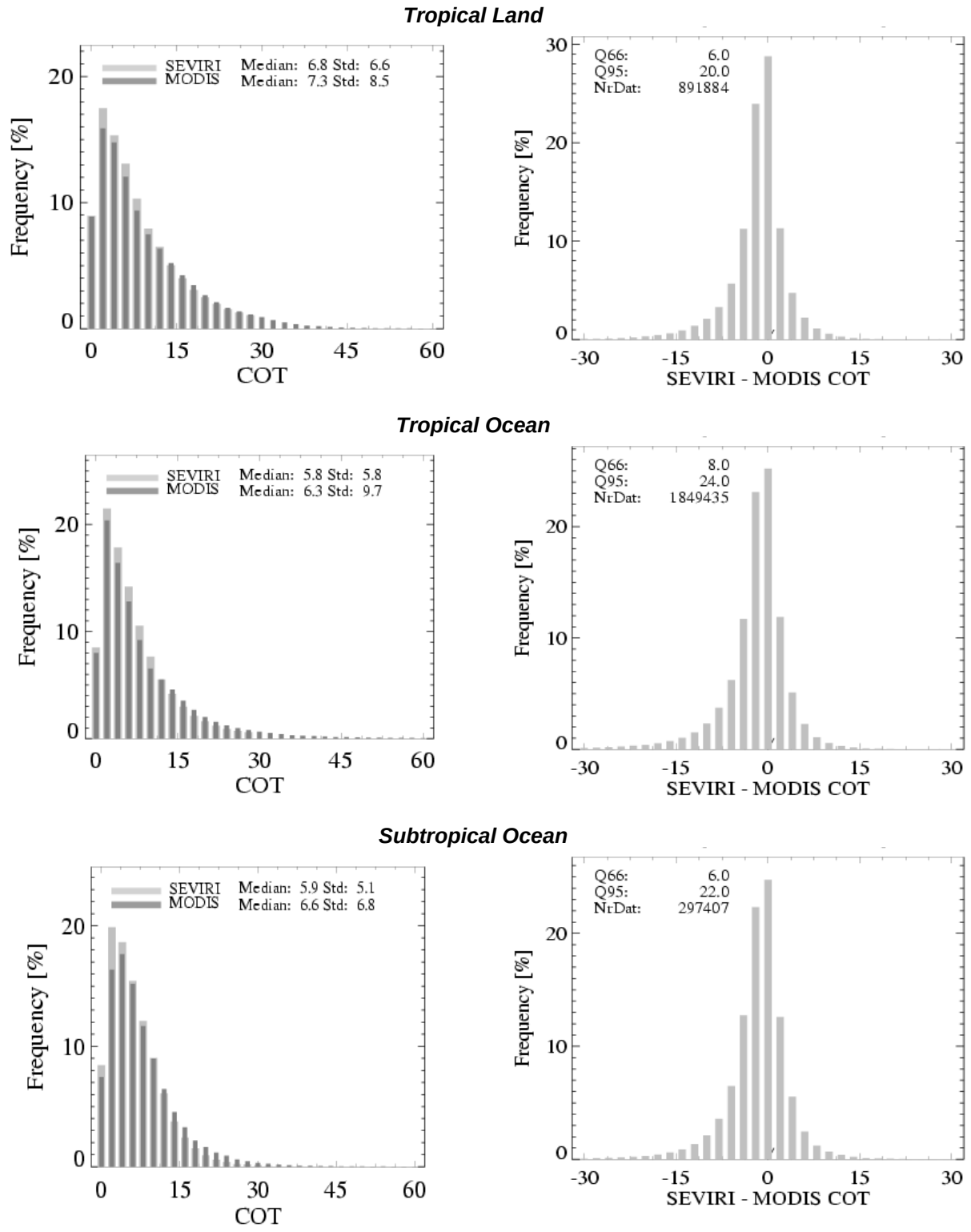

Figure 8.3 *Frequency distributions of SEVIRI and MODIS COT (left panels) and frequencies of differences (right panels) for tropical land (upper panels), tropical ocean (middle panels), and subtropical ocean (lower panels).*

9 Validation of cloud liquid water path – CWP

This section is divided into the following paragraphs:

- Introduction
- Validation task
- Validation against ground-based measurements
- Comparison against MODIS observations
- Summary
- Discussion

9.1 Introduction

This product provides information on the Cloud liquid Water Path (CWP) given in g m^{-2}. The CWP is derived from the cloud optical thickness (τ_{vis}) and the droplet effective radius estimates (r_e). More details on the retrieval method are given in section 3.4. Please note that in most presentations of CM-SAF the Cloud liquid Water Path product is abbreviated as CWP (as in the section heading). However, in this section the LWP notation is used to clearly indicate that this validation study does not cover the ice water path due to lack of suitable data for a comparison.

9.2 Validation method

The differences between the LWP retrievals from SEVIRI and MWR for the CloudNET sites of Chilbolton and Palaiseau are assessed for a summer period, covering May – August 2004. The LWP retrievals from MWR were averaged over 20 minutes. When Taylor's hypothesis of frozen turbulence is assumed and the windspeed is about 10 m s^{-1} this corresponds to a tracklength of about 12 km, which is considered representative for the field of view of SEVIRI (4x7 km^2). The LWP values from SEVIRI were retrieved at a temporal resolution of 15-minutes for the pixel that coincided with the ground station. The retrievals were done between 6 and 18 UTC at solar zenith angles smaller than 72°. During the summer, most observations had solar zenith angles smaller than 60° and scattering angles between 120 and 150°. Pixels that were identified as "clear certain", were excluded from the comparison. Because of the insensitivity of MWR observations to ice clouds, the comparison is restricted to water clouds. The cloud thermodynamic phase retrievals from SEVIRI were used to select observations with water clouds overhead the CloudNET stations. The analysis of the MWR retrieved LWP values was restricted to non-precipitating clouds with LWP values smaller than 800 g m^{-2}. The MWR measurements that were disturbed by rain were identified with rain gauge observations.

To evaluate the annual cycle of the accuracy and precision of the SEVIRI retrievals one year of MWR and SEVIRI retrieved LWP values were compared This comparison could only be done for Chilbolton, where MWR retrieved LWP and raingauge observations were available during the period May 2004 until April 2005. During this period, more than 3800 coincident sets of ground-based and SEVIRI retrieved observations were available. The comparison was restricted to the daily and monthly median LWP retrievals. The daily median LWP values were calculated for all days with more than six coincident sets of SEVIRI and MWR observations of LWP. The monthly median values were calculated from the instantaneous LWP retrievals from SEVIRI and MWR, which varied between 70 and 700 observations per month. There were no LWP retrievals from SEVIRI during the entire month of December 2004 and part of January because LWP was only retrieved at solar zenith angles smaller than 72°.

The statistics examined in this report include the mean and median of the LWP retrievals and the 50^{th} (Q50), 66^{th} (Q66) and 95^{th} (Q95) interquantile range of the deviation between the LWP retrievals from SEVIRI and MWR. The median is preferably used instead of the mean because the LWP is highly asymmetrically distributed. The Q50 is the difference between the 25% and 75%

quantiles of the deviations, Q66 and Q95 mutatis mutandis. The Q50 is an alternative measure of one standard deviation. The fact that the upper and lower 25% of the dataset are ignored makes Q50 a more robust estimator of variance than the standard deviation, and the preferred one for non-Gaussian distributions. The Q66 value is used to indicate twice the standard deviation, which would exactly be the case for a Gaussian distribution. In this study, the Q50, Q66 and Q95 values are calculated from the instantaneous observations, but for different sampling periods, i.e. day (Q66-D), month (Q66-M) and season (Q66-S). The accuracy is defined as the bias between the median SEVIRI and MWR retrieved LWP values over the observation period, whereas the precision is given by the Q50 value of the deviations between SEVIRI and MWR retrieved LWP values.

9.3 Results ground-based comparison

9.3.1 Instantaneous LWP values

A statistical analysis of frequency distributions of LWP retrievals from MWR and SEVIRI is performed to evaluate the differences between Chilbolton and Palaiseau. Figure 9.1 presents the distributions of LWP retrieved from SEVIRI and MWR over the period May – August 2004 for both CloudNET sites. The LWP distributions from SEVIRI and MWR are log-normally distributed and have similar shapes. The lower tails of the distributions reveal differences that are mainly related to differences between the LWP retrieval algorithms. The LWP retrievals from MWR can become slightly negative due to small calibration drifts, whereas the LWP retrievals from SEVIRI are always positive. Clouds with low LWP values dominate the distribution of Palaiseau, while thicker clouds that could be associated with convection (LWP > 100 g m^{-2}) rarely occur. The distribution of Chilbolton exhibits a much wider range of LWP values. Although the majority of the clouds at Chilbolton have LWP values smaller than 30 g m^{-2}, a considerable fraction of clouds (about 10%) have LWP values larger than 100 g m^{-2}. At Chilbolton, SEVIRI observes a 20% higher frequency of clouds with LWP values between 0 and 30 g m^{-2} than the MWR. This difference reduces to an overestimation of about 5% when the negative values are taken into account. The MWR retrieves negative LWP values for about 15% of the observations. The right graph in Figure 9.1 shows that the SEVIRI overestimation is compensated by an underestimation of the frequency of thick clouds (LWP > 50 g m^{-2}). Note that sampling differences partly explain why SEVIRI observes high frequencies of clouds with low LWP values than the MWR. The variations in the LWP values from MWR often occur at sub-pixel level. Although the LWP values from MWR are averaged over a 20 minutes period, aiming to represent more or less the field of view of the SEVIRI, the MWR samples a substantially different portion of the cloud (~0.1x15 km^2) than SEVIRI (~4x7 km^2). For example, cloud fields that appear as homogeneous fields of thin clouds at the 4x7 km^2 resolution of SEVIRI may show up either as cloud free or as cloud filled along the 0.1x15 km^2 sample track of the MWR.

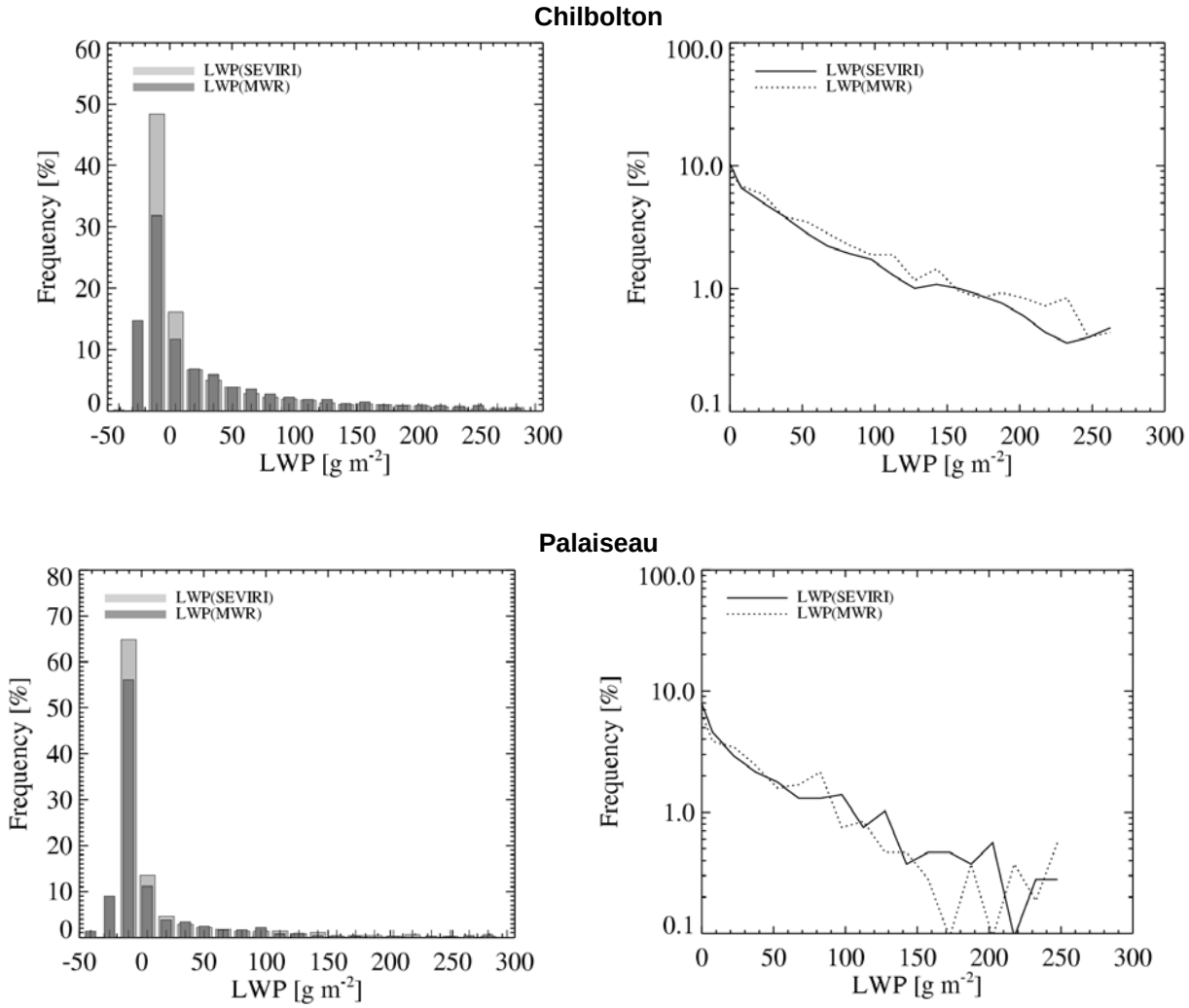

Figure 9.1 *Frequency distributions of SEVIRI and MWR retrieved LWP and their corresponding distributions plotted on a logarithmic scale for Chilbolton and Palaiseau over the period May – August 2004.*

Figure 9.2 shows that the frequency distributions of differences are non-Gaussian. This is seen best from the strongly peaked frequency at differences around zero and the rapid drop in the frequency of occurrence as the differences increase. The slightly negative skew suggests larger LWP values from MWR than from SEVIRI. At Chilbolton and Palaiseau, the Q66-S values of about 55 and 26 g m^{-2} are in the same order of magnitude as the mean LWP values from MWR of about 58 and 33 g m^{-2}, respectively. The Q95-S values are about six times larger than the Q66−S value, with 289 g m^{-2} for Chilbolton and 206 g m^{-2} for Palaiseau. This indicates that for a limited number of observations the differences between the LWP retrievals from SEVIRI and MWR are very large. Possible reasons for these large Q95-S values are the nature of cloud in-homogeneity, multi-layer clouds, and the decreasing accuracy of both ground-based and SEVIRI retrievals of LWP with increasing cloud optical thickness.

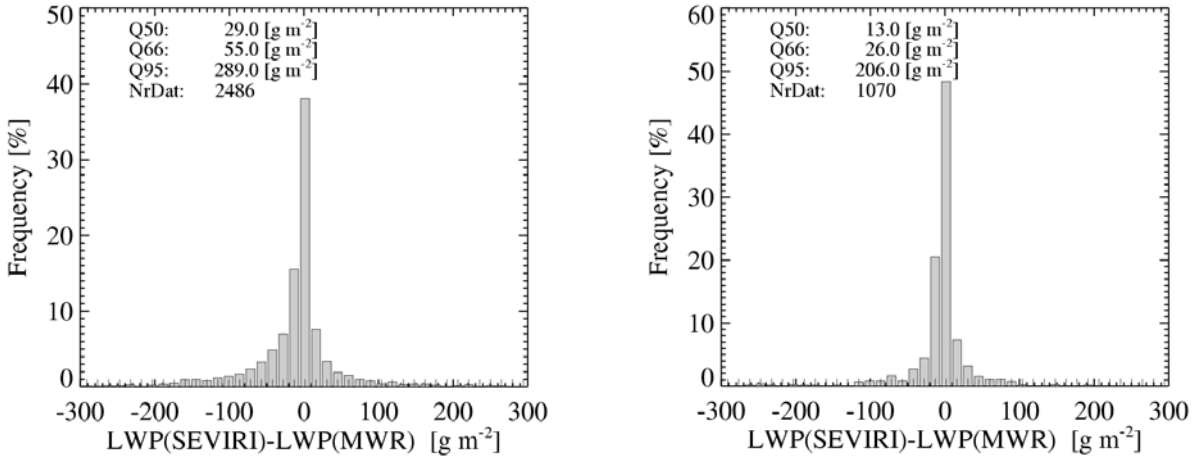

Figure 9.2 *Frequency distributions of differences between SEVIRI and MWR retrieved LWP for Chilbolton (left) and Palaiseau (right) over the period May – August 2004.*

Figure 9.3 presents the accuracies of SEVIRI retrieved LWP values as a function of the LWP values retrieved from MWR. These values are calculated for bins of 20 g m^{-2} in MWR retrieved LWP values. The number of coincident observations and the Q66-S values are also given. The Figure shows a substantial reduction in accuracy with increasing LWP values from MWR, with an underestimation of about 30 g m^{-2} at MWR retrieved LWP values of about 100 g m^{-2}. However, the majority of the observations are made at MWR retrieved LWP values smaller than 40 g m^{-2}, where the accuracies are better than 10 g m^{-2}. In general, the Q66-S values (error bars) are about equal to the MWR retrieved LWP values, both at Chilbolton and Palaiseau. If the Q66-S value represents twice the standard deviation, the relative precision of the instantaneous LWP retrievals from SEVIRI is about 50%.

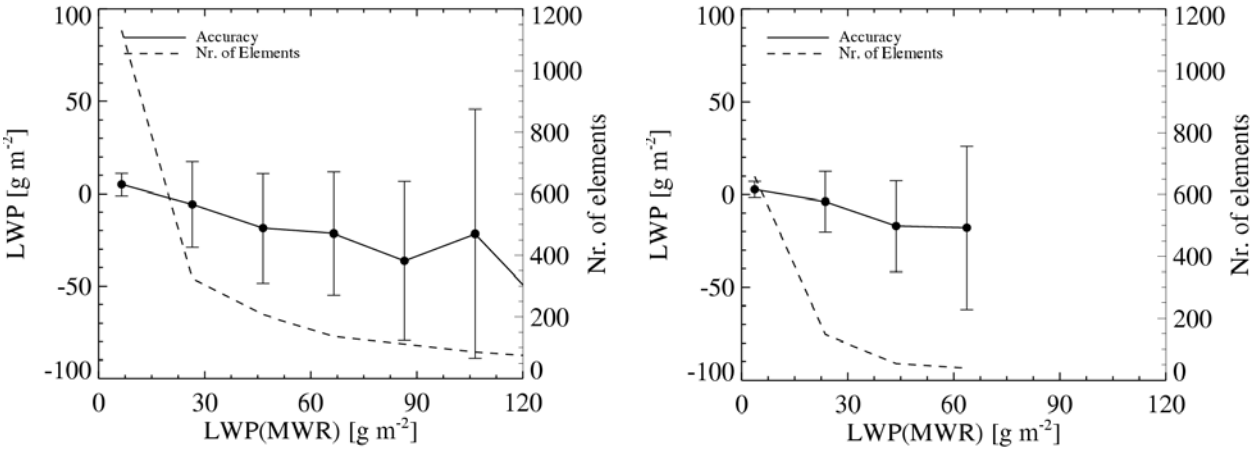

Figure 9.3 *The accuracies and number of observations of the instantaneous LWP retrievals from SEVIRI as function of the instantaneous LWP values from MWR for Chilbolton (left) and Palaiseau (right). The accuracies are calculated for bins of 20 g m^{-2} in LWP values from MWR over the period May – August 2004. The error bars give the Q66-S values for each bin.*

9.3.2 Daily LWP values

Comparing daily median LWP retrievals instead of instantaneous retrievals can reduce the effect of spatial mismatching. The unique characteristic of SEVIRI is that the high sampling frequency (15 minutes) combined with the spectral channels similar to AVHRR allows for the calculation of daily median LWP values. The daily median LWP values were calculated from SEVIRI and MWR

retrievals for days with at least six observations. Figure 9.4 presents the daily median LWP values from MWR and SEVIRI for 83 days at Chilbolton and 44 days at Palaiseau during the summer period. At both locations large variations in daily median LWP values are observed, ranging from 0 to 400 g m^{-2}. However, for about 90% of the days the daily median LWP values are below 100 g m^{-2}. In general, the agreement between the daily median LWP values from MWR and SEVIRI is very good, with a correlation of 0.94 at Chilbolton and 0.95 at Palaiseau. This is surprisingly high, considering the fact that the MWR and SEVIRI sample different portions of the cloud. With the exception of a few days at both sites, the differences between the daily median LWP retrievals from SEVIRI and MWR are smaller than 30 g m^{-2}. The Q66-D values (error bars), which indicate the variance of the differences between the instantaneous retrievals during the observation days, are smaller than 100 g m^{-2} for most days, but larger than the median LWP values. Both at Palaiseau and Chilbolton, the daily median LWP values from SEVIRI are retrieved with an almost perfect accuracy (better than 3 g m^{-2}). The precision of the daily median LWP values are better than 15 g m^{-2}, which corresponds to relative precisions better than 30%. This indicates that the precision of the daily median LWP values from SEVIRI is significantly better than the precision of the instantaneous retrievals.

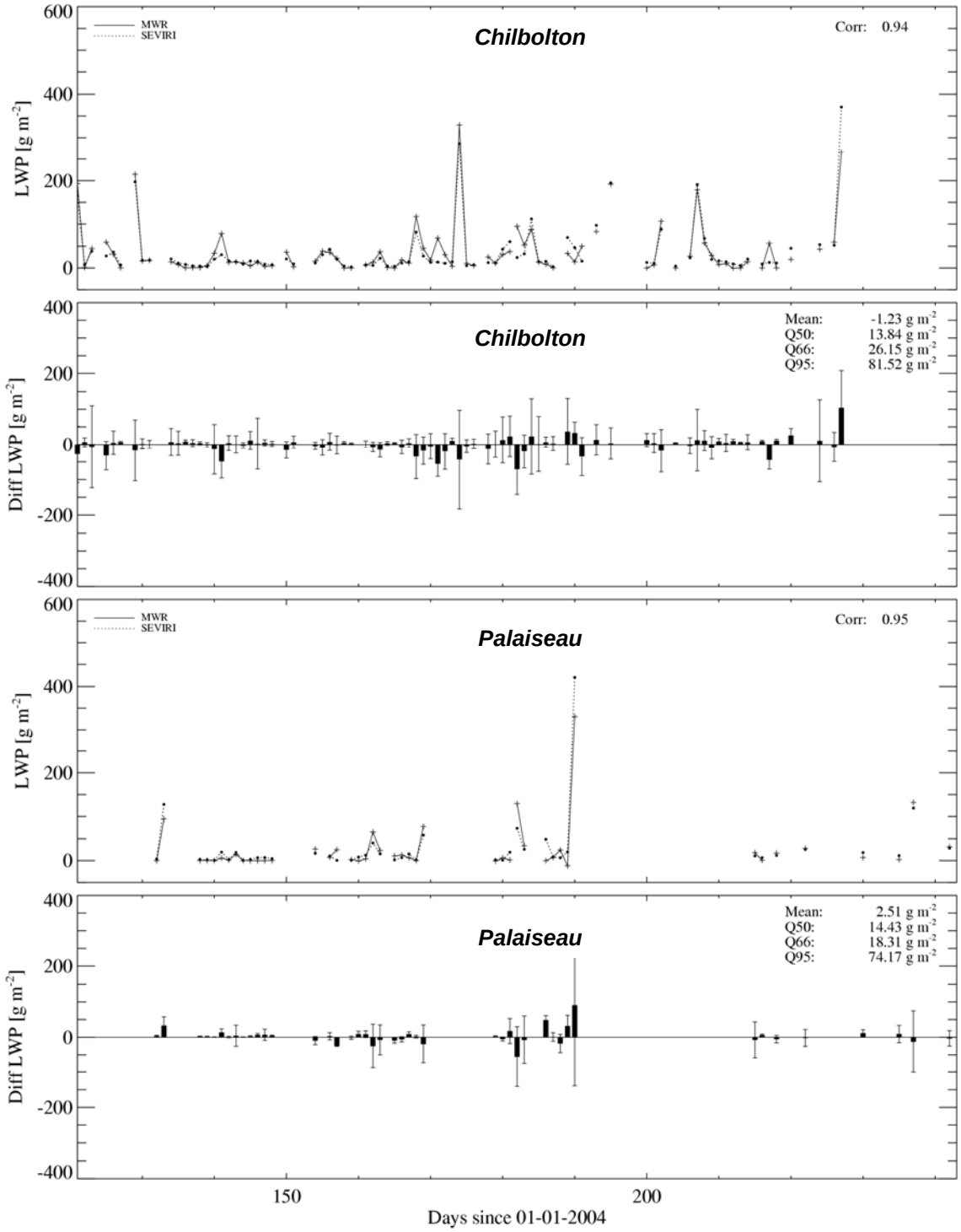

Figure 9.4 *Time series of daily median LWP values from SEVIRI and MWR, and their corresponding difference in LWP for Chilbolton and Palaiseau over the period May-August 2004. The error bars indicate the Q66-D values.*

9.3.3 Monthly LWP values

The high number of observations per month (> 400) allows for the calculation of statistically significant values of the monthly median LWP. Figure 9.5 presents the monthly median LWP retrievals from MWR and SEVIRI for Palaiseau over the 4 summer months. The values are directly calculated from the instantaneous retrievals that have been presented in Figure 9.1. The dominance of thin clouds during the summer months at Palaiseau is reflected in the magnitude of monthly median LWP values from MWR, which vary between 1 and 20 g m^{-2}. The difference between the LWP retrievals from SEVIRI and MWR is slightly positive for Palaiseau.

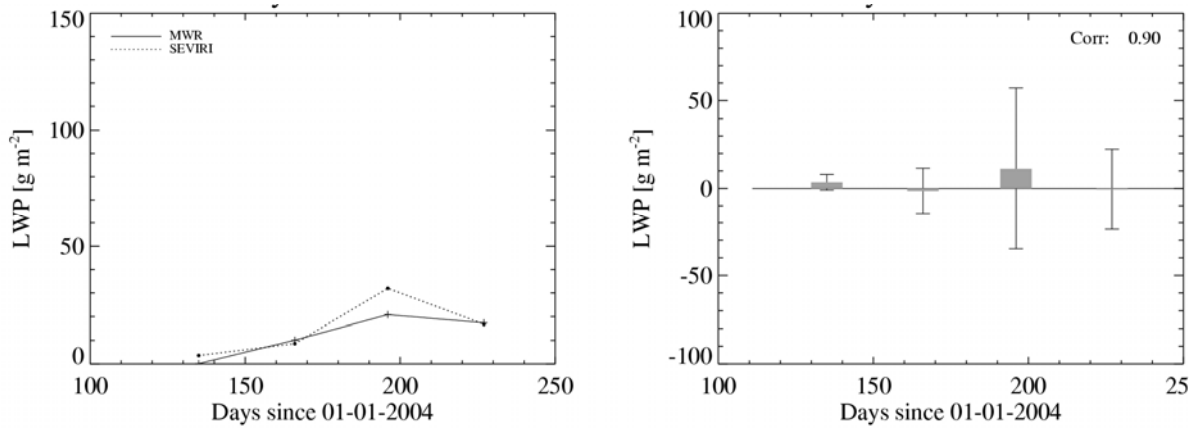

Figure 9.5 *Time series of monthly median LWP from SEVIRI and MWR and their difference for Palaiseau. The error bars indicate the Q66-M values.*

Figure 9.6 is similar to Figure 9.5, but then presents the results for the monthly median LWP values over the observation year. The monthly median LWP values from SEVIRI are in the same order of magnitude as the MWR values, and vary between 10 and 60 g m^{-2}. In general SEVIRI slightly underestimates the LWP values from MWR, with as exceptions November 2004 and January 2005. The accuracies during the summer months (~5 g m^{-2}) are significantly better than the during the winter months (~25 g m^{-2}). Besides the lower accuracies during winter, the precision, as indicated by the error bars, also reveals a strong annual cycle. During the summer months the precisions are better than 20 g m^{-2}, whereas during the winter months (September - March) these values are larger than 50 g m^{-2}.

Contrary to the results presented for the daily median LWP values, the results of the comparison of monthly median LWP values are somewhat different for Chilbolton and Palaiseau. These differences could be related to the differences between the MWRs at the CloudNET sites. Löhnert and Crewell (2003) showed that differences of 5 to 10 g m^{-2} between different MWRs are common. However, the meteorological conditions at Palaiseau and Chilbolton differ too much to attribute the observed differences to instrumental differences. To quantify the accuracies of the MWRs at the CloudNET sites would require either a longer dataset, or even better, a microwave intercomparison study at one of the measurement sites.

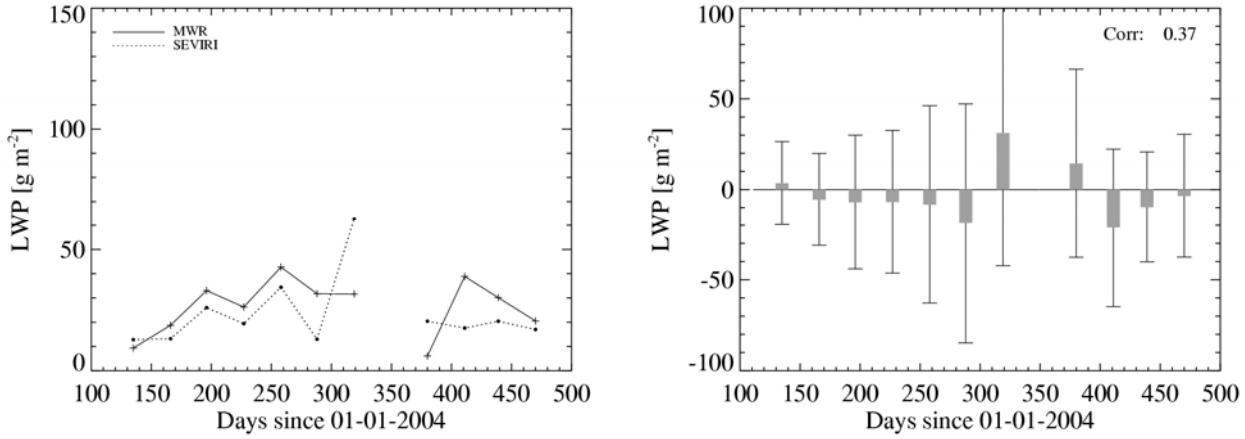

Figure 9.6 *Time series of monthly median LWP from SEVIRI and MWR and their difference for Chilbolton over the period May 2004 – April 2005. The error bars in the difference plots indicate the Q66-M values.*

9.3.4 Diurnal variations of LWP

Figure 9.7 shows the diurnal variations in median LWP values from SEVIRI and MWR as function of the fraction of the day for the CloudNET sites over the summer period. The fraction of the day is the normalized period between sunrise (fraction = 0) and sunset (fraction = 1). The median LWP values from MWR exhibit a clear diurnal trend. At both CloudNET sites, the LWP values of either early morning (fraction < 0.2) or late afternoon (fraction > 0.8) observations are about six times smaller than the values at local solar noon (fraction = 0.5). The LWP values from MWR exhibits a sharp increase till the fraction is about 0.4, which corresponds during summer to 10 hr local solar time. Note that the thickest clouds are observed around local solar noon, when the continental boundary layer is thickest and convective activity highest. There is a slight asymmetry between the LWP values before and after local solar noon. The afternoon LWP values are somewhat higher than the morning values, which is probably the result of increased convection from morning to afternoon. Throughout the day there are significantly thinner clouds at Palaiseau than at Chilbolton, which can be seen from the median LWP values from MWR that are about two times lower at Palaiseau than at Chilbolton.

In general, the median LWP values from SEVIRI exhibit similar diurnal variations as the MWR values. However, the amplitude of the diurnal variations in LWP is smaller from SEVIRI than from MWR. During early morning of late afternoon SEVIRI always observes higher median LWP values than the MWR. It is suggested that cloud in-homogeneities may be responsible for the observed differences at these observation times. This is consistent with the results of Loeb and Coakley (1998) who found that cloud property retrievals from one-dimensional schemes, such as CPP, systematically increase at the solar zenith angles (□0) that are observed during early morning or late afternoon (□0 > 60°). For most observations at Palaiseau, the median LWP values from SEVIRI are higher than the corresponding MWR values, with a maximum difference of 5 g m–2. This does not agree with the results of Chilbolton, where SEVIRI overestimates LWP during early morning and late afternoon, while LWP is underestimated around local solar noon.

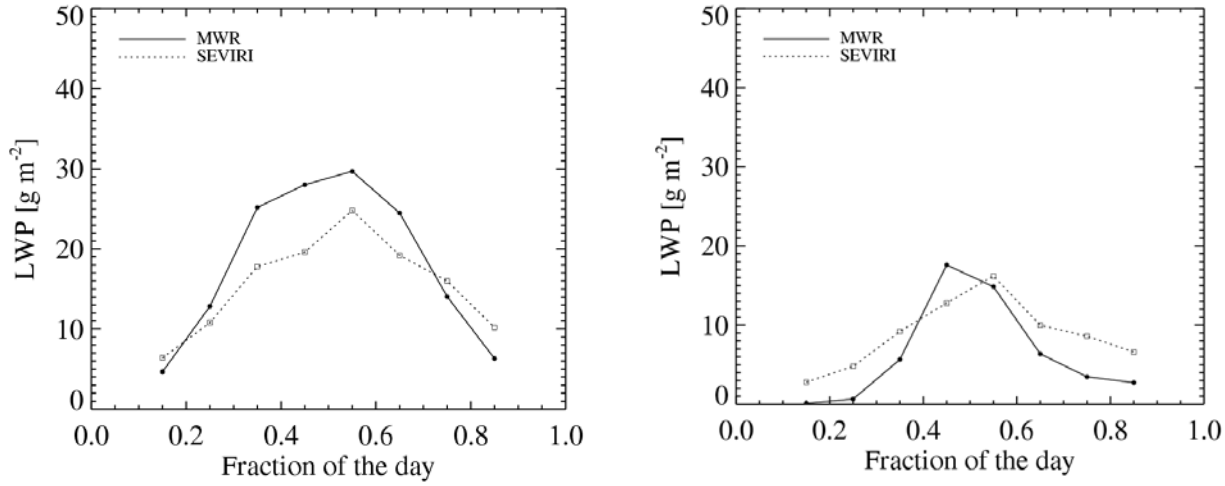

Figure 9.7 *The median LWP retrieved from MWR and SEVIRI as function of the fraction of the day for Chilbolton (left) and Palaiseau (right) during the period May–August 2004. The fraction of the day is the normalized period between sunrise (fr. = 0) and sunset (fr. = 1).*

9.4 Results comparison SEVIRI and MODIS

The results of the comparison between SEVIRI and MODIS retrieved LWP values are presented in Figure 9.8. The comparison results are given for the selected regions over Tropical Land, Tropical Ocean and Subtropical Ocean using the coinciding overpasses during July 2004. The left panels in the Figure show the frequency distributions, whereas the right panels show the distributions of the absolute difference between SEVIRI and MODIS retrieved LWP values.

The frequency distributions of LWP from both satellites are similar. For all regions SEVIRI observers significantly higher frequencies of clouds with LWP values smaller than 50 $g.m^{-2}$ than MODIS. Over Tropical Land this difference is about 10%, whereas differences of about 20% are observed over the Ocean regions. On the other hand, for clouds with LWP values larger than 100 $g.m^{-2}$ MODIS observers about 10 -20 % higher frequencies than SEVIRI. It is remarkable that this effect is largest over the Ocean regions. This is also confirmed by the median LWP values from MODIS, which are about 10 $g.m^{-2}$ larger over land and 20 g m^{-2} over Ocean from MODIS than those from SEVIRI. Considering collocation errors and differences in spatial resolution and viewing conditions, the agreement between the SEVIRI and MODIS retrievals of COT can be regarded very well.

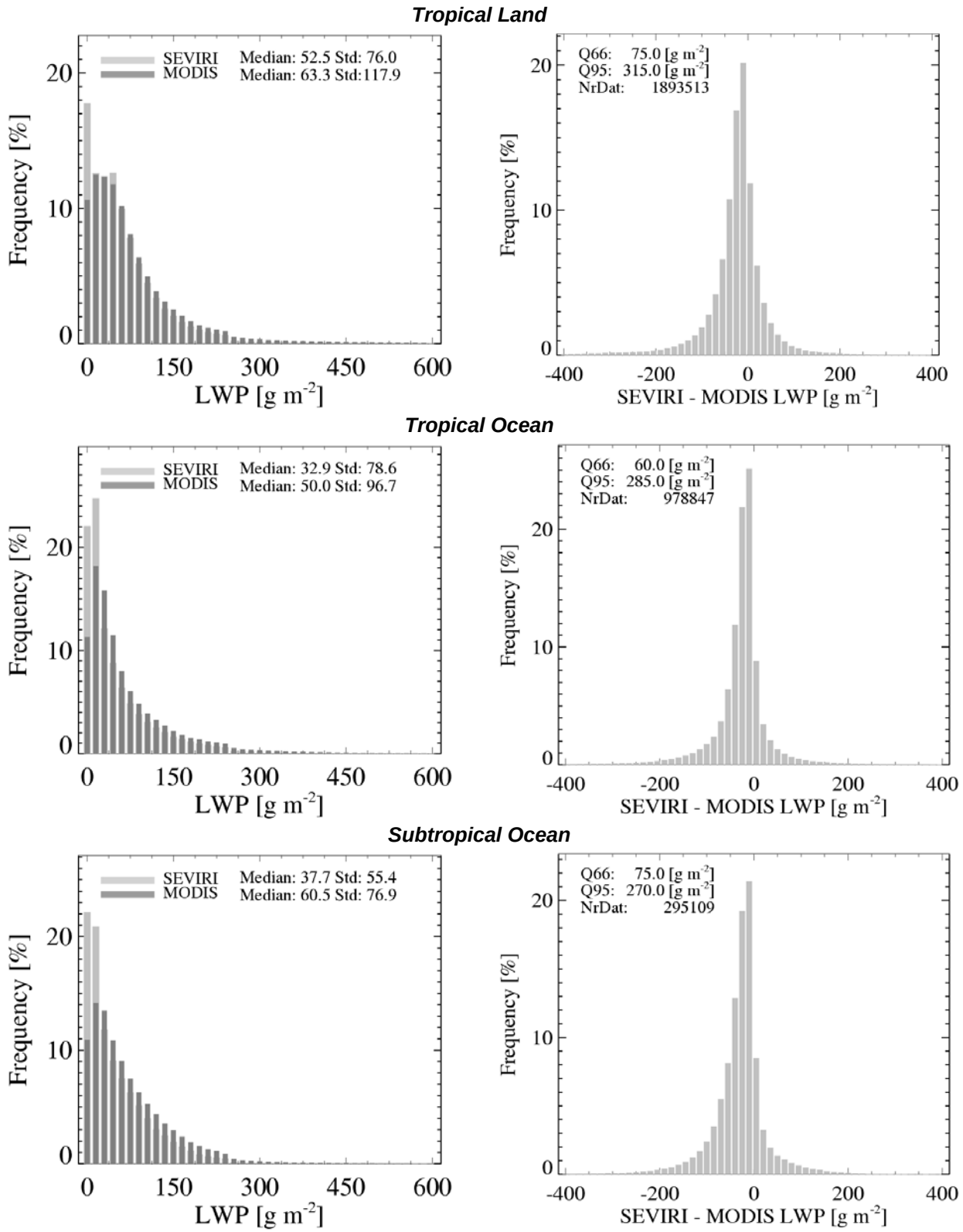

Figure 9.8 *Frequency distributions of SEVIRI and MODIS LWP (left panels) and frequencies of differences (right panels) for tropical land (upper panels), tropical ocean (middle panels), and subtropical ocean (lower panels).*

The differences between the frequency distributions of LWP from MODIS and SEVIRI are much larger than those between the COT values. The differences are smaller than what was observed in the SIVVRR_V3 report, where part of the differences was related to a bug in the MODIS collection-4 calculation of cloud water path. In collection-4 the LWP is incorrectly calculated as 3/4xCOTxReff, whereas the correct equation is 2/3xCOTxReff. Where, COT is the optical thickness and Reff the effective radius. This bug was solved in MODIS collection-5. Since the COT retrievals from SEVIRI and MODIS are similar over the observation areas and periods, the differences between the SEVIRI and MODIS retrieved LWP values are due to differences in retrieved Reff values. There are three reasons for MODIS to retrieve larger Reff values than SEVIRI. First, the MODIS particle size retrievals tend to be slightly too high (personal communication with Steve Platnick). Second, MODIS uses the 2.2 μm instead of the 1.6 μm channel for the retrieval of effective radius. The observed effective radii in this channel are more weighted towards the cloud top than the effective radii observed at 1.6 μm. Platnick, 2000 showed that for clouds with effective radii increasing towards the top and COT > 10 that the effective radii observed at 2.2. μm are about 10% larger that those observed at 1.6 μm. Finally, for thin clouds (COT < 8) the CPP algorithm uses an assumed effective radius of 8 μm (see section 3.4), whereas the MODIS algorithm retrieves an effective radius for these clouds.

9.5 Summary

This chapter presents the validation of SEVIRI retrieved LWP values using MWR retrieved LWP values from the CloudNET sites in Palaiseau and Chilbolton. The ability of SEVIRI to make accurate retrievals of LWP over Northern Europe has been examined. A high agreement is found during the summer months between instantaneous LWP retrievals from MWR and SEVIRI for both Palaiseau and Chilbolton. The accuracy of these retrievals is about 5 g m-2, whereas the precision is about 40 g m^{-2}. The added value of the 15-minute sampling frequency of METEOSAT-8 is especially evident in the validation of the daily and monthly median LWP retrievals from SEVIRI. These retrievals agree significantly better with the MWR retrieved LWP values than the instantaneous ones. This is indicated by the large improvement of the precision, which is better than 15 g m^{-2} for the daily median LWP retrievals from SEVIRI. The analysis of one year of daily median LWP retrievals for Chilbolton reveals a clear annual cycle of accuracy, with much lower accuracies during winter than during summer.

9.6 Discussion

The precisions of instantaneous validation presented in this report are well within the range of expected precisions, and similar to the precisions of the LWP values retrieved from MWR of about 30 g m^{-2}. However, the precisions significantly improve when, instead of instantaneous values, the daily median LWP values are compared. This improvement suggests that part of the observed differences is related to validation uncertainties. Roebeling et al. (2006b) examined the uncertainties in validation studies due to cloud in-homogeneities and the overlap between MWR and SEVIRI retrieved LWP values. The latter is determined by uncertainties in co-location, parallax, position of the ground station and differences due to sampling of different portions of the cloud. They used LWP fields from MODIS to simulate SEVIRI and MWR retrieved LWP values, and calculate the Q66 values for above described uncertainties. For marine stratocumulus clouds they found that the validation causes errors similar to or larger than those of the SEVIRI retrieval process, with uncertainties due to co-location and parallax of about 50 g m^{-2} and uncertainties due to sampling different portions of the clouds of about 20 g m^{-2}. Part of these differences may be alleviated through improving the sampling strategy. In this report, a simple sampling strategy is used, in which the LWP retrievals from SEVIRI over the ground station are compared to 20 minute mean LWP values from MWR. Therefore a substantial part of the Q66 values could be due to co-location mismatch. Improvements in the validation may be obtained by determining the optimum ground track length that corresponds with the track that overlaps best with the SEVIRI

pixel. Thus, for an optimal correspondence ground-based observations need to be averaged over different periods depending on the wind speed and direction at cloud altitude.

The validation of one year of LWP retrievals from SEVIRI exhibited large differences in accuracy between summer and winter. We suggest two possible reasons for the decreased accuracy of LWP retrievals from SEVIRI during the winter months. First, the LWP retrievals from SEVIRI are much more sensitive to errors at the low solar zenith angles and backward scattering geometries that prevail during the winter months over Northern Europe. Second, cloud in-homogeneities influence the reflectances most at these viewing geometries and may cause large errors in one-dimensional retrievals of LWP, such as the CPP algorithm.

During winter, the solar zenith angles are high (θ_0> 60°) and the scattering angles are also often in backward scattering directions. Although similar low solar zenith angles are observed during summer in the morning or late afternoon, these observations do not coincide with scattering angles close to the backward scattering peak. Loeb et al. (1998) found that the relative difference between three-dimensional and one-dimensional cloud reflectances can be large due to subpixel variations in cloud-top height (i.e., cloud bumps). Depending on the structure of the cloud field and its optical thickness, the three-to-one-dimensional difference may be as large as 10% in backward scattering directions. These differences are largest at viewing zenith angles > 60°, where it may lead to a significant overestimation of optical thickness. Radiative transfer simulations with DAK show that one-to-three-dimensional differences of 10% may lead to very large errors in COT retrievals at low solar zenith angles for thick cloud (COT > 30). Because of the non-linear relationship between the simulated reflectances and COT, an increase of the reflectance of 5% at azimuth difference angle (ϕ 160 and solar zenith angle (θ_0 70° results in a COT increase from about 50 to 250 (about 500%). The DAK simulations show that this sensitivity is much lower at low solar zenith angles, where the reflectances saturate at larger COT values. In addition, the one-to-three-dimensional differences are smaller at low solar zenith angles. Thus, it is likely that one-to-three-dimensional differences at high solar zenith angles in the backward scattering direction, the viewing geometries that correspond to SEVIRI observations during the winter season, leads to higher LWP values from SEVIRI that have lower accuracy.

10 Discussion and conclusions

10.1 Macrophysical cloud products (CFC, CTY and CTx)

Chapter 4 describes the validation of the fractional cloud cover product CFC based on comparisons with surface observations and observations from two other satellite sensors (MODIS and CALIOP). In general, the results of this study and the initial validation study (see AD 9) agree quite well.
Taking into account all the differences between ground-based measurements and satellite observations (viewing geometry, subjective man observations, rules for ground-based observations) the overall performance of instantaneous CFC results is rather good which is confirmed by a high Kuiper skill score of around 0.82 and a bias of only 0.044 or 0.008 if tuned for partially cloudy satellite footprints. Daily and monthly mean values and the monthly mean diurnal cycle were further calculated for both ground-based and space-based records and we could show that the given target accuracy (bias) of CFC products is reached for all product types.
However, the detailed results clearly show that satellite measurements overestimate the cloud coverage over sea surfaces while some underestimation is found over land. We found also a negative bias of CFC in the tropical belt which is compensated by a positive bias towards the edges of the earth disk. A similar behaviour was also seen when comparing to corresponding result from NOAA AVHRR data. Best performance of all products is found over Northern Midlatitudes, mainly over European land surfaces.
Comparison of SEVIRI CFC with MODIS and CALIOP CFC gave in principle similar results: However, there is some evidence that the matching of SEVIRI and CALIOP observations is problematic which causes somewhat lower skill scores in comparison to the MODIS-SEVIRI case.
The MODIS-SEVIRI intercomparison provides some evidence that SEVIRI often misses optically thin cirrus clouds and, in particular, misclassifies tropical semi-transparent cirrus clouds as being fractional instead of fully covering the cloudy pixel.
We believe that we can now answer the open question about the validity of using a bias correction factor outside the Baseline area. Our analysis confirmed this since the bias of the full disk results is minimised if the tuning parameter is set to 75% (0.75). The comparison with corresponding results for the NOAA AVHRR CFC product (also testing different bias correction factors) largely supported these findings. However, a final decision on if this correction should be implemented for the operational products or not must also take into account the fact that for full consistency also other cloud products could be subject of a bias correction. This concerns in particular the COT product which today is derived assuming fully cloudy pixels. If correcting CFC but not COT new problems arise for external users with a desire to use both parameters for cloudy radiative transfer applications. If keeping both parameters uncorrected one could still claim that the usage in radiative transfer applications would be reasonably justified (overestimation of cloud coverage will at least partly compensate for underestimated cloud optical thickness).

Chapter 5 describes the validation of the MSG/SEVIRI CTY product using detailed information about cloud layer occurrence and cloud layer altitudes derived from cloud radar measurements. In general, the results confirm previous SIVVRR V2 and ORR-V2 validation results (AD 7, AD 9). Especially low-level and high-level cloud assignments appear to work well but problems are evident for mid-level clouds. However, this problem is largely explained by limitations of the validation method rather than by wrong MSG CTY assignments. The limitation of having cloud radar measurements only for two sites and restricted to four months made the estimation of the accuracy for monthly means uncertain and the estimation of the corresponding precision not feasible. Furthermore, radar measurements were only made during daytime which precluded an investigation of the diurnal cycle of cloud types.

It is important to mention that the cloud type product will be subject of a major revision during the upcoming CDOP phase, mainly because of the fact that we are currently not utilizing the results provided by the complete set of cloud products (e.g. corrected cloud top information and the complete set of cloud physical products).

Chapter 6 describes the validation of the MSG/SEVIRI cloud top product CTx (with its three realisations CTH, CTT and CTP) using cloud radar observations (similar to the CTY validation) and two other satellite-based cloud top retrievals from the MODIS and CALIOP sensors.
The validation of cloud-top height from instantaneous MSG/SEVIRI data using cloud radar observations showed that SEVIRI observations seem to overestimate the cloud-top height, especially if semitransparent clouds are present. However, there are remaining uncertainties caused by problems to ensure a reasonable comparison between the two observation sources. Similarly to the case of CTY, we are here not able to confidently estimate monthly accuracies and precisions as well as the quality of the diurnal cycle.
In general, we confirm the main results of the previous validation study indicating that target accuracies are generally reached (although exceptions do occur for some individual months). While scenarios with high-level and low-level clouds agree rather well (except for the noticed overestimation of thin cirrus) problems increase for mid-level clouds which are difficult to retrieve from radar data (discussed also in the context of cloudy type validation).
Comparisons of MODIS and SEVIRI cloud-top pressure (CTP) results give an overall low bias (< 10 hPa) while the standard deviation of about 170 hPa is rather high. Depending on the height level of typical clouds this translates into a statistical variability of a cloud layer of more than 3 km. SEVIRI CTP is often higher for low clouds between 700 hPa and 1000 hPa but the height of high clouds is also often overestimated (SEVIRI CTP lower) if compared to MODIS results.
SEVIRI CTP is systematically higher over water surfaces in the Midlatitudes (Northern and Southern hemisphere) and there is a clear trend towards a negative bias (SEVIRI CTP lower) for increasing satellite observation zenith angles and for high southern latitudes. One possible explanation could be that thin cirrus clouds above water clouds are not properly detected and corrected for by the SEVIRI algorithm but also other influencing factors are likely to be important.
CALIOP and SEVIRI cloud-top height results were compared under the same premises. We found again a low overall bias of the cloud-top height of below 50 m that goes along with a high standard deviation of more than 2900 m. Partly in contrast to the CTP results of the SEVIRI-MODIS intercomparison, we observe an overall overestimation of the SEVIRI cloud-top height relative to CALIOP.

10.2 Products on Cloud Physical Properties (CPH, COT and CWP)

Chapter 7 describes the validation of the SEVIRI-derived CPH retrievals with both ground-based observations over the CloudNET sites and a comparison with MODIS retrievals. For the CloudNET sites the CPH retrieval from SEVIRI agree fairly well with the ground-based target classification data during the summer months. This is indicated by monthly Kuiper Skill Scores in the range of 0.7 to 0.8 (misclassification of 20-25%) for the ground-based comparison. During winter months, the dependence of the CM-SAF CPH product on visible and near-infrared reflectance reduces the skill significantly due to the decrease of the signal-to-noise ratio with decreasing sun elevations, leading to a drop in skill score to 0.3 (misclassification about 50%). Although the Kuipers Skill Scores are fair, the monthly CPH statistics show that SEVIRI underestimates the percentage of ice clouds, with differences of 30-40%, as compared to the ground-based observations. It is discussed that a large part of this bias is attributed to the difference in detection of ice clouds between SEVIRI and the ground-based instruments used.
The comparison of SEVIRI and MODIS cloud phase retrievals shows a much better agreement than the comparison to ground based observations. Over Tropical and Subtropical Ocean surfaces both instruments retrieve very similar percentages of ice and water clouds. Over the Tropi-

cal Land surface SEVIRI tends to underestimate the percentage of ice clouds by 10% as compared to MODIS. This bias is significantly smaller than the bias between SEVIRI and ground-based cloud phase retrievals. The MODIS cloud phase tests, which are based on infra-red spectral channels, may be used to further improve the SEVIRI cloud phase product during the winter months and over land surfaces. Since SEVIRI and MODIS have many spectral channels in common, most MODIS tests for the cloud phase product can serve as base line for additional infra-red cloud phase tests, and will be adapted for use on SEVIRI in future updates to the CM-SAF CPH algorithm.

Chapters 8 and 9 describe detailed validation studies of SEVIRI-derived COT and CWP values with ground-based pyranometer and MWR measurements, as well as a comparison with MODIS retrieval results. Calculating the atmospheric transmission from retrieved COT based on the SICCS retrieval, for a cloudy atmosphere the atmospheric transmission is well-reproduced, with an accuracy better than 3 percent and a precision for hourly values of about 9 percent throughout the year. However, the validation indicates too high cloud optical thicknesses and too low atmospheric transmission for thin clouds. The comparison of ground-based MWR and SEVIRI induced LWP shows that both the instantaneous and daily median LWP values from SEVIRI have a high accuracy of about 5 g m^{-2}. The precision of the instantaneous LWP retrievals from SEVIRI is better than 30 g m^{-2}. Daily median LWP values from MWR and SEVIRI agree with precisions better than 15 g m^2 and with almost no bias. Good agreement is also found between SEVIRI and MODIS COT and CWP path retrievals. The COT values derived by MODIS and SEVIRI show very similar frequency distributions, and their corresponding differences peak at a value close to zero. The CWP retrievals from MODIS are 20-50% higher than the SEVIRI retrieved CWP values for optically thin clouds. These differences for these clouds are explained by larger effective radii found from MODIS than from SEVIRI.

Overall, the three CM-SAF products CPH, COT and CWP show a high level of consistency with the satellite-based MODIS retrieval. While this should be expected, as they are based on similar input data and retrieval methods, it also provides confidence in the quality of implementation of the algorithms. This finding is not trivial, because larger differences were found between different SEVIRI-based retrievals that where compared for the EUMETSAT MSG Cloud Workshop in Norrköping, May 2006. The validation with ground-based reference data is much more problematic due to different scales of measurements, as well as the different measurement principles and the resulting differences in sensitivity. Nevertheless, the agreement found is rather encouraging, in particular for the COT and CLWP product.

The observed high bias of COT versus ground-based pyranometer measurements is likely caused by a mismatch of observed and expected clear-sky TOA reflectance. This issue is currently being investigated, together with the discrepancy of MODIS- and SEVIRI-retrieved effective radii. Possible explanations include bi-directional variations of surface reflectance, calibration differences between SEVIRI and MODIS leading to problems in applying the MODIS surface reflectance, or an unrealistic modelling of the radiative transfer in clear-sky situations. If these investigations reveal shortcomings in the products, this will be fixed by a future update to the algorithms.

The validation also shows that the instantaneous retrieval results are affected by a fairly large random scatter. While part of the imprecision might be attributable to the choices made in the implementation, the majority is inherent in the retrieval method of Nakajima and King (1990), which is very sensitive to input errors due to the underlying non-linear retrieval function, in combination with the highly variable nature of clouds. The magnitude of this scatter indicates that the overall accuracy is also highly dependent on systematic uncertainties, such as radiometric calibration of the SEVIRI sensor. Biases can reach up to 30 percent for the COT for realistic estimates of the calibration uncertainty (Roebeling, 2005). Future efforts towards the CM-SAF COT and CWP products will therefore be directed at a more precise quantification of the uncertainties

of these products. This will allow better judgement of the significance of differences observed between different datasets by users, or the selection of subsets according to accuracy requirements.

10.3 Achieved accuracies and precisions

The results of the ORR V3 validation activities for the CM-SAF cloud products are here finally summarised in Table 10.1. Concerning the macrophysical products (CFC, CTY, CTx) we omitted results from the SEVIRI-MODIS and the SEVIRI-CALIOP intercomparison studies. Note, that we have not been able to calculate all required averaged error numbers for CTY and CTX which were only available from two ground stations and a limited number of months. Similar restrictions exist also for some of the cloud physical products.

Table 10.1 *Validation results for the CM-SAF cloud products for the ORR V3 review. Results are given for the evaluation of individual (instantaneous), daily and monthly estimations (where applicable). Only validation results against ground-based measurements were considered.*

Product (unit)	**Acronym**	**Resolution/accuracy/precision**							
		Temporal							
		Instantaneous		**Daily**		**Monthly**		**MMDC (Monthly Mean Diurnal Cycle)**	
		accur.	prec.	accur.	prec.	accur.	prec.	accur.	prec.
Fractional cloud cover (%)	CFC[1]	4.4		4.9	18.7	4.9	11.7	4.6	14.7
Cloud type (low) (%)	CTY[2]	-	-	<15	<35	< 15	-	-	-
Cloud type (mid) (%)	CTY[2]	-	-	<5	<35	< 5	-	-	-
Cloud type (high) (%)	CTY[2]	-	-	<10	<40	< 10	-	-	-
Cloud top temperature, height, and pressure (K, m, hPa)	CTH, (CTT), (CTP)[3]			<15	<35	< 15	-	-	-
Cloud phase (%)	CPH	30	-	-	-	30	-	-	-
Cloud optical thickness (dimensionless) (+ transmissivity in %)	COT[4] (T)	20 (5)	40 (10)	10 (5)	30 (7)	10 (3)	20 (5)	-	-
Cloud water path (gm^{-2})	CWP	5	40	5	15	5	-	-	-

[1] *Validation results derived without bias correction factor*

[2] *Cloud type observations compared to daytime cloud radar data (i.e., diurnal cycle not available)*

[3] *Results given only for CTH and in relative terms (% of mean CTH). No diurnal cycle available (as for cloud type).*

[4] *Results given in relative terms for COT (% of mean COT) and in absolute terms for T*

References

Anderson, G. P., S. A. Clough, F. X. Kneizys, J. H. Cheywynd, and E. P Shettle, 1986: AFGL Atmospheric Constituent Profiles (0-120km). Tech. Rep. AFGL-TR-86-0110, 43 pp.

Baum, B.A., P.F. Soulen, K.I. Strabala, M.D. King, S.A. Ackerman, W.P. Menzel and P. Yang, 2000: Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS: 2. Cloud thermodynamic phase, *J. Geophys. Res.,* **105***,* 11781-11792.

Boers R., A. van Lammeren, and A. Feijt, 2000: Accuracy of cloud optical depth retrievals from ground based pyranometers. *J. Atmos. Ocean. Tech.*, **17**, 916-927.

Cahalan, R., W. Ridgway, W. Wiscombe, T. Bell, and J. Snider, 1994: The albedo of fractal stratocumulus clouds. J. Atmos. Sci. 51(16), 2434–2455.

Crewell, S., and U. Löhnert, 2003: Accuracy of cloud liquid water path from ground-based microwave radiometry. Part II. Sensor accuracy and synergy, *Radio Sci.*, **38**, (3), 8042, 7–10.

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

Deneke, 2002: Influence of clouds on the solar radiation budget, *KNMI Scientific Report, WR-2002-09*.

Deneke, H., A. Feijt, A. van Lammeren and C. Simmer 2005: Validation of a physical retrieval scheme of solar surface irradiances from narrowband satellite radiances, *J. Appl. Meteor.,* **44***,* 1453-1466.

Derrien M. and H. LeGléau, 2005: MSG/SEVIRI cloud mask and type from SAFNWC, *Int. J. Rem. Sens.*, **26**, 470–4732.

Donovan, D.P., H. Bloemink, and A.C.A.P. van Lammeren, 1998: Analysis of ERM synergy by use of CLARA observations, *Final Report, ESA Contract No. 12953/98/NL/6D*.

Feijt, A.J., D. Jolivet, R. Koelemeijer and H. Deneke, 2004: Recent improvements to LWP retrievals from AVHRR, *Atmos. Res.,* **72**, 3-15.

Gaussiat, N., R. J. Hogan and A. J. Illingworth, 2006: Accurate liquid water path retrieval from low-cost microwave radiometers using additional information from LIDAR and operational forecast models, *J. Atmos. Oceanic Technol*, (submitted August 2006)

Hansen, J. E. and J. W. Hovenier, 1974: Interpretation of the polarization of Venus, *J. Atmos. Sci.*, **31**, 1137-1160.

Hanssen A.J. and W.J. Kuipers, 1965: On the relationship between the frequency of rain and various meteorological parameters. *Koninklijk Nederlands Meteorologist Institua Meded, KNMI*. Verhand, 81-2-15.

Hess, M., R. Koelemeijer, and P. Stammes, 1998: Scattering Matrices of Imperfect Hexagonal Ice Crystals, *J. Quant. Spectrosc. Radiat. Transfer*, **60**, 301-308.

Illingworth and Coauthors, 2006: Continuous evaluation of cloud profiles in seven operational models using ground-based observations, *Bull. Am. Meteor. Soc.*, accepted for publication.

Jolivet, D. and A. Feijt, 2005: Quantification of the accuracy of LWP fields derived from NOAA-16 advanced very high resolution radiometer over three ground stations using microwave radiometers, *J. Geophys. Res.,* **110**, D11204, doi:10.1029/2004JD005205.

Jolivet D., A.J. Feijt and P. Watts, 2000: Requirements for synergetic use of the ERM imager, *ESA Contract RFQ/3-9439/98/NL/GD*.

King, M.D., 1987: Determination of the scaled optical thickness from reflected solar radiation measurements,, *J. Atm. Sci.,* **44***,* 1734-1751.

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

Leontyeva, E., and K. Stamnes, 1994: Estimation of cloud optical thickness from ground-based measurements of incoming solar radiation in the Arctic, *J. Clim.,* **7***, 566-578*.

Li, Z., M. Cribb, F.L. Chang, A. Trishchenko, and Y. Luo, 2005: Natural variability and sampling errors in solar radiation measurements for model validation over the Atmoshperic Radiation Measurement Soutern Great Plains region, *J. Geophys. Res.,* **110***, D*D15S19, doi:10.1029/2004JD005028.

Li Z., K. Masuda, and T Takashima, 1993: Estimation of SW flux absorbed at the surface from TOA reflected flux, *J. Clim.,* **6***, 317-330*.

Loeb, N. G., T. Varnai and D. M. Winker, 1998: Influence of subpixel-scale cloud top structure on reflectances from overcast stratiform cloud layers, *J. Atmos. Sci.*, **55**, 2960-2973.

Löhnert, U., and S. Crewell, 2003: Accuracy of cloud liquid water path from ground-based microwave radiometry. Part I. Dependency on Cloud model statistics, *Radio Science*, **38**, (3), 8041, 6–11.

Majewski D., D. Liermann, P. Prohl, B. Ritter, M. Buchhold, T. Hanisch, G. Paul, W. Wergen, and J. Baumgardner, 2002: The operational global icosahedral-hexagonal grid point model GME: Description and high resolution tests, *Mon. Wea. Rev*., **130**, 319-338.

Marchand R., T. Ackerman, E.R. Westwater, S.A. Clough, K. Cady-Pereira, and J.C. Liljegren, 2003: An assessment of microwave absorption models and retrievals of cloud liquid water using clear-sky data, *J. Geophys. Res.*, **108** (D24), 4773, doi:10.1029/2003JD003843.

Matrosov, S.Y., A.V. Korolev, and A.J. Heymsfield, 2002: Profiling Cloud Ice Mass and Particle Size from Doppler Radar Measurements, *J. Atmos. Ocean. Technol., 19,* 1003-1018.

Menzel W.P., W.L. Smith, and T. R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment using VAS. *J. Clim. Appl. Meteor.*, **22**, 377-384.

Menzel W.P. R.A. Frey, B.A. Baum, and H. Zhang, 2006: Cloud Top Properties and Cloud Phase, MODIS Algorithm Theoretical Basis Document, Version 7 (http://modis.gsfc.nasa.gov/data/atbd)

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

Minnis, P., D.P. Garber, D.~F. Young, R.~F. Arduini, and Y. Takano, 1998: Parameterizations of reflectance and effective emittance for satellite remote sensing of cloud properties, *J. Atm. Sci.*, *55,* 3313-3339.

Mlawer, E., and S. Clough, 1997: On the extension of rapid radiative transfer model to the short-wave region, in *Proceedings of the 6th Atmospheric Radiation Measurement(ARM) Sciene Team Meeting*, U.S. Deparment of Energy, CONF-9603149.

Naisiri, S.L. and B.A. Baum, 2004: Daytime multi-layered cloud detection using multi-spectral imager data, *J. Atmos. Ocean. Technol., 21,* 1145-1155.

Nakajima T. and M.D. King, 1990: Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements, part 1: Theory, *J. Atm. Sci.,* **47***,* 1878-1893.

Nakajima T.Y. and T. Nakajima, 1995: Determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEX regions, *J. Atm. Sci,* **52***,* 4043 – 4059.

Nunez, M. K. Fienberg, and C. Kuchinke, 2005: Temporal Strucutre of the Solar Radiation Field in Cloudy Conditions: Are Retrievals of Horuly Averages from Space Possible?, *J. Appl.. Meteor.,* **44**(1), 167-178.

Platnick, S., 2000: Vertical photon transport in cloud remote sensing problems. *J. Geophys. Res*., **105** (D18), 22919-22935.

Platnick, S., 2001: A Superposition Technique for deriving Mean Photon Scattering Statistics in Plane-Parallel Cloudy Atmospheres, *J. Quant. Spectros. Radiat. Transfer,* **68**, 57-73.

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

Roebeling, R.A., A. Berk, A.J. Feijt, W. Frerichs, D. Jolivet, A. Macke, and P. Stammes, 2005: Sensitivity of cloud property retrievals to differences in narrow band radiative transfer simulations, *Sci. Rep WR 2005-02*, Koninklijk Ned. Meteorol. Inst., De Bilt, the Netherlands.

Roebeling, R. A., A. J. Feijt, and P. Stammes, 2006a: Cloud property retrievals for climate monitoring: Implications of differences between Spinning Enhanced Visible and Infrared Imager (SEVIRI) on METEOSAT-8 and Advanced Very High Resolution Radiometer (AVHRR) on NOAA-17, *J. Geophys. Res.*, **111**, D20210, doi:10.1029/2005JD006990

Roebeling, R.A., N. Schutgens, and A.J. Feijt, 2006b: Analysis of uncertainties in SEVIRI cloud property retrievals for climate monitoring, Proceedings: published, 12th Conference on Atmospheric Radiation, 2006, Madison, USA, AMS

Roebeling, R.A., Deneke, H. and A.J. Feijt, 2006c: Validation of cloud liquid water path retrievals from SEVIRI using one year of cloudNET observations, *Submitted to JAMC*.

Roebeling R.A., H. Hauschildt, D. Jolivet, E.A. van Meijgaard and A.J. Feijt, 2003: Retrieval and validation of METEOSAT-8 and AVHRR based cloud physical parameters in the CM-SAF. *Proceedings 2003 EUMETSAT Data Users Conference, 2003, Weimar, Germany*.

Roebeling R.A., A.J. Feijt, and P. Stammes, 2006: Cloud property retrievals for climate monitoring: implications of differences between SEVIRI on METEOSAT-8 and AVHRR on NOAA-17, J. Geophys. Res., (accepted).

Rosenfeld, D., E. Cattani, S. Melani, and V. Levizzani, 2004: Considerations on daylight operation of 1.6-versus 3.7-μm channel on NOAA and Metop satellites, *Bull. Amer. Meteor. Soc.*, **85**, 873–881.

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

Stammes, P. 2001: Spectral radiance modelling in the UV-visible range, Proceedings IRS-2000: Current problems in atmospheric radiation, edited by W.L. Smith and Y.M. Timofeyev, 181-184, A. Deepak Publ., Hampton, Virginia, USA.

Stamnes, K., S. Tsay, W. Wiscombe, and K. Jayaweera, 1988: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emittting layered media, *Appl. Opt.*, **27**, 2502-2509.

Stephens G.L., G.W. Paltridge and C.M.R. Platt, 1978: Radiation profiles in extended water clouds. III. Observations. *J. Atm. Sci*, **35**, 2133-2141.

Tian, L., and J.A. Curry, 1989: Cloud overlap statistics, *J. Geophys. Res., 94*, 9925-9935.

Trolez. M., A. Tetzlaff, and K.-G. Karlsson, 2005: Validating the Cloud Top Height product using LIDAR data, CM-SAF Visiting Scientist report, *SMHI Reports Meteorology*, No. 118.

Watts, P.D., C.T. Mutlow, A.J. Baran, and A.M. Zavody, 1998: Study on Cloud Properties derived from Meteosat Second Generation Observations. Final Report, EUMETSAT ITT no. 97/181.

Wolters, E.L.A., R.A. Roebeling, and A.J. Feijt, 2006: Quantitative assessment of cloud phase determination from SEVIRI onboard METEOSAT-8 using ground-based LIDAR and cloud radar data, *submitted to J. Appl. Meteor. Clim.*