Atmospheric
Measurement
Techniques

Open Access

Comparing satellite- to ground-based automated and manual cloud coverage observations – a case study

A. Werkmeister[1,4], M. Lockhoff[2], M. Schrempf[1], K. Tohsing[1], B. Liley[3], and G. Seckmeyer[1]

[1]Institut für Meteorologie und Klimatologie, Leibniz Universität Hannover, Herrenhaeuser Str. 2, 30419 Hannover, Germany
[2]Deutscher Wetterdienst, Frankfurter Str. 135, 63067 Offenbach, Germany
[3]National Institute of Water and Atmospheric Research (NIWA), Lauder, Central Otago, New Zealand
[4]now at: Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, 3149 Miami, USA

Correspondence to: G. Seckmeyer (seckmeyer@muk.uni-hannover.de)

Abstract. In this case study we compare cloud fractional cover measured by radiometers on polar satellites (AVHRR) and on one geostationary satellite (SEVIRI) to ground-based manual (SYNOP) and automated observations by a cloud camera (Hemispherical Sky Imager, HSI). These observations took place in Hannover, Germany, and in Lauder, New Zealand, over time frames of 3 and 2 months, respectively.

Daily mean comparisons between satellite derivations and the ground-based HSI found the deviation to be $6 \pm 14\%$ for AVHRR and $8 \pm 16\%$ for SEVIRI, which can be considered satisfactory. AVHRR's instantaneous differences are smaller ($2 \pm 22\%$) than instantaneous SEVIRI cloud fraction estimates ($8 \pm 29\%$) when compared to HSI due to resolution and scenery effect issues. All spaceborne observations show a very good skill in detecting completely overcast skies (cloud cover ≥ 6 oktas) with probabilities between 92 and 94 % and false alarm rates between 21 and 29 % for AVHRR and SEVIRI in Hannover, Germany. In the case of a clear sky (cloud cover lower than 3 oktas) we find good skill with detection probabilities between 72 and 76 %. We find poor skill, however, whenever broken clouds occur (probability of detection is 32 % for AVHRR and 12 % for SEVIRI in Hannover, Germany).

In order to better understand these discrepancies we analyze the influence of algorithm features on the satellite-based data. We find that the differences between SEVIRI and HSI cloud fractional cover (CFC) decrease (from a bias of 8 to almost 0 %) with decreasing number of spatially averaged pixels and decreasing index which determines the cloud coverage in each "cloud-contaminated" pixel of the binary map. We conclude that window size and index need to be adjusted in order to improve instantaneous SEVIRI and AVHRR estimates. Due to its automated operation and its spatial, temporal and spectral resolution, we recommend as well that more automated ground-based instruments in the form of cloud cameras should be installed as they cover larger areas of the sky than other automated ground-based instruments. These cameras could be an essential supplement to SYNOP observation as they cover the same spectral wavelengths as the human eye.

1 Introduction

Clouds play an important role for solar and terrestrial radiation. As a consequence, clouds have an impact on the energy budget and global climate. A small change in cloud parameters may significantly change the temperature variation (Seinfeld and Pandis, 1998). High clouds, in general, act as a greenhouse gas and warm the Earth, whereas low clouds can cool the Earth by reflecting the radiation back to space (Liou, 1991). Several researchers proposed that the effect of clouds enhances global and UV radiation (Calbo et al., 2005; Schafer et al., 2012; Poetzsch-Heffter et al., 1995; Solomon et al., 2007). Clouds mediate the indirect effect of aerosol on radiation (Forster et al., 2007). Albrecht (1989) explained that increases in aerosol concentrations over the oceans increase the amount of low-level cloudiness. Further-

more, analysis by Clement et al. (2009) shows observational and model evidence that changes in low-level clouds act as a positive feedback over the ocean. Since the feedback of cloud coverage on the climate is the biggest uncertainty in climate research and forecasts (Solomon et al., 2007), it is essential to investigate and improve cloud coverage measurements, both ground- and space-based.

Ground observations and measurements are not provided in a sufficient spatial coverage although they often cover longer time scales than satellite observations. Only space-based observations can deliver the necessary global coverage with sufficient quality and long time frames. Particularly over the ocean and inaccessible regions satellites are largely the only data source (Ohring et al., 2005).

Already in the past 100 years the study of cloud parameters has been of high interest. Human observations were the first method of cloud coverage determination. Observers classify clouds according to the subjective view of shape and appearance (Robaa, 2008) and estimate sky coverage. During the last years more and more of these human-based observations were replaced by ground-based automated instruments to obtain a higher consistency in cloud coverage estimations (Orsini et al., 2002; Dürr and Philipona, 2004).

In the early 1970s Malberg (1973) compared cloud cover from satellite photographs to ground-based synoptic cloud observations and found mean annual differences of about 9 % over northern Europe and 15 % over southern Europe. He explained the differences between ground-based observations and satellite imagery with geometric, synoptic and orographic factors. With increasing availability of satellite data, scientists started deriving cloud properties (Ackerman et al., 1998; Christodoulou et al., 2003; Ebert, 1987; Gao and Wiscombe, 1994; Garand, 1988; Parikh, 1977; Porcú and Levizzani, 1992; Romano et al., 2007; Saunders and Kriebel, 1988; Schröder et al., 2002; Welch et al., 1992). The cloud detection threshold test by Derrien et al. (1993) is a real-time processing scheme that is applied to the different channels of irradiances from the NOAA-11 satellite. This algorithm was further developed and adjusted to new instruments on further satellites (Derrien and LeGleau, 2005, 2013).

Dürr and Philipona (2004) developed an automatic partial cloud amount detection algorithm that estimates cloud coverage from surface long-wave downward radiation, surface temperature and relative humidity. Schade et al. (2009) validated the algorithm by Dürr and Philipona (2004) against human observations and digital all-sky imaging. The results show that the differences between algorithm and imaging are lower than between algorithm and human cloud estimations.

Boers et al. (2010) conducted ground-based measurements with five different methods that were either performed by passive or active remote sensing instruments. These measurements were compared to a 30-year climatology of human observations. They concluded that of course it is unrealistic to expect complete similarity between observer and instrumental outputs. The lack of sunlight during night compounds the difficulty of cloud detection for the observer. Observers as well as some instruments were unable to detect very high thin and wispy clouds.

Schutgens and Roebeling (2009) analyzed the influence of cloud inhomogeneity on intercomparisons of liquid water distribution retrievals by a geostationary satellite imager and a ground-based microwave radiometer. They classified the validation errors due to this inhomogeneity into two categories: retrieval process for satellite observations (plane-parallel bias and field-of-view mismatches between the radiometer's channels) and differences in observed scenery (by satellite- and ground-based measurements). Schutgens and Roebeling (2009) conclude that the dominating error is due to scene differences and that smaller pixel sizes increase this behavior unless the parallax effect is corrected. Greuell and Roebeling (2009) established standards for validation procedures to minimize these errors by determining the optimum statistical agreement between satellite- and ground-based liquid water path measurements. The parallax correction led to a significant improvement in validation. However, the same correction did not significantly improve results for relatively homogeneous cloud fields.

Martinez-Chico et al. (2011) performed comparisons of cloud classification from different ground-based instruments. In this case they used radiation data and hemispherical sky images to determine different cloud types. They also proposed to use this kind of studies to determine sites for solar panels to improve solar resource assessment models. Kazantzidis et al. (2012) compared an automatic estimation of the cloud coverage and classification derived from a simple whole sky imaging system to synoptic data. According to their results, 83 % (broken cloudiness) and 94 % (overcast cloudiness) of the analyzed images agreed within ±1 and ±2 oktas, respectively, compared to the weather observations. They also concluded that the total cloud cover is underestimated when cirrus clouds are present.

The Satellite Application Facility on Climate Monitoring (CM SAF) which is part of the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) SAF network, generates, archives and distributes satellite-derived products for climate monitoring in an operational mode. CM SAF distributes, among others, cloud products (Cloud Fractional Cover, Cloud Type, Cloud Top Pressure, etc.) derived from the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on the first Meteosat Second Generation (MSG) geostationary spacecraft and from the Advanced Very High Resolution Radiometer (AVHRR) from the polar-orbiting NOAA (National Oceanic and Atmospheric Administration) satellites. These products are directly derived from satellite radiance measurements.

CM SAF published several validation reports on cloud products in the past years. Deneke et al. (2007) examined cloud fractional cover (CFC) comparisons over an 8-month period between SEVIRI and SYNOP (surface synoptic observations) in 2007 with focus on instantaneous, daily mean

(DM) and monthly mean (MM) times scales. The bias (mean difference) for the instantaneous and DM CFC of each month was approximately 4 and 12 %, respectively, which is consistent with previous works.

Reuter et al. (2009) validated SEVIRI with synoptic CFC and also initial CFC comparisons with MODIS (Moderate Resolution Imaging Spectrometer) and CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) CFC measurements. These results show that the CFC from CM SAF agreed well with synoptic data (within 1 okta difference) and polar orbiting satellite data over mid-latitudes. However, the CFC was found to be overestimated towards the edges of the visible Earth disk. They concluded that the clouds might be identified correctly by SEVIRI instrument but are interpreted incorrectly by the algorithm. The results show that the horizontal cloud coverage seems larger than in reality just by geometrical viewing effects. The parallax effect results in a displacement of the cloud positions in relation to the sensor when approaching the edges of foot print. In the same case the position as well as the length of these clouds is misinterpreted and the CFC is overestimated.

Amato et al. (2008) performed a statistical analysis of cloud detection from SEVIRI imagery. Their discriminant analysis showed a good performance in cloud detection.

In a contrail study by Mannstein et al. (2010) instantaneous comparisons to a Wolkam camera showed that the SEVIRI cloud detection algorithm detected 15 % of 79 contrails. Note that contrails are hard to detect with passive sensors, since they are very thin. The same study for the AVHRR algorithm showed better results due to a higher spatial resolution of the instrument. Of the contrails, 27 % were confirmed by the Wolkam camera (detailed results in Mannstein et al., 2010).

We compare the CFC products provided by CM SAF with ground-based observations. This means the CFC data are checked in a process of comparisons in order to determine the resemblance of instantaneous and DM satellite data.

We will describe the instruments – SEVIRI, AVHRR and Hemispherical Sky Imager (HSI) – and how the data were retrieved and processed in this work. After introducing the methodology, we will present comparisons between the different data sets (SEVIRI, AVHRR, SYNOP and HSI). We will continue with an analysis of the characteristics of the CFC retrieval algorithm and finally discuss and conclude the results.

2 Instruments and data

2.1 The Hemispherical Sky Imager

The HSI is composed of a digital compact charge-coupled device camera, a fish-eye objective with a field of view of 183° and a steering unit to provide a hemispherical image of the entire sky. This system is installed on the roof of the Insti-

tute of Meteorology and Climatology (IMUK) in Hannover, Germany (52.4° N, 9.7° E), and is protected by a waterproof enclosure. More details of the HSI system are described in Tohsing et al. (2013). The image acquisition for the cloud coverage determination is performed within 10 s intervals. An identical system is mounted at NIWA (National Institute of Water and Atmospheric Research) in Lauder, New Zealand (45.0° S, 169.7° E). A camera projection, which describes the relationship between the incoming light ray and the incident angle, needs to be considered in order to estimate the cloud cover from the HSI image. Tohsing et al. (2013) analyzed the camera projection of this camera system and found it to be adequate for the cloud cover determination. The equidistant camera projection has the advantage that the acquired image is only minimally distorted and clouds can be analyzed to zenith angles of 80°. The cloud cover of the sky with a zenith angle greater than 80° is not analyzed due to horizontal brightening and hazy sky. The spatial horizontal coverage of the HSI instrument depends on the considered field of view and the cloud base height. By assuming a field of view of 160° – thus ignoring the sky between the horizon and the elevation angle of 10° – and a cloud base height of 3 km, the spatial horizontal coverage can be up to 900 km^2. With these assumptions the radius of the circular area is 17 km. With a decreasing cloud base height the spatial horizontal coverage is reduced. At a height of 1.5 km the coverage is approximately 225 km^2.

In order to extract the CFC from red-green-blue signal counts we used an algorithm based on the approach by Yamashita et al. (2004). We define the SkyIndex in order to separate blue sky and cloud areas.

Since the SkyIndex by Yamashita et al. (2004) cannot analyze hemispheric images with an adequate accuracy, we further developed the algorithm. In addition to a sun filter, a haze filter was implemented in the algorithm to analyze uncertain or hazy areas in the digital image by taking into account the green signal counts. The haze filter defines a hazy area if the value of the green signal count is greater than the average of red and blue. A cloud is defined by the haze filter if the green signal count is smaller than the average.

The position of the sun in the image is calculated in order to evaluate the mostly bright circular solar area with an additional sun filter. In contrast to the SkyIndex, the sun filter uses different thresholds which are optimized for the higher and saturated signal. The algorithm is computing the CFC with a spatial resolution of approximately 3 megapixels.

2.2 Surface synoptic observations

SYNOP is a numerical code introduced by the World Meteorological Organization for weather observations made at manual and automated weather stations. Besides many meteorological parameters (local temperature, precipitation, visibility etc.) the CFC is reported at standard synoptic times. At these times a synoptic observer at a specific location reads the

instruments and also estimates several variables, including visibility and CFC. CFC is reported in oktas, ranging from completely clear (0 oktas) to completely overcast (8 oktas). The hourly CFC SYNOP observations used within the study take place about 10 km from the HSI site at Hannover Airport, denoted hereafter HAJ.

2.3 Space-based measurements

2.3.1 Instruments

Spinning Enhanced Visible and Infrared Imager)

On board the geostationary MSG satellites at 36 000 km height above the equator, SEVIRI provides full disc imagery (at 0° latitude and longitude) over Europe and Africa every 15 min (HH : 00, HH : 15, HH : 30 and HH : 45). The first MSG was launched in 2002 and has been delivering data ever since. The second MSG was launched in 2005 and in 2012 MSG-3 was launched. The fourth MSG is scheduled for launch in 2015.

SEVIRI is an optical imaging radiometer with 12 channels in the visible, near-infrared and thermal infrared part of the spectrum, between 0.6 and 13.4 µm (Aminou, 2002), and provides unique capabilities for cloud imaging and tracking, fog detection, measurement of the Earth-surface and cloud-top temperatures, tracking of ozone patterns and many other improved measurements. SEVIRI has a spatial resolution of 3 km × 3 km at the nadir (Aminou, 2002). A complete image of the Earth's full disk consists of 3712 × 3712 pixels.

Advanced Very High Resolution Radiometer

AVHRR is one of the longest operating satellite instruments to date. It operates on board the polar orbiting NOAA satellites and is also carried by the Meteorological Operational Satellites (MetOp)-A and MetOp-B polar orbiter operated by EUMETSAT since 2006. These measurements began already in the late 1970s and have continued until today (Kogan et al., 2011). The NOAA satellites 15, 16, 17, 18, 19 and MetOp-A and MetOp-B (MetOp-C will be launched in 2017) belong to the Polar Operational Environmental Satellite program. These satellites are all equipped with the third version of the AVHRR. The AVHRR is a scanning radiometer, meaning it makes calibrated measurements of upwelling radiation from small areas (scan spots or pixel) which are scanned across the sub-satellite track. The operation of the AVHRR is representative of many scanning radiometers on low Earth orbiters (Kidder and Von der Haar, 1995). This scanning radiometer uses six detectors that collect different bands of radiation at wavelengths between 0.58 and 12.50 µm.

2.3.2 Algorithms

The CMSAF cloud mask (CMa) products are based on algorithm packages provided by the Satellite Application Facility in supporting NoWCasting and very short range forecasting (NWCSAF). Two different algorithms have been developed for the two different radiometers (SEVIRI and AVHRR) because of their different channel characteristics. Both algorithms are based on the same concept: the cloud detection is performed by a multi-spectral thresholding technique. This means that a series of threshold tests allow the identification of pixels which are contaminated by clouds, snow or ice. These tests are applied to land or sea pixels depending on the illumination conditions (daytime, nighttime, etc.). Most thresholds are dynamically determined from ancillary data using radiative transfer models. If one test is well above its threshold, the process is stopped. The tests with the respective thresholds are detailed in Derrien and LeGleau (2005) and Derrien and LeGleau (2013) for SEVIRI and Dybbroe et al. (2005) for AVHRR, respectively.

2.3.3 Data

SEVIRI data

We use Level 2 data provided by CM SAF. We are using a 3-month extract of the data set CLAAS (CLoud property dAtAset using SEVIRI) which is an 8-year record of satellite-based cloud properties. The SEVIRI cloud products are derived from the space-based radiometers using the MSG NWC software package version v2010 (Stengel et al., 2014).

For the calculation of CFC the original CMa fields were transformed into an equal-area (sinusoidal) projection with a spatial resolution of 3 km × 3 km resulting in a field of 5925 × 5925 pixels. Each pixel contains information of the cloud situation (cloud-free, cloud-contaminated, cloud-filled, ice-contaminated and no data). The 5925 × 5925 CMa field is then transformed into a binary map. This is done by assigning the value "1" to the pixels classified "cloud-filled" and "cloud-contaminated", whereas "cloud-free" and "ice-contaminated" pixels are assigned the value of "0". This assigned value will later be introduced as the cloud layer index. No-data pixels are assigned N/A values. Linear averaging of the CMa binary map over 5 × 5 grid boxes leads to the final 1185 × 1185 pixel grid. In accordance with CM SAF processing the CFC was calculated as the fraction of cloudy pixels (cloud-filled and cloud-contaminated) per subregion (5 × 5 grid boxes) compared to the total number of analyzed pixels per same subregion, which means that the CFC is computed as the cloudy fraction of all pixels within a 15×15 km^2 grid square and is expressed in percent (Derrien and LeGleau, 2005).

AVHRR data

Concerning the AVHRR-based cloud cover information, we used CM SAF's new cloud climate data record which is based on AVHRR Global Area Coverage data: CLARA-A1 (CLoud, Albedo and RAdiation data set, AVHRR-based, version 1, Karlsson et al., 2013). This particular AVHRR data set has its strengths in its long duration (28 years) and foundation upon a homogenized AVHRR radiance data record. The instantaneous CLARA-A1 retrievals have a spatial resolution of $4\,\text{km} \times 4\,\text{km}$. This spatial resolution results from averaging over 4 out of 5 pixels and skipping three lines in the original high-resolution picture transmission (Karlsson et al., 2013). This data set is based on the adjusted NWC-SAF/PPS version 2010 algorithm for polar orbiters (Karlsson et al., 2005).

3 Methodology

3.1 Data processing

Both instantaneous satellite data sets (SEVIRI and AVHRR) had to be temporally and spatially analyzed and sorted in order to be compared to ground-based observations (SYNOP and HSI).

3.1.1 SEVIRI data processing

Since the CM SAF SEVIRI CFC is distributed for HH : 45 (scan starting time) and the scan by SEVIRI takes 12 min (Schmetz et al., 2002) and reaches the area over Hannover after approximately 10 min, the time HH : 00 was chosen for the CFC calculation by HSI and also SYNOP. Only values measured under the conditions of a solar zenith angle lower than 80° were accepted, since HSI and SYNOP data are based on the visible spectrum of the solar radiation (camera and human eye). Especially during dusk and dawn the cloud state can be misinterpreted due to reflection at the horizon for example.

The CMa's cloudy pixels can be labeled as either cloud-filled or contaminated. CM SAF assumes a 100 % cloud coverage for both cases. We define BCLI (broken-cloud layer index) as the value that is assigned to a "cloud-contaminated" pixel (in case of CM SAF: BCLI = 100 %). We will also analyze the influence of different BCLI as well of different sizes of averaging windows on the CFC. Window sizes of 3×3, 5×5 and 7×7 pixels are included in the calculations. We also replace the original BCLI of 100 % with BCLIs of 50 and 75 % in order to determine the influence of BCLI on CFC estimations in SEVIRI data. Subsequently this CFC is compared to the CFC from HSI.

3.1.2 AVHRR data processing

The HSI data points are chosen according to the overflight time of the polar satellites (NOAA satellites 15, 16, 17, 18, 19 and MetOp-2) during daylight as in the SEVIRI case. On average, 20 values of AVHRR CFC are computed for each day, for the box centered over Hannover, Germany. The AVHRR CFC pictures are chosen according to the position of the HSI. Only AVHRR and HSI CFC with a zenith angle lower than 40° are chosen for comparisons. These are approximately 10 values per day. The first step of our algorithm consists in searching these auxiliary data in order to find the temporally correlated HSI image to which the CFC will be compared. After finding the central pixel ($4 \times 4\,\text{km}^2$), 3×3 pixels ($\approx$ spatial resolution of SEVIRI CFC) are averaged. The result is a $12 \times 12\,\text{km}^2$ box containing the CFC in percent.

3.2 Statistics

We use different statistic relations in order to compare the different data sets. We distinguish between instantaneous and daily mean cloud coverage data which are calculated from the instantaneous data.

The daily mean) is defined by

$$\text{DM} = \frac{1}{k} \sum_{n=1}^{k} \text{CFC}(n), \tag{1}$$

where k is the number of CFC values in 1 day.

In order to quantify over- and underestimation of the CFC by the instruments, we define the bias as the mean of differences. The equation becomes

$$\text{Bias} = \frac{1}{B} \sum_{m=1}^{B} x_m, \tag{2}$$

where B equals the number of available match-ups between two data sets and x_m is the difference between these data sets.

The standard deviation (SD) is defined by

$$\text{SD} = \sqrt{\frac{1}{l} \sum_{n=1}^{l} (x_n - \mu)^2}, \tag{3}$$

where x_n is the difference between two data sets, l is the number of available values and $\mu = \frac{1}{l} \sum_{n=1}^{l} x_n$ is the mean of these differences.

The correlation coefficient $R(y, z)$ is used for comparisons between instantaneous data and is defined by

$$R(y, z) = \frac{\sum_{j=1}^{N} (z_j - \overline{z})(y_j - \overline{y})}{\sqrt{\sum_{j=1}^{N} (z_j - \overline{z})^2 \cdot \sum_{j=1}^{N} (y_j - \overline{y})^2}}, \tag{4}$$

where z_j are the values of CFC by SEVIRI, y_j are the values of CFC by HSI and $\overline{z}, \overline{y}$ are the arithmetical means of CFC by the respective instrument.

Table 1. Contingency table as shown in Reuter et al. (2009). Cloud-free (clear) is defined as 0–2 oktas, cloud-contaminated (broken) as 3–5 oktas and cloud-covered (cloudy) stands for cloud coverage between 6–8 oktas.

		Reference data		
	Scenario	Clear	Broken	Cloudy
	Clear	a	b	c
Satellite or SYNOP	Broken	d	e	f
	Cloudy	g	h	i

For the instantaneous cloud coverage data we distinguish in our work between three different scenarios: cloud-free (clear) (CFC ≤ 2 oktas), cloud-contaminated (broken) (3 oktas $\leq$ CFC ≤ 5 oktas) and cloud-covered (cloudy) (CFC ≥ 6 oktas). Comparing two data sets, the variables a, b, c, d, e, f, g, h and i give the number of observations for different combinations of scenarios of cloud-free, cloud-contaminated and cloud-covered sky (as shown in the contingency matrix in Table 1). The probability of detection (POD) indicates the probability of correctly detecting the cloud scenario as seen by the following reference:

$$\text{POD}_{\text{clear}} = \frac{a}{a+d+g}, \tag{5}$$

$$\text{POD}_{\text{broken}} = \frac{e}{b+e+h} \tag{6}$$

and

$$\text{POD}_{\text{cloudy}} = \frac{i}{c+f+i}. \tag{7}$$

The false alarm rate (FAR) is a measure of observation performance (just as POD) and describes for each scenario the ratio between the number of false alarm events and the total number of events:

$$\text{FAR}_{\text{clear}} = \frac{b+c}{a+b+c}, \tag{8}$$

$$\text{FAR}_{\text{broken}} = \frac{d+f}{d+e+f} \tag{9}$$
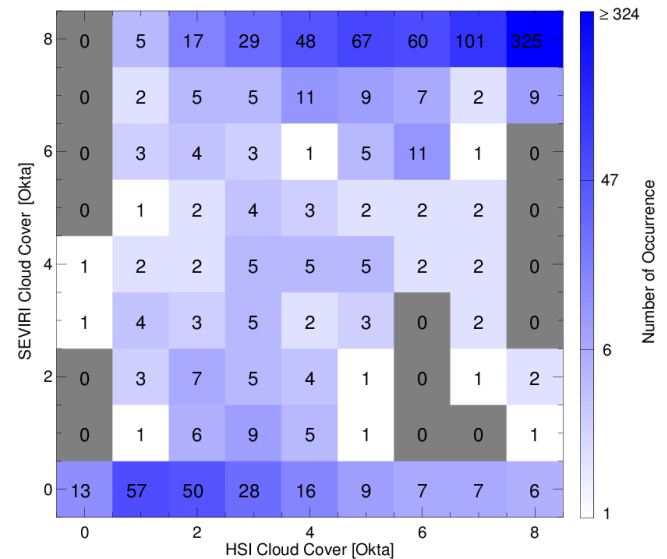
and

$$\text{FAR}_{\text{cloudy}} = \frac{g+h}{g+h+i}. \tag{10}$$

4 Results

4.1 Hannover, Germany – comparison

This section will present all the results of the comparisons between HSI, SEVIRI, SYNOP and AVHRR in Hannover, Germany, for the months July through September 2009. The results of the instantaneous data will be followed by the results for the daily mean data sets.

Figure 1. Density plot of the occurrences of the CFC by HSI as a function of instantaneous CFC in oktas by SEVIRI. Each color of one box represents the amount of matches at the respective CFCs by HSI and SEVIRI. The total of all valid matches is 957 and represents the results from 1 July to 30 September 2009 for Hannover.
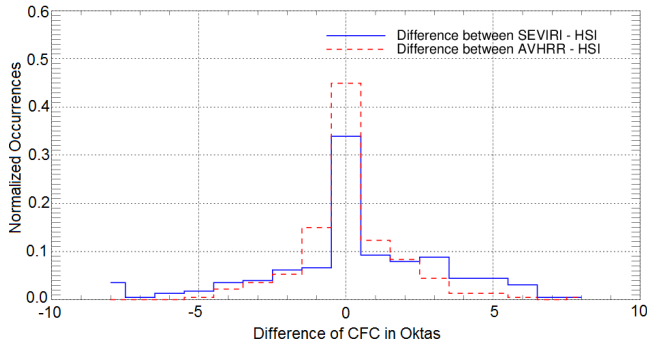
4.1.1 Instantaneous

SEVIRI

Figure 1 shows a density plot for instantaneous CFC of HSI and SEVIRI for the months July through September 2009. The number of occurrences for both instruments measuring 8 oktas is 325. We find that from all 1029 valid measurements, 371 ($\approx 36\%$) match a difference of 0 okta and so agree with each other. Of the 652 measurements by SEVIRI that show a CFC of 8 oktas, in 50 % HSI measures a CFC between 1 and 7 oktas, which means that half of SEVIRI's cloud-covered skies are overestimated. In 193 cases we find a CFC of 0 okta by SEVIRI and 93 % of these measurements show a difference higher than 1 okta from HSI CFC. When SEVIRI and HSI return CFCs between 1 and 7 oktas, SEVIRI overestimates 61 % of 148 match-ups.

Figure 2 shows (among other results) a histogram of the differences between SEVIRI and HSI (here: red dashed line). The reader should notice that positive differences between SEVIRI and HSI are on average higher than the negative differences. This observation shows that SEVIRI often overestimates the CFC when compared to HSI. The bias is 8 % and also shows that SEVIRI tends to overestimate CFC. Martinez-Chico et al. (2011) explain that such an overestimation can be due to off-nadir effects and different viewing angles. We suggest that these deviations are also due to different spatial resolutions of the instruments as well as cloud-contaminated pixels in the cloud mask which are assigned a

Figure 2. Occurrences of the instantaneous differences (blue solid line: SEVIRI minus HSI; red dashed line: AVHRR minus HSI) in oktas in Hannover from 1 July to 30 September 2009. Occurrences are normalized with their total count. Negative differences express an underestimation of the CFC by the satellites compared to HSI; positive differences express an overestimation.

Table 2. POD_{clear}, POD_{broken}, POD_{cloudy}, FAR_{clear}, FAR_{broken} and FAR_{cloudy} (clear: cloud-free, broken: cloud-contaminated, cloudy: cloud-covered) in percent between HSI and SEVIRI in Hannover, Germany, from 1 July to 30 September 2009; total of 1027 valid matches (not matched to AVHRR).

POD_{clear}	POD_{broken}	POD_{cloudy}	FAR_{clear}	FAR_{broken}	FAR_{cloudy}
72 %	12 %	94 %	43 %	43 %	29 %

100 % cloud coverage. This problem will be discussed further in Sect. 4.3.

Statistical analysis of these two data sets for all three scenarios shows that clear and complete covered sky are the cases of good agreement whereas broken-cloud coverage show the highest deviations. $POD_{cloudy} = 94$ % shows the highest value in this comparison, whereas the $FAR_{cloudy} = 29$ %. This means when HSI detects cloudy sky, 94 % are detected by SEVIRI as well but 29 % of the cases identified by SEVIRI as cloudy are false alarms.

POD_{broken} (12 %) is almost one-eighth of POD_{cloudy}. These results indicate that SEVIRI shows only poor skill in detecting broken-cloud events. A summary of these values is shown in Table 2.

We also believe that moisture in the upper atmosphere and spatial resolution differences influence the order of magnitude of deviations. We show an example of high altitude moisture in the atmosphere in Fig. 3 where we present CFC measurements on 15 July 2009. The four HSI images on the top of Fig. 3 indicate the cloud coverage at 7:00, 8:00, 11:50 and 12:00 UTC. The CFC by SEVIRI, HSI and SYNOP is displayed in the plot underneath. Here we find differences between SEVIRI and HSI between −30 and 40 % (positive = overestimation by SEVIRI). In the HSI image of 7:00 UTC we can see cirrostratus clouds which indicate moisture in the high altitudes. Although SYNOP observes a 50 % CFC, HSI measures a CFC of around 75 % and SE-
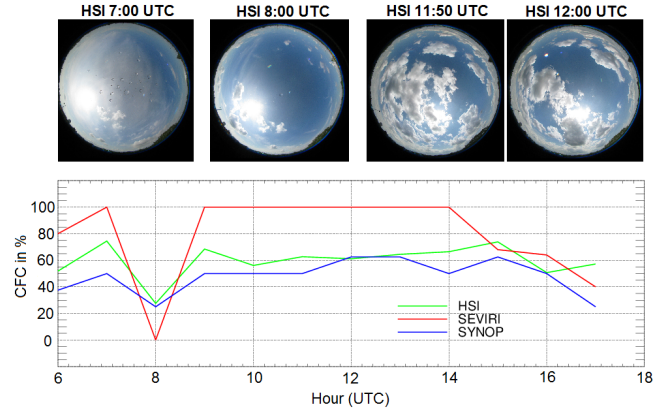
Figure 3. Instantaneous CFC by HSI (green), AVHRR (purple), SEVIRI (red) at IMUK and SYNOP (blue) at HAJ for 15 July 2009. Upper HSI images show the sky at 8:00, 9:00, 11:50 and 12:00 UTC (from left to right).

Table 3. Contingency table comparing HSI and SEVIRI CFC for the cases of cloud-free (0–2 oktas), broken clouds (3–5 oktas) and cloud-covered (6–8 oktas). Values represent ratio between matchups and the total of 957 valid measurements. Similar numbers are found for AVHRR.
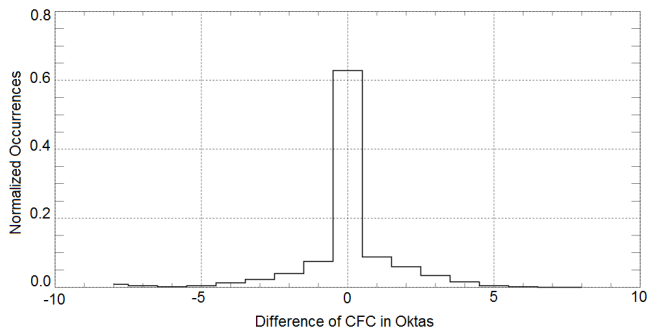
		HSI Hannover		
	Scenario	Clear	Broken	Cloudy
	Clear	13 %	8 %	2 %
SEVIRI	Broken	2 %	3 %	1 %
	Cloudy	3 %	17 %	50 %

VIRI of 100 %. HSI CFC is already overestimated due to dew on the dome. At 8:00 UTC SEVIRI measures a CFC of 0 % whereas HSI and SYNOP estimate a CFC between 25 and 30 %. Due to its coarser resolution, SEVIRI is unable to correctly detect small clouds as seen in the HSI picture. Increasing occurrence of these clouds, which cannot be detected by the satellite instrument but by ground-based observers, also increase differences between the measurement systems. This can also be seen in the examples of 11:50 and 12:00 UTC. Obviously HSI can capture even small changes in cloud coverage whereas SEVIRI, due to its coarser spatial resolution and viewing conditions, does not detect the same changes in CFC.

In conclusion, the results of instantaneous CFC show that there is a chance of one in three that SEVIRI measurements differ from HSI measurements by at least 1 okta (refer to Table 3: 100 % $-a-e-i$). In 74 % of these cases these differences will occur during cloud-contaminated sky $((b+h)/33$ %).

Table 4. POD_{clear}, POD_{broken}, POD_{cloudy}, FAR_{clear}, FAR_{broken} and FAR_{cloudy} (clear: cloud-free, broken: cloud-contaminated, cloudy: cloud-covered) in % between HSI and AVHRR in Hannover, Germany, from 1 July to 30 September 2009 (not matched to SEVIRI) and Lauder, New Zealand, from 1 November to 31 December 2009.

	POD_{clear}	POD_{broken}	POD_{cloudy}	FAR_{clear}	FAR_{broken}	FAR_{cloudy}
Hannover	76 %	32 %	92 %	42 %	38 %	21 %
Lauder	68 %	26 %	85 %	29 %	68 %	22 %

Figure 4. Histogram of instantaneous CFC differences between AVHRR and HSI in Hannover, Germany, from 1 July to 30 September 2009. The occurrences are normalized with their respective total count. Negative differences express an underestimation of the CFC by AVHRR compared to HSI; positive differences express an overestimation.

AVHRR

We calculate the CFC as described in Sect. 3.1 for comparisons between CFC from AVHRR and HSI in Hannover.

Figure 4 displays the distribution of instantaneous CFC differences in oktas between AVHRR and the ground-based HSI in Hannover, Germany. The count is normalized by the total number of match-ups. We find that the count of differences equal to 0 is approximately equal to the number of differences not equal to 0. Nonetheless there is a slight overestimation by AVHRR as shown by the greater occurrences of positive differences and by the positive bias equal to 2 % with a standard deviation equal to 22 %.

In the cloud-covered scenario both POD and FAR show relatively good results (as shown in Table 4). $POD_{cloudy} = 92$ % has the highest value which indicates that 92 % of all data pairs agrees on a CFC between 6 and 8 oktas. For the same scenario, $FAR_{cloudy} = 21$ % represents the percentage of events that are false alarms and shows that cloud detection in this case is satisfactory. Nevertheless we find the $POD_{broken} = 32$ % and $FAR_{broken} = 38$ %, which means that more than one-third of AVHRR CFC between 3 and 5 oktas are false alarms.

AVHRR vs. SEVIRI

We compare instantaneous CFC by AVHRR to 1 h results by SEVIRI. Both data sets were temporally matched according to the satellite overflight time over Hannover, Germany, and the temporal resolution of SEVIRI imagery. Maximum temporal differences were ± 10 min. The HSI picture which was temporally closest to AVHRR was chosen and compared to AVHRR and SEVIRI CFC. A total of 227 match-ups were found.

It can be seen in the histograms in Fig. 2 that AVHRR has a lower count for overestimating CFC in respect to HSI and a significantly lower count in underestimating compared to SEVIRI measurements. Compared to SEVIRI, AVHRR shows in total a lower frequency at differences greater than 1 okta (positive and negative). We can also read from this figure, that in the few cases AVHRR underestimates CFC, it is underestimated mostly by 1 okta compared to HSI. The bias also shows this slight underestimation and is equal to -2 % with a standard deviation equal to 21 %. SEVIRI, however, shows a higher count for positive differences which implies that SEVIRI tends to overestimate CFC, which has already been shown in Sect. 4.1.1 and is also confirmed in this case with the positive bias of 3 % with a standard deviation equal to 38 %. In this case SEVIRI's standard deviation is higher than in all the other cases due to temporal matching. POD and FAR comparisons between SEVIRI/AVHRR and HSI reveal that all observations mostly agree in cases of cloud-covered sky. We find that AVHRR's FAR_{clear} is significantly lower than SEVIRI's FAR_{clear} (31 vs. 56 %). As well PODs in the same scenario are greater for AVHRR than SEVIRI (82 vs 64 %, respectively). SEVIRI's POD for cloud-contaminated skies (6 %) is almost one-fifth of AVHRR's POD (29 %). We conclude that AVHRR overall shows better skill in capturing the three different cloud scenarios.

SYNOP vs. SEVIRI

The CM SAF algorithm, as described in Sects. 2.3.2 and 3.1, has been used to calculate hourly instantaneous CFC to compare CFC between SYNOP and SEVIRI in Hannover–Langenhagen (HAJ).

We analyzed the deviations in CFC between the adjacent boxes over IMUK and HAJ in order to determine whether we can use SYNOP observations at HAJ for comparisons against SEVIRI and HSI estimates at IMUK. In Fig. 5 we present the
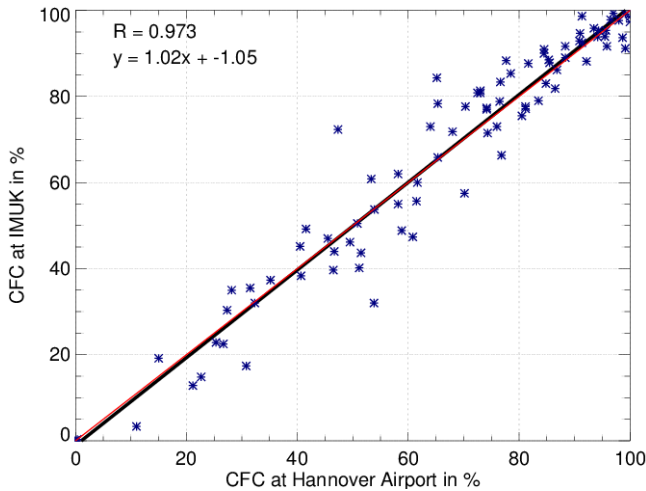
Figure 5. Relation between instantaneous SEVIRI cloud fractional cover (CFC) of pixels over Hannover Airport (HAJ) and the Institute of Meteorology and Climatology (IMUK).
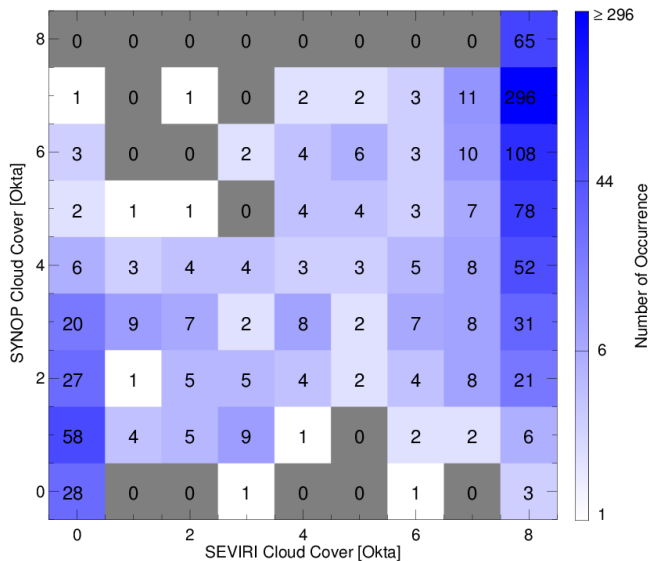
Figure 6. Density plot of the occurrences of the CFC by SYNOP as a function of instantaneous CFC in oktas by SEVIRI. Each color of one box represents the amount of matches at the respective CFCs by SYNOP and SEVIRI. The total of all valid matches is 996 and represents the results from 1 July to 30 September 2009 for Hannover, Germany.

relation between the CFCs at both sites and find that these measurements are indeed correlated with a correlation coefficient $R = 0.98$. The standard deviation is 5 % and the bias is equal to 1 %. We conclude that we can use the box over IMUK for SYNOP comparisons to satellite and HSI measurements.

Evaluations reveal that SYNOP data show a high variation in CFC when SEVIRI measures a CFC of 0 or 8 oktas (see Fig. 6). Of a total 996 match-ups, in 296 SEVIRI shows
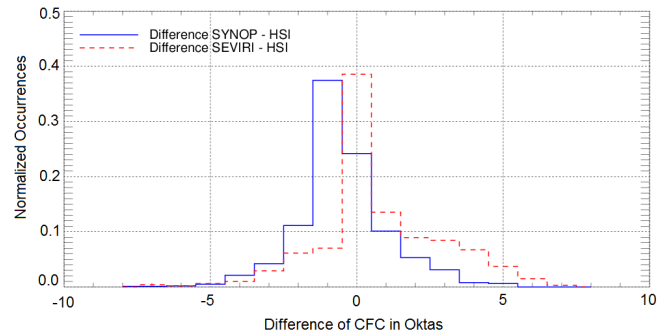
Figure 7. Histogram of instantaneous CFC differences (solid blue: SYNOP minus HSI at HAJ; dashed red: SEVIRI minus HSI at IMUK) from 1 July to 30 September 2009. Occurrences are normalized with their total count. Negative (positive) differences indicate that SEVIRI or SYNOP underestimate (overestimate) CFC compared to HSI.

a CFC of 8 oktas while SYNOP estimates a CFC of 7 oktas. That is an 1 okta underestimation by SYNOP. Figure 6 also reveals that in a total of 117 cases SYNOP estimates a CFC between 1 and 7 oktas while SEVIRI measures a CFC of 0 okta; in alone 85 of these cases SYNOP estimates a CFC between 1 and 2 oktas, which shows that for small cloud coverages SYNOP tends to overestimate CFC with respect to SEVIRI. In total around 70 % of the CFCs are underestimated by SYNOP by at least 1 okta whereas only 17 % are overestimated. This tendency of underestimation by SYNOP is confirmed by the negative bias of −15 % with a standard deviation equal to 26 % with respect to SEVIRI.

Also with respect to HSI, SYNOP tends to underestimate CFC by 1 okta. This observation is confirmed in the histogram of Fig. 7. In 222 cases SYNOP estimates a CFC of 7 oktas while HSI measures a CFC of 8 oktas. Overall SYNOP underestimates 56 % of all 996 match-ups, while only 20 % are overestimated. With a bias equal to −6 % and a standard deviation equal to 19 % we conclude that differences between HSI and SYNOP are smaller than differences between SEVIRI and SYNOP.

These observations are also shown in the results of PODs and FARs. In Table 5 we show that the probability of SEVIRI detecting cloudy sky with respect to HSI is 97 % and represents the highest POD for all three scenarios and is 30 % higher than SEVIRI-SYNOP POD and 16 % higher than HSI-SYNOP POD. However, in the case of broken-cloud coverage, HSI-SEVIRI has the lowest POD of 15 % in contrast to HSI-SYNOP POD of 52 %.

A 15 % POD for cloud-contaminated skies shows that SEVIRI only detects 15 % of the broken-cloud cases by HSI. The SEVIRI-SYNOP and HSI-SYNOP PODs for the same scenario are 44 and 52 %, respectively. However, the FAR by SEVIRI-SYNOP is 89 % which indicates that the amount of false alarms is very large for the broken-cloud scenario.

Table 5. Table of all results of the comparisons between instantaneous SEVIRI, SYNOP and HSI CFC in Hannover, Germany. The results are presented as POD_{clear}, POD_{broken}, POD_{cloudy}, FAR_{clear}, FAR_{broken} and FAR_{cloudy} (clear: cloud-free, broken: cloud-contaminated, cloudy: cloud-covered).

	POD_{clear}	POD_{broken}	POD_{cloudy}	FAR_{clear}	FAR_{broken}	FAR_{cloudy}
HSI-SYNOP	72 %	52 %	81 %	36 %	48 %	16 %
HSI-SEVIRI	70 %	15 %	97 %	34 %	38 %	29 %
SEVIRI-SYNOP	69 %	44 %	67 %	35 %	89 %	4 %
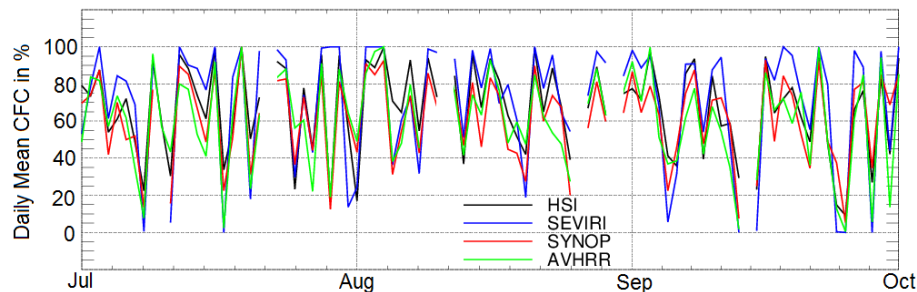
Figure 8. Daily mean CFC of HSI, SEVIRI, SYNOP and AVHRR in percent in Hannover from 1 July to 30 September 2009.

4.1.2 Daily Mean

Comparing only DM CFC by SEVIRI to HSI we find a bias of 8 % which indicates a rather small overestimation by SE-VIRI but is nevertheless the highest bias we find for all daily mean comparisons. The standard deviation is 16 % and we can conclude that these data sets do agree well and that the differences in the instantaneous data diminish to a minimum.

A comparison of the DMs of the CFC by AVHRR, SEVIRI and HSI shows that AVHRR–HSI DM CFC and SEVIRI–HSI DM CFC have almost the same maxima and minima on the same days (Fig. 8). The differences of AVHRR–HSI DM CFC are generally larger in July and smaller in September, than the differences from SEVIRI–HSI DM CFC. The SD of the difference for the DM CFC derived from the AVHRR is 14 % and the bias is 6 %. SEVIRI also shows a bias of 6 % but a slightly higher standard deviation equal to 19 %.

The differences of the DM CFC between SEVIRI and SYNOP are slightly higher than the differences between SE-VIRI and HSI. The SD of the difference between HSI and SYNOP is 11 %. The bias is −6 %. Although SEVIRI shows a higher SD (15 %), both SYNOP and SEVIRI DM CFC show a good agreement to HSI DM CFC.

All SD and biases for instantaneous and daily mean data in all different match-ups are presented in Table 6.

4.2 Lauder, New Zealand – comparison

In addition, we also performed HSI measurements in Lauder, New Zealand, and compared these to AVHRR data. The distribution of differences between instantaneous CFC in oktas from AVHRR and HSI for November and December 2009 in Lauder, New Zealand, are presented in Fig. 9. About 35 %

Table 6. Summary of standard deviations and bias for the comparisons between instantaneous and daily mean CFC in Hannover, Germany. Results show the deviations of SEVIRI, AVHRR and SYNOP to HSI. The corresponding data set has been matched to the data set in parentheses. A lack of parentheses indicates that the data set has only been matched to HSI.

	SD	Bias	DM SD	DM bias
AVHRR	22 %	2 %	14 %	6 %
AVHRR (SEVIRI)	21 %	−2 %	14 %	6 %
SEVIRI	29 %	8 %	16 %	8 %
SEVIRI (SYNOP)	27 %	9 %	15 %	10 %
SEVIRI (AVHRR)	38 %	3 %	19 %	6 %
SYNOP	19 %	−6 %	11 %	−6 %

of all HSI–AVHRR match-ups show a difference at 0 okta. When underestimating, it seems that AVHRR underestimates mostly by 1 okta whereas match-ups are decreasing exponentially with increasing difference for an overestimating AVHRR. On average, AVHRR slightly overestimates CFC as shown in the contingency table (Table 7). In contrast to 13 % of CFC underestimation, we find that in 18 % of all cases AVHRR determines a higher CFC than HSI.

We obtain the same conclusions as in the Hannover comparisons from the POD and FAR results shown in Table 4: in the cloud-covered scenario both POD and FAR show relatively good results, whereas cloud-contaminated scenarios show a low POD (26 %) and a high FAR (68 %). However, we find that in all cases (clear, broken and cloudy) the Lauder, New Zealand, PODs are lower compared to Hannover, Germany.
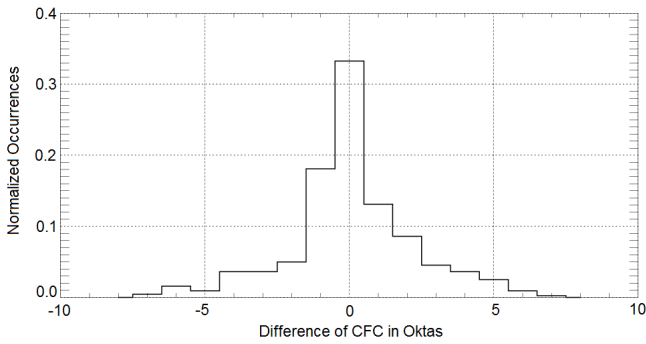
Figure 9. Histogram of instantaneous CFC differences between AVHRR and HSI in Lauder, New Zealand, from 1 November to 31 December 2009. The occurrences are normalized with their according total count. Negative differences express an underestimation of the CFC by AVHRR compared to HSI; positive differences express an overestimation.

Table 7. AVHRR in Lauder, New Zealand: contingency table comparing HSI and AVHRR CFC for the cases of cloud-free (0–2 oktas), broken clouds (3–5 oktas) and cloud-covered (6–8 oktas). Values represent ratio between match-ups and total of 442 valid measurements.

| | Scenario | HSI Lauder | | |
		Clear	Broken	Cloudy
	Clear	19 %	5 %	3 %
AVHRR	Broken	5 %	5 %	5 %
	Cloudy	4 %	9 %	45 %

The difference between DM CFC of the AVHRR and HSI in Lauder, New Zealand, states a maximum deviation of approximately ± 25 % at the beginning of November as shown in Fig. 10. During the first month it seems that the AVHRR either under- or overestimates the CFC when compared to the HSI. However, during the majority of December AVHRR is only slightly underestimating the CFC compared to the HSI. The SD between AVHRR and HSI is 15 % and the bias is equal to 5 %, which shows a good agreement between the daily mean CFC series and also agrees with the findings in the Hannover case.

4.3 SEVIRI algorithm – variation of broken-cloud layer index

We analyzed the influence of different features (e.g., averaging window size, BCLI) in the CM SAF algorithm on the resulting CFC (in 1 h match-ups). For these comparisons the original features were used to examine the original product, which is currently published by CM SAF.

Since a cloud-contaminated pixel is not completely cloud-filled by definition, a BCLI of 100 % is not accurate. Hence, there is a need to examine whether a BCLI of 100 % is the best choice. In the following analyses we vary the BCLI be-

tween 50, 75 and 100 % for CFC calculations. We then compare the different CFC time series to HSI CFC. We combine this last analysis with variable averaging window sizes (3×3, 5×5 and 7×7).

We find that changing the BCLI does only have a minimal effect on PODs and FARs and shows small changes in the differences between SEVIRI and HSI. However, changing the averaging window size has an influence on PODs and FARs, even though these results also depend on the scenario type. Figure 11 shows box-averaged PODs and FARs. The results reveal that with increasing averaging window sizes, POD and FAR are about the same for cloudy scenarios (92 %). In the case of a clear sky the POD is highest for the smallest window size and lowest for 5×5 (83 vs. 75 %). Between the window sizes of 3×3 and 7×7 we find a change in POD_{broken} from 4 to 15 % and a change in FAR_{broken} from 28 to 42 %. This means that instantaneous CFC should be treated with caution.

In Table 8 we present the standard deviations and bias for all nine different cases. At 50 and 75 % the mean bias is equal to −1 and 1 %, respectively. It seems that the window size influences the SD, which decreases with increasing window size. We suggest that a BCLI lower than 75 % should be considered for further usage in instantaneous, daily mean and monthly mean data and that the averaging window size should be decreased to 3×3. All these results are summarized in Table 8.

5 Discussion

The differences observed can be explained by three main sources of uncertainty: spatial resolution, algorithm deficiencies and viewing geometry.

5.1 Spatial resolution and algorithm issues

The comparison of the CFC from SEVIRI with HSI data showed up to 100 % deviation in instantaneous measurements. However, SEVIRI and HSI mostly agree on the CFC especially in the case of completely cloudy skies. Whereas on clear-sky days, when SEVIRI shows no amount of clouds, HSI still notices a CFC of up to 5 %. This particular deviation is due to a solar filter (in the HSI CFC algorithm), which does not entirely exclude the sun's influence (appears white in the picture). The highest deviations are shown during partially cloud-covered skies.

Both satellite instruments, SEVIRI and AVHRR, are sensitive to the same weather conditions (convective clouds, high winds and fast-changing weather conditions). Humidity in the upper atmosphere is a cause which leads to an incorrect interpretation of the cloud situation by HSI. Some cirrus clouds are too thin to be detected in the visible spectrum. Spatial resolution, limits and form of one box by AVHRR
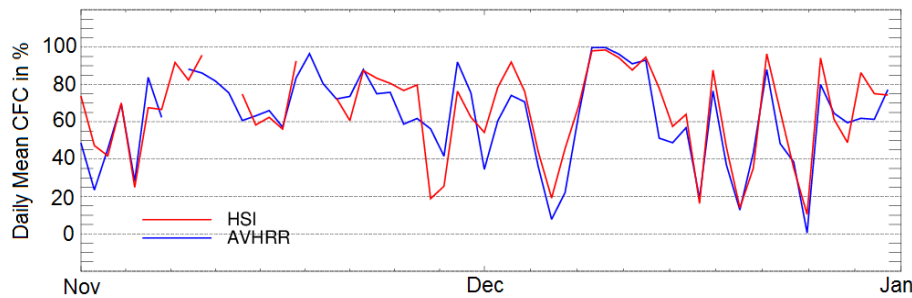
Figure 10. Daily mean CFC in percent by HSI (red) and AVHRR (blue) in Lauder, New Zealand, between 1 November and 31 December 2009.
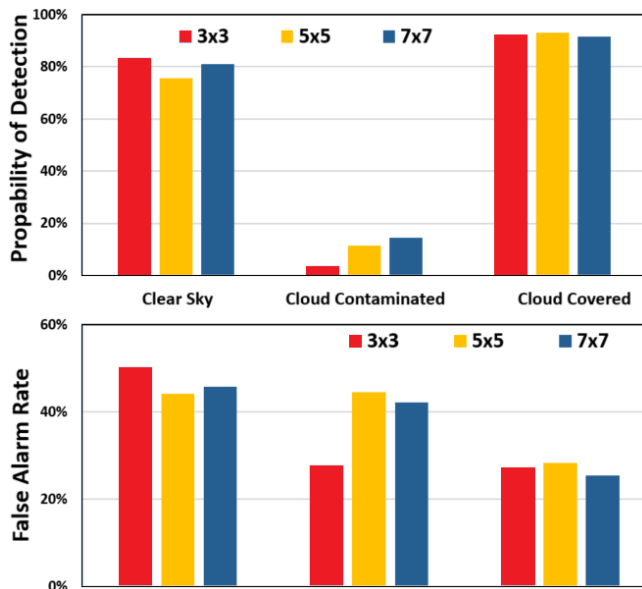
Figure 11. Bar plot of POD (up) and FAR (below). Each bar represents the average POD/FAR calculated from the three different BCLIs scenarios (BCLI = 50, 75 and 100 %). The different colors of the bars stand for the corresponding averaging window sizes (red 3×3, yellow 5×5 and blue 7×7, respectively). The bars are organized into three groups representing clear (0–2 oktas), cloud-contaminated (3–5 oktas) and cloud-covered skies (6–8 oktas).

Table 8. Standard deviation (SD) and bias (as well as mean SD and mean bias) between HSI and SEVIRI as a function of different broken-cloud layer indexes (BCLI = 50, 75 or 100 %) and averaged window sizes (3×3, 5×5 and 7×7) in Hannover, Germany, from 1 July to 30 September 2009.

BCLI	Window size	SD	Bias	Mean SD	Mean bias
	3×3	30 %	−2 %		
50 %	5×5	28 %	0 %	28 %	−1 %
	7×7	27 %	1 %		
75 %	3×3	31 %	0 %	29 %	1 %
	5×5	28 %	1 %		
	3×3	32 %	3 %		
100 %	5×5	29 %	8 %	30 %	5 %
	7×7	28 %	3 %		

which differ from the field of view and form of the HSI also cause further misinterpretation of the cloud situation.

Another effect contributing to deviations between HSI and SEVIRI/AVHRR is the occurrence of stratocumulus. These clouds have small areas where blue sky is exposed, which are detected by HSI. Because of the small scale and only slight transparency of these areas, the corresponding pixel is assigned a wrong BCLI of a 100 %. This causes a higher computed CFC. It has been shown that decreasing both the BCLI and the averaged window size can lead to improvement in the cases of cloud-free to cloud-contaminated sky. Less readily detectable cirrus clouds are the third effect contributing to these deviations. These clouds are either not detected by the HSI because of their high transparency for the blue portion of the spectrum of the sky radiance or are not detected by the satellite instrument because of the lack of cloud particles per volume of the cloud. The circumsolar area of the sun that is not perfectly analyzed by the sun filter in the HSI algorithm also contributes to the instantaneous and therefore to the DM deviation of the CFCs. As a result, this bright area is incorrectly characterized as cloud-contaminated or filled, leading to an error up to 5 %.

Another factor contributing to the lower deviations between AVHRR and HSI (compared to SEVIRI–HSI) is the originally higher resolution of $1 \times 1 \, \text{km}^2$ of AVHRR in comparison to SEVIRI with a resolution of $3 \times 3 \, \text{km}^2$.

In the case of SYNOP observations we believe that the subjective estimation by different weather observers, who are working in shifts, is one of the major factors that contributes to the deviations between SYNOP and SEVIRI CFC. Therefore, these estimations depend on the physical conditions of humans. Even trained observers tend to over- or underestimate cloud coverage (Dybbroe et al., 2004). In our case, the results show that SYNOP underestimates CFC by $6 \pm 19 \, \%$ compared to HSI and SEVIRI. This underestimation is mostly due to SYNOP's CFC definition of cloud-covered sky (only 8 oktas). Only skies that are completely

overcast (100 % CFC) will be reported as 8 oktas, whereas SEVIRI and HSI define 8 oktas as a CFC between 93.75 and 100 %. Overestimations in the lower CFC are due to the fact that SYNOP observers estimate a CFC of 1 okta once a cloud is present. These observers also consider different areas for different observations. These areas are highly dependent on visibility and topography. This also means that the scenery effect plays an important role in the observation of clouds. The differences in the data sets can also be explained by the changes in these viewing conditions.

5.2 Influence of viewing angles and geometry

Instantaneous CFC from HSI was compared to AVHRR NOAA satellite data. The AVHRR shows nearly the same deviations as SEVIRI with slightly smaller magnitudes. However, in the case of broken-cloud coverage we find far less agreement between SEVIRI and HSI than we find between AVHRR and HSI. Since AVHRR is operating on polar satellites we can conclude that the large viewing angle obviously does influence the results of SEVIRI. The parallax effect influences the quality of the data notably for broken-cloud coverage. This effect describes a geometric dislocation of high cloud layers. This effect depends on the cloud layer height and thickness and increases with increasing distance from the satellite nadir, i.e., with an increase in the oblique angle. This means that the ground-based cloud layer observation is horizontally displaced in the satellite image. SEVIRI's IMUK and HAJ pixels have a satellite zenith angle of about 60°. With a mean cloud top height lower than 3 km (excluding clear-sky days) the mean parallax displacement is at maximum 5 km. In comparison, we restricted the maximum zenith angle for AVHRR to 40°, which under the same circumstances leads to a maximum displacement of 2.5 km. Since this effect is highly dependent on the zenith angle, we see an advantage in the polar orbiting instrument measurements. In an ideal case the zenith angle decreases to 0 and there will be no parallax displacement. Whereas SEVIRI on a stationary satellite will always have the same zenith angle (for Hannover it is about 60° decreasing towards the tropics). In contrast to Greuell and Roebeling (2009), who validated SEVIRI against ground-based microwave radiometers, we believe that the parallax effect in our case study is not the major error source since HSI images cover an area of $\pm 15 \times 15$ km^2 which corresponds to the box size of SEVIRI. Ground-based microwave radiometers only have view cross sections of 90×90 and 220×220 m^2 for 2 and 5 km cloud top height, respectively (Greuell and Roebeling, 2009).

We also need to consider the scenery effect. This effect describes the overestimation of CFC caused by a slanted view at convective cloud towers, for example. The contribution of this effect increases with increasing viewing angles and therefore especially influences SEVIRI results towards the edges of the MSG disk, thus at high latitudes. However, AVHRR and the surface observations also encounter prob-

lems in correctly estimating the cloud amount in case of convective clouds (that shield the cloud-free gaps in between the individual clouds) at large viewing angles (Malberg, 1973).

6 Conclusions

We compared instantaneous and daily mean CFC derived from satellite-based instruments to ground observations by an automated camera and SYNOP data.

We find in general good agreement between satellite-derived estimates compared to HSI with biases ranging from 2 % (AVHRR) to 8 % (SEVIRI) and standard deviations of 22 % (AVHRR) and 29 % (SEVIRI) for instantaneous results. SYNOP underestimates CFC by 6 ± 19 % compared to HSI and SEVIRI. All DM CFC comparisons showed lower standard deviations than the instantaneous comparisons, which are mostly around 10 % lower. We conclude that the averaged climatology may well be used for comparison against ground-based observations. Yet in the case of broken-cloud fields (3–5 oktas) the instantaneous CFC should be treated with caution.

We find that both SEVIRI and AVHRR show good skill when detecting cloud-free and cloud-covered skies. We only find poor skill, though, whenever broken clouds occur. In the case of broken-cloud fields, major influences on performance are viewing angles, spatial resolution, the broken-cloud layer index and the averaging window size. We showed that BCLIs of 50 and 75 % show lower biases for SEVIRI CFC compared to HSI. It has been shown as well that changing the averaging window size to 7×7 leads to overall smaller deviations between SEVIRI and HSI. The largest impact on the performance of the satellite products, however, can most likely be attributed to the scenery effect (especially for SEVIRI) and the rather low spatial resolution (compared to HSI). It is worth while to remember that clouds often have a complex small-scale structure that cannot be detected well with a pixel size of 1 km^2 or more. As a consequence, in case of broken-cloud fields, gaps between clouds are not detected by the satellite instruments and therefore classified as "cloudy", which leads to an general overestimation of the actual cloud cover in these cases. However, there may be better ways to deal with this problem by systematic and long-term comparisons with all-sky camera data on the ground. The differences between SYNOP and satellite observations can partly be explained by differences of the viewing conditions. We showed that SYNOP's 1 okta underestimation of CFC is due to the definition of 8 oktas. In case of SYNOP 8 oktas are only reported for complete overcast skies (100 % CFC) whereas SEVIRI and HSI define 8 oktas as CFC between 93.75 and 100 %. Therefore, we can to a certain extend reconfirm Dürr and Philipona (2004) results and we conclude that synoptic observations of CFC are comparable to instantaneous comparison of instrument-based SEVIRI CFC for clear and overcast skies. However, strong deviations remain in the case of

broken cloudiness and we also conclude that a continuously operated all-sky camera will be better suited for comparisons to spaceborne observations.

References

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, J. Geophys. Res.-Atmos., 103, 32141–32157, 1998.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, 1989.

Amato, U., Antoniadis, A., Cuomo, V., Cutillo, L., Franzese, M., Murino, L., and Serio, C.: Statistical cloud detection from SEVIRI multispectral images, Remote Sens. Environ., 112, 750–766, 2008.

Aminou, D.: MSG's SEVIRI Instrument, ESA Bulletin, 111, 15–17, 2002.

Boers, R., de Haij, M., Wauben, W., Baltink, H. K., van Ulft, L., Savenije, M., and Long, C. N.: Optimized fractional cloudiness determination from five ground-based remote sensing techniques, J. Geophys. Res.- Atmos., 115, D24116, doi:10.1029/2010JD014661, 2010.

Calbo, J., Pages, D., and Gonzales, J.: Empirical Studies of Cloud Effects on UV Radiation: A review, Rev. Geophys., 43, 1–28, 2005.

Christodoulou, C. I., Michaelides, S. C., and Pattichis, C. S.: Multi-feature texture analysis for the classification of clouds in satellite imagery, Geoscience and Remote Sensing, IEEE Transactions on, 41, 2662–2668, 2003.

Clement, A. C., Burgman, R., and Norris, J. R.: Observational and Model Evidence of Positive Low-Level Cloud Feedback, Science, 325, 460–464, 2009.

Deneke, H., Johnston, S., Reuter, M., Roebeling, R., Tetlaff, A., Thomas, W., and Wolters, E.: Validation of CM SAF Cloud Products Derived from MSG/SEVIRI Data, Technical Report, CM SAF Deutscher Wetterdienst, available at http://www.cmsaf.eu/EN/Documentation/Documentation/AnnualValidationRep/AnnualValidationReports.html, 2007.

Derrien, M. and LeGleau, H.: MSG/SEVIRI Cloud Mask and Type from SAFNWC, Int. J. Remote Sens., 26, 4707–4732, 2005.

Derrien, M. and LeGleau, H.: Algorithm Theoretical Basis Document for Cloud Products, Technical Report, NWCSAF, 2013.

Derrien, M., Farki, B., Harang, L., LeGleau, H., Noyalet, A., Pochic, D., and Sairouni, A.: Automatic cloud detection applied to NOAA-11/AVHRR imagery, Remote Sens. Environ., 46, 246–267, 1993.

Dürr, B. and Philipona, R.: Automatic cloud amount detection by surface longwave downward radiation measurements, J. Geophys. Res.- Atmos., 109, 1–9, 2004.

Dybbroe, A., Karlsson, K. G., and Thoss, A.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling – Part II: Tuning and Validation, J. Appl. Meteor, 44, 55–71, 2004.

Dybbroe, A., Thoss, A., and Karlsson, K.-G.: NWCSAF AVHRR Cloud Detection and Analysis Using Dynamic Thresholds and Radiative Transfer Modeling – Part I: Algorithm Description, J. Appl. Meteorol., 44, 39–54, 2005.

Ebert, E.: A pattern recognition technique for distinguishing surface and cloud types in the polar regions, J. Clim. Appl. Meteorol., 26, 1412–1427, 1987.

Forster, P., Ramaswamy, V., Artaxo, P., Berntsen, T., Betts, R., Fahey, D., Haywood, J., Lean, J., Lowe, D., Myhre, G., Prinn, R., Raga, G., Schultz, M., and VanDorland, R.: Changes in Atmospheric Constituent and in Radiative Forcing. Chapter 2, In: Climate Change 2007, The Physical Science Basis, 2007.

Gao, B.-C. and Wiscombe, W.: Surface-Induced Brightness Temperature Variations and Their Effects on Detecting Thin Cirrus Clouds Using IR Emission Channels in the 8-12 micron Region., J. Appl. Meteorol., 33, 568–572, 1994.

Garand, L.: Automated recognition of oceanic cloud patterns. Part I: Methodology and application to cloud climatology, J. Climate, 1, 20–39, 1988.

Greuell, W. and Roebeling, R.: Toward a standard procedure for validation of satellite-derived cloud liquid water path: A study with SEVIRI data, J. Appl. Meteorol. Climatol., 48, 1575–1590, 2009.

Karlsson, G. K., Wolters, E., Albert, P., Tetzlaff, A., Roebeling, R., Thomas, W., and Johnston, S.: Validation of CM SAF Cloud Products Using MSG/SEVIRI, Technical Report, CM SAF Deutscher Wetterdienst, available at http://www.cmsaf.eu/EN/Documentation/Documentation/AnnualValidationRep/AnnualValidationReports.html, 2005.

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

Kazantzidis, A., Tzoumanikas, P., Bais, A., Fotopoulos, S., and Economou, G.: Cloud detection and classification with the use of whole-sky ground-based images, Atmos. Res., 113, 80–88, 2012.

Kidder, S. and Von der Haar, T.: Satellite Meteorology: An Introduction, Gulf Professional Publishing, 1995.

Kogan, F., Vargas, M., and Guo, W.: Comparison of AVHRR-Based Global Data Records, in Use of Satellite and In-Situ Data to Improve Sustainability, Springer Netherlands, 267-272, 2011.

Liou, K.: Atmospheric radiation: Causes and Effects, in Encyclopedia of Earth System Science, Academic Press, 261-272, 1991.

Malberg, H.: Comparison of mean cloud cover obtained by satellite photographs and ground-based observations over Europe and the Atlantic, Mon. Weather Rev., 101, 893–897, 1973.

Mannstein, H., Brömser, A., and Bugliaro, L.: Ground-based observations for the validation of contrails and cirrus detection in satel-

lite imagery, Atmos. Meas. Tech., 3, 655–669, doi:10.5194/amt-3-655-2010, 2010.

Martinez-Chico, M., Batlles, F., and Bosch, J.: Cloud classification in a mediterranean location using radiation data and sky images, Energy, 36, 4055–4062, 2011.

Ohring, G., Wielicki, B., Spencer, R., Emery, B., and Datla, R.: Satellite Instrument Calibration for Measuring Global Climate Change, B. Am. Meteorol. Soc., 86, 1303–1313, 2005.

Orsini, A., Tomasi, C., Calzolari, F., Nardino, M., Cacciari, A., and Georgiadis, T.: Cloud cover classification through simultaneous ground-based measurements of solar and infrared radiation, Atmos. Res., 61, 251–275, 2002.

Parikh, J.: A comparative study of cloud classification techniques, Remote Sens. Environ., 6, 67–81, 1977.

Poetzsch-Heffter, C., Liu, Q., Ruperecht, E., and Simmer, C.: Effect of Cloud Types on the Earth Radiation Budget Calculated with the ISCCP C1 Dataset: Methodology and Initial Results, J. Climate, 8, 829–843, 1995.

Porcú, F. and Levizzani, V.: Cloud classification using METEOSAT VIS-IR imagery, Int. J. Remote Sens., 13, 893–909, 1992.

Reuter, M., Thomas, W., Albert, P., Lockhoff, M., and Weber, R.: The CM SAF and FUB Cloud Detection Schemes for SE-VIRI: Validation with Synoptic Data and Initial Comparison with MODIS and CALIPSO, J. Appl. Met. Clim., 48, 301–316, 2009.

Robaa, S.: Evaluation of sunshine duration from cloud data in Egypt, Energy, 33, 785–795, 2008.

Romano, F., Cimini, D., Rizzi, R., and Cuomo, V.: Multilayered cloud parameters retrievals from combined infrared and microwave satellite observations, J. Geophys. Res., 112, D08210, doi:10.1029/2006JD007745, 2007.

Saunders, R. and Kriebel, K.: An improved method for detecting clear sky and cloudy radiances from AVHRR data, Int. J. Remote Sens., 9, 123–150, 1988.

Schade, N. H., Macke, A., Sandmann, H., and Stick, C.: Total and partial cloud amount detection during summer 2005 at Westerland (Sylt, Germany), Atmos. Chem. Phys., 9, 1143–1150, doi:10.5194/acp-9-1143-2009, 2009.

Schafer, J., Saxena, V., Wenny, B., Barnard, W., and DeLuisi, J.: Observed Influence of Clouds on Ultraviolet-B Radiation, Geophys. Res. Lett., 23, 2625–2628, 2012.

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., and Ratier, A.: An Introduction to Meteosat Second Generation (MSG), Am. Meteorol. Soc., 83, 977–992, 2002.

Schröder, M., Bennartz, R., Schüller, L., Preusker, R., Albert, P., and Fischer, J.: Generating cloudmasks in spatial high-resolution observations of clouds using texture and radiance information, Int. J. Remote Sens., 23, 4247–4261, 2002.

Schutgens, N. and Roebeling, R.: Validating the validation: the influence of liquid water distribution in clouds on the intercomparison of satellite and surface observations, J. Atmos. Ocean. Technol., 26, 1457–1474, 2009.

Seinfeld, J. and Pandis, S.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, Wiley & Sons, New York, USA, 1998.

Solomon, S., Qun, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., and Tignor, M. (Eds.): Climate Change 2007: the Physical Science Basis: Contribution of Worki Group I to the Forth Assessment Report of the Intergovernmental Panel of Climate Change, Cambridge University Press, Cambridge, UK, 2007.

Stengel, M., Kniffka, A., Meirink, J. F., Lockhoff, M., Tan, J., and Hollmann, R.: CLAAS: the CM SAF cloud property data set using SEVIRI, Atmos. Chem. Phys., 14, 4297–4311, doi:10.5194/acp-14-4297-2014, 2014.

Tohsing, K., Schrempf, M., Riechelmann, S., Schilke, H., and Seckmeyer, G.: Measuring High-Resolution Sky Luminance Distributions with a CCD Camera, Appl. Opt., 52, 1564–1573, 2013.

Welch, R., Sengupta, S., Goroch, A., Rabindra, P., Rangaraj, N., and Navar, M.: Polar cloud and surface classification using AVHRR imagery: An intercomparison of methods, J. Appl. Meteorol., 31, 405–420, 1992.

Yamashita, M., Yoshimura, M., and Nakashizuka, T.: Cloud Cover Estimation Using Multitemporal Hemisphere Imageries, International Archives of Photogrammetry, Remote Sens. Spat. Inf., 35, 826–829, 2004.