

**Algorithm Theoretical Basis Document**

**CM SAF Cloud, Albedo, Radiation data record,**

**AVHRR-based, Edition 2.1 (CLARA-A2.1)**

**Probabilistic cloud masks – CMA-prob**

**A complementary product to level-2/level-2b versions of**

[DOI: 10.5676/EUM\\_SAF\\_CM/CLARA\\_AVHRR/V002\\_01](https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V002_01)

Fractional Cloud Cover

CM-11015

## Document Signature Table

	Name	Function	Signature	Date
<b>Author</b>	Karl-Göran Karlsson	CM SAF scientist		15/10/2020
<b>Editor</b>	Rainer Hollmann	Science Coordinator		15/10/2020
<b>Approval</b>	Rainer Hollmann	Science Coordinator		15/10/2020
<b>Release</b>	Martin Werscheck	CM SAF Manager		16/10/2020


## Distribution List

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

External Distribution		
Company	Name	No. Copies
Public		1

## Document Change Record

Issue/ Revision	Date	DCN No.	Changed Pages/Paragraphs
1.0	24/04/2015	SAF/CM/SMHI/ATBD/GAC/CMA-prob	Initial version, submitted for PCR 2.2
1.1	27/05/2016	SAF/CM/SMHI/ATBD/GAC/PBCM	Version submitted for DRR 2.2
1.2	19/08/2016	SAF/CM/SMHI/ATBD/GAC/PBCM	Close-out version with DRR 2.2 RIDs implemented
1.3	15/10/2020	SAF/CM/SMHI/ATBD/GAC/PBCM	Editorial changes

	<b>Algorithm Theoretical Basis Document</b> <b>CLARA Edition 2</b> <b>Probabilistic cloud masks CMA-prob</b>	Doc. No: SAF/CM/SMH/ATBD/GAC/PBCM Issue: 1.3 Date: 15.10.2020
-----------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------

### Applicable Documents

Reference	Title	Code / Date
AD 1	CM SAF Product Requirement Document	SAF/CM/DWD/PRD/2.7

### Reference Documents

Reference	Title	Code
RD 1	Algorithm Theoretical Basis Document SAFNWC/PPS “Cloud mask”, PPS version 2014 patch 20150327 (OBS: in the text referred to as patch 1).	SAF/NWC/CDOP2/PPS/SCI/ATBD/1, Issue 1.1, 13 March 2015
RD 2	Algorithm Theoretical Basis Document CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 2 (CLARA-A2) Cloud Products processing chain (level-1 – level-2/2b – level-3).	SAF/CM/DWD/ATBD/GAC/CLD/ Issue 2.3

## Table of Contents

1	The EUMETSAT SAF on Climate Monitoring.....	6
2	Introduction .....	8
3	Theoretical description of the CMA-prob method.....	9
3.1	Background – problems with traditional cloud masking and suggested new approaches 9	
3.2	Bayesian theory .....	10
3.3	The CMA-prob Naïve Bayesian approach .....	10
3.4	Estimating conditional cloud probabilities from CALIPSO measurements.....	11
3.5	Definition of a basic sub-set of constrained AVHRR image features.....	13
3.6	Training the classifier using CALIPSO-CALIOP cloud data.....	17
3.7	Achieved probability distributions.....	19
4	First results.....	23
4.1	Product demonstration.....	23
4.2	Preliminary validation results .....	25
5	Limitations and areas for future improvements .....	28
6	References.....	29
7	Glossary.....	31

## List of Tables

Table 3-1: Spectral channels of the Advanced Very High Resolution Radiometer (AVHRR). Three different versions of the instrument are described as well as corresponding satellites. ....	15
Table 3-2: Used transformed AVHRR image features for daytime probabilistic cloud masking. ....	15
Table 3-3: Used transformed AVHRR image features for night-time probabilistic cloud masking. ....	16
Table 3-4: Geographical regions used when training the probabilistic classifiers. ....	17

## List of Figures

Figure 3-1: Cloud probabilities estimated from global CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR 0.6 $\mu\text{m}$ visible reflectances (denoted $R_{\text{vis}}$ ) over Low Latitude ocean surfaces (left - defined in text) and over High Latitude mountain areas (right - defined in text). ....	12
Figure 3-2: Cloud probabilities estimated from CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR temperature differences (denoted $T_{\text{diff}}$ ) between AVHRR 11 $\mu\text{m}$ brightness temperatures and the ERA-Interim (Dee et al., 2011) surface skin reference temperature over Low Latitude ocean surfaces during day (left) and over High Latitude mountain areas during night (right). ....	13
Figure 3-3: Cloud probabilities estimated from CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR temperature differences between AVHRR channel 4 and 5 (denoted $T_{\text{diff}}$ in the plots) over Low Latitude ocean surfaces during night. Left panel shows results in original form and right panel if plotting results as a function of temperature differences related to PPS thresholds (consisting of dynamic threshold plus a tuning offset value). ....	14
Figure 3-4: Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the $R_{\text{swir\_3b}}$ feature (see Table 3-2) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009. ....	19
Figure 3-5: Frequency of cloud occurrence over High Latitude Ice-free ocean (left) and over Polar Snow-covered surfaces (right) as a function of the $R_{\text{nir\_3a}}$ feature (see Table 3-2) derived from 385 globally matched NOAA-17/CALIPSO orbits in the period 2006-2009. ....	20
Figure 3-6: Frequency of cloud occurrence over Tropical Ocean as a function of the texture feature (see Table 3-2 and Table 3-3) during day (left) and night (right) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009. ....	21
Figure 3-7: Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the $T_{\text{diff}}$ feature (see Table 3-3) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009. ....	21
Figure 3-8: Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the $T_{\text{diff}}$ feature (see Table 3-3) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009. ....	22
Figure 4-1: Part of an original NOAA-18 AVHRR GAC scene in satellite projection over the North American west coast (with Gulf of California and Baja California in the center) registered in ascending mode (i.e., North is down, South is up) from 26 January 2010. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100%). ....	23

Figure 4-2: Part of an original NOAA-18 AVHRR GAC scene in satellite projection over Spain and northern Africa registered in ascending mode (i.e., North is down, South is up) from 16 May 2007. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100 %). .....24

Figure 4-3: Part of an original NOAA-17 AVHRR GAC scene in satellite projection over Greenland registered in descending mode (i.e., North is up, South is down) from 4 June 2009. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100 %). .....25

Figure 4-4: Hitrate and Kuipers Skill Scores plotted as a function of filtered CALIPSO-CALIOP cloud optical thickness (explained in text). Results for CMA-prob are compared to results of the official PPS version 2014 Patch 1 ([RD 1]). All results were derived for 99 NOAA-18 orbits in the period 2006-2009. ....26

Figure 4-5: Hitrate and Kuipers Skill Scores plotted as a function of filtered CALIPSO-CALIOP cloud optical thickness (explained in text). Results for CMA-prob are compared to results of the official PPS version 2014 Patch 1 ([RD 1]). All results were derived for 385 NOAA-17 orbits in the period 2006-2009. ....27

## 1 The EUMETSAT SAF on Climate Monitoring

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


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

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

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

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

- Application of highest standards and guidelines as lined out by GCOS for the satellite data processing,
- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,

	<p align="center"><b>Algorithm Theoretical Basis Document CLARA Edition 2 Probabilistic cloud masks CMA-prob</b></p>	<p>Doc. No: SAF/CM/SMH/ATBD/GAC/PBCM Issue: 1.3 Date: 15.10.2020</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------


- Intensive validation and improvement of the CM SAF climate data records,
- Taking a major role in data record assessments performed by research organisations such as WCRP. This role provides the CM SAF with deep contacts to research organizations that form a substantial user group for the CM SAF CDRs,
- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

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

## 2 Introduction

This CM SAF Algorithm Theoretical Basis Document (ATBD) describes a new probabilistic cloud masking product - denoted CMA-prob – which has been developed by CM SAF during the CDOP-2 phase. It is based on Bayesian theory and it is complementary to the SAFNWC PPS cloud mask which was used when defining the CM SAF CLARA-A1 data record and which is planned for use also for the CLARA-A2 data records (i.e., the Fractional Cloud Cover product CM-11011). The idea is that on a longer term (beyond CLARA-A2) this new probabilistic cloud mask will replace the current one in order to improve the error characterisation of cloud masking and its influence on downstream cloud, surface radiation and surface albedo products. For CLARA-A2 only a demonstration product is intended for provision to users for preliminary evaluation.

The ATBD follows largely the detailed description of the method published by Karlsson et al. (2015).

	<b>Algorithm Theoretical Basis Document CLARA Edition 2 Probabilistic cloud masks CMA-prob</b>	Doc. No: SAF/CM/SMH/ATBD/GAC/PBCM Issue: 1.3 Date: 15.10.2020
-----------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------

### 3 Theoretical description of the CMA-prob method

#### 3.1 Background – problems with traditional cloud masking and suggested new approaches

For many years the definition of fixed cloud masks or cloud masks with a small set of quality flags has been the most common way of solving the cloud screening problem in applications based on passive multispectral satellite imagery. Many examples of this exist in the literature, e.g., Dybbroe et al., (2005a, 2005b), Kriebel et al. (2003), Derrien and LeGleau (2005), Frey et al. (2008) and Pavolonis et al. (2005). The use of a fixed cloud mask is straightforward for downstream applications (e.g. for Sea Surface Temperature (SST), surface albedo, clear sky radiance and NDVI vegetation index retrievals) meaning that all cloudy pixels should be discarded in the retrieval of the actual parameter. However, the drawback is that no or very limited information about the uncertainty in the cloud screening is generally available. Consequently, also for parameters derived downstream the error characteristics are generally unknown even if specific internal parameter-specific algorithm uncertainties may be known. Furthermore, various cloud masks have generally been defined aiming for different purposes and applications. Consequently, the performance may vary considerably from method to method regarding whether the cloud screening is executed in a clear conservative way (i.e., defining clear pixels with high confidence) or in a cloud conservative way (i.e., defining cloudy pixels with high confidence). The desire to instead define a more flexible cloud mask, suitable for any (or at least most) downstream applications, has become increasingly important recently. Such a cloud mask can either be expressed as a cloud index (as suggested by Khlopenkov and Thrishchenko, 2007) or a cloud probability (Merchant et al., 2005) meaning that any user should be able to define the most suitable mode of operation. In other words, it could be used anywhere in the range from the clear conservative mode to the cloud conservative mode by just changing the tolerance level of the required cloud probabilities.

Although statistical and probabilistic (Bayesian) theories have been well established since decades (or even centuries) a problem has been to find a good observational reference to represent the true global cloud occurrence from which a firm statistical cloud distribution database can be built. However, with the 2006 launch of the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite, the situation has improved considerably. CALIOP offers global cloud observations with higher detection sensitivity than any other passive instrument (Winker et al., 2009). Furthermore, observations can be matched simultaneously in time (however, restricted to certain conditions) to observations by current operational AVHRR sensors. This has triggered numerous studies examining AVHRR-based cloud detection methods in detail (e.g., Karlsson and Dybbroe, 2010, Karlsson and Johansson, 2013 and Stengel et al., 2014). It has also paved the way for more systematic attempts to provide cloud probabilities rather than fixed cloud masks (Heidinger et al., 2012, and Musial et al., 2014). The CMA-prob development in the CM SAF is another example of this.

### 3.2 Bayesian theory

Let us first recapitulate some fundamentals of the probabilistic statistical theory. The theory is based on the pioneering work by Thomas Bayes who already in 1763 formulated his famous theorem (nowadays referred to as Bayes' Theorem) for estimation the posteriori probability of an event as a function of likelihoods (conditional probabilities) and a priori probabilities of other events. In the context of analysis of radiance feature vectors measured by satellite sensors we may express Bayes' Theorem as follows after introducing a number of definitions. If  $\mathbf{F}$  is a vector of satellite radiances or image features (e.g., brightness temperature differences or reflectances) we may denote the posteriori conditional probability that it is cloudy when  $\mathbf{F}$  is given as  $P(\text{cloudy}|\mathbf{F})$ . In the same sense we may denote the conditional probability that vector  $\mathbf{F}$  occurs given it is cloudy as  $P(\mathbf{F}|\text{cloudy})$ . If also introducing the overall probability (climatological mean) that it is cloudy as  $\overline{P(\text{cloudy})}$  and the overall probability that any given value of  $\mathbf{F}$  occurs as  $P(\mathbf{F})$  we may write Bayes' Theorem as follows:

$$P(\text{cloudy}|\mathbf{F}) = \frac{\overline{P(\text{cloudy})}P(\mathbf{F}|\text{cloudy})}{P(\mathbf{F})} \quad (1)$$

Despite its simple form, the solution of Eq. 1 is not easy to find in a situation with multispectral measurements (i.e., when the dimension of  $\mathbf{F}$  is large). The estimation of parameters in the right-hand side of Eq. 1 (especially  $P(\mathbf{F}|\text{cloudy})$ ) becomes increasingly difficult the more image features that are chosen. It then requires extraction of very large statistical training datasets to really describe the dependence on individual image features and, in addition, also the effect of their mutual correlation. What complicates things even further is that, even with one specific realisation of feature vector  $\mathbf{F}$ , probabilities may differ depending on different environmental situations (e.g. if the pixel measurement is made in winter or in summer, over land or over ocean, in mountainous terrain or over desert, etc.). Thus, the training process needs to consider additional ancillary information for a correct description of environmental conditions.

To reduce complexity of the problem some approximations may be utilised. One way to go could be the entirely empirical approach of estimating  $P(\text{cloudy}|\mathbf{F})$  directly from predefined Lookup Tables composed during training with some stratification based on ancillary data. Such a method has been demonstrated by Musial et al. (2014). Alternatively, some simplifications and approximations could be made to Eq. 1. One such simplification is denoted *The Naïve Bayesian approach* and this is used for the CMA-prob method.

### 3.3 The CMA-prob Naïve Bayesian approach

If assuming that individual image feature components  $f_i$  in  $\mathbf{F}$  are all independent (i.e., image features are uncorrelated), individual probabilities may now be multiplied to get the total probability, following the fundamental statistical rule for "Compound Probability of Independent Events". Thus, Eq. 1 reduces to

$$P(\text{cloudy}|\mathbf{F}) = \frac{\overline{P(\text{cloudy})} \prod_i P(f_i|\text{cloudy})}{P(\mathbf{F})} \quad (2)$$

This approximation of Bayes' Theorem is denoted **the Naïve Bayesian approximation**.

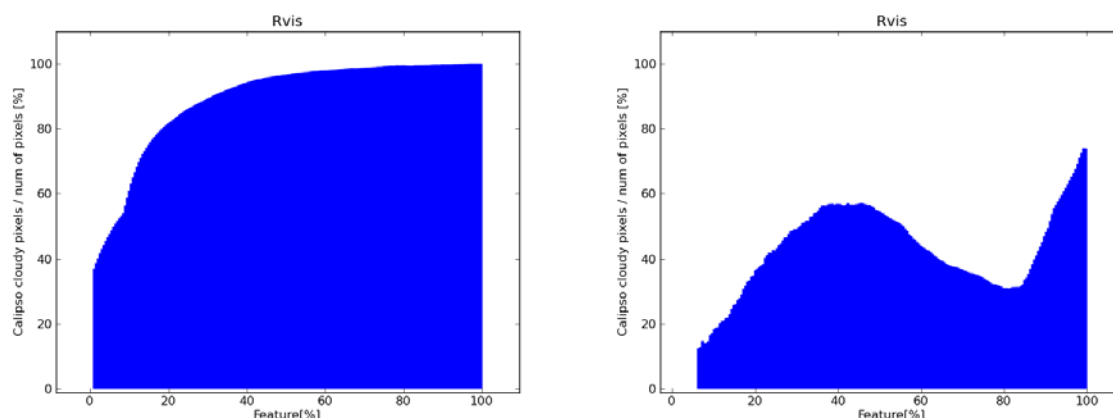
The problem has now been reduced to estimating individual probabilities  $P(f_i|cloudy)$  and then simply multiplying them. However, it must be emphasized that even if we have achieved a very simple equation for calculation of the probabilities, the big scientific challenge lies in defining and estimating the conditional probabilities in the right-hand side of the equation. This includes the very fundamental choice of appropriate image features  $f_i$ . These must be chosen in an optimal way with all of them having documented capabilities to provide essential information about cloud occurrence. The following sub-sections will describe the methodology used and the choice of optimal image feature components  $f_i$ .

We notice also that there must be a mutual inter-dependence between  $P(f_i|cloudy)$  and  $P(cloudy|f_i)$ . More clearly, if knowing the conditional probability that it is cloudy given a certain image feature value, we can also calculate it the other way around from the same statistical training dataset (provided that both absolute and relative frequencies of cloud occurrences are stored). This fact will be utilised when defining the method. Remaining factors on the right-hand side of Eq. 2 may also be calculated from training data. An estimation of the mean cloud occurrence  $P(\overline{cloudy})$  is possible and the factor  $P(F)$  may be estimated by summing contributions from both cloudy and clear cases and then compute the overall frequency that any particular realisation of vector  $F$  occurs.

The Naïve Bayesian approximation has been successfully applied to many scientific applications (e.g., Kossin and Sitkowski, 2009) and it has also been applied to the AVHRR cloud screening problem (Heidinger et al., 2012). The main difference between CMA-prob and the latter method lies in the choice of image features and the used ancillary information.

### 3.4 Estimating conditional cloud probabilities from CALIPSO measurements

If having access to a system that can match and co-locate CALIPSO/CALIOP and NOAA/METOP AVHRR measurements it is relatively straight-forward and simple to estimate conditional cloud probabilities, i.e. the frequency that it is cloudy at a certain image feature value which is formally expressed as  $P(cloudy|f_i)$ . However, as previously mentioned, this should only be done by using some restrictions on e.g. illumination conditions and the geographical coverage in order to avoid too broad distributions and distributions with a limited dynamical range of probabilities. If not doing this, the final ability to separate cloudy from cloud-free radiances would be reduced (i.e., too often give cloud probabilities close to 50 %).

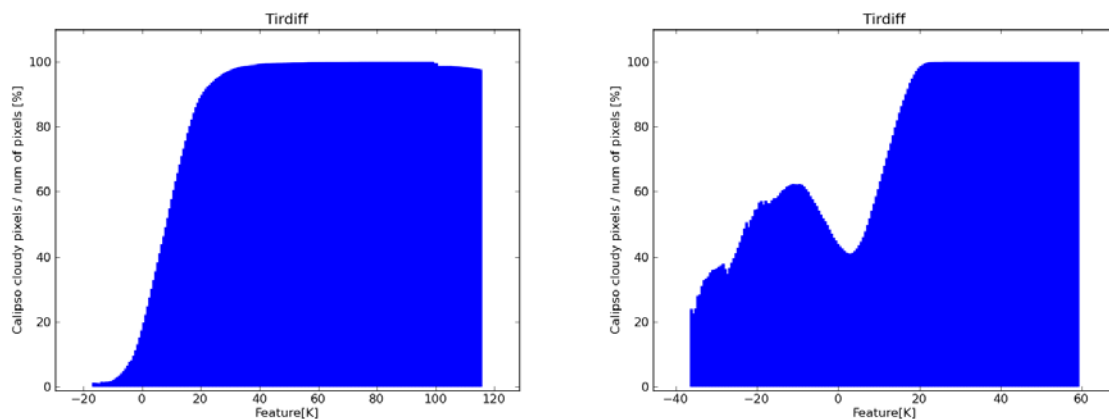


**Figure 3-1:** Cloud probabilities estimated from global CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR 0.6  $\mu\text{m}$  visible reflectances (denoted  $R_{\text{vis}}$ ) over Low Latitude ocean surfaces (left - defined in text) and over High Latitude mountain areas (right - defined in text).

Figure 3-1 shows the estimated cloud probabilities as a function of the AVHRR visible reflectances of the 0.6  $\mu\text{m}$  channel over low-latitude ocean surfaces (left panel) and over high-latitude snow-covered mountain surfaces (right panel). The distinction between low- and high latitudes is made at  $\pm 45$  degrees latitude and mountainous terrain is defined as areas with topography above 500 m. Information on snow-cover is taken from the National Snow and Ice Data Center (NSIDC) dataset provided with the CALIPSO-CALIOP cloud product.

From Figure 3-1 we conclude that cloud probabilities increase rapidly with reflectance over a very dark surface such as the ice-free ocean surface. Probabilities exceed 50 % already at a very low reflectance value (at approximately 6 % reflectance) and reach the 80 % level at approximately 18 % reflectance. Thus, conditions for cloud-screening appear almost ideal. This is not the case for the second situation in Figure 3-1 (right) showing conditions over snow-covered ground in mountainous regions. Here, we hardly see any reflectance value where cloud probability exceeds 50 % (which would be needed for this image feature to be useful for cloud screening purposes). This occurs only for moderately high reflectances close to 40 % and for very high reflectances (approaching 100 %). For the inter-mediate region of high reflectances probabilities are actually rather low which mainly is explained by the effect of non-isotropic reflection at very high solar zenith angles caused by illuminated snow-covered mountain sides.

A similar situation is seen over the same Earth surfaces in Figure 3-2 for the infrared brightness temperature difference of the 11  $\mu\text{m}$  channel with regard to the surface skin temperature. Very good separability conditions are seen over low latitude ocean surfaces while they are very problematic over mountainous terrain. Notice in particular the effect of near-surface temperature inversions over mountainous terrain leading to a specific peak in cloud probability (although just slightly exceeding 50 %) for negative values of the temperature difference (i.e., showing that clouds may then frequently be warmer than the surface temperature).

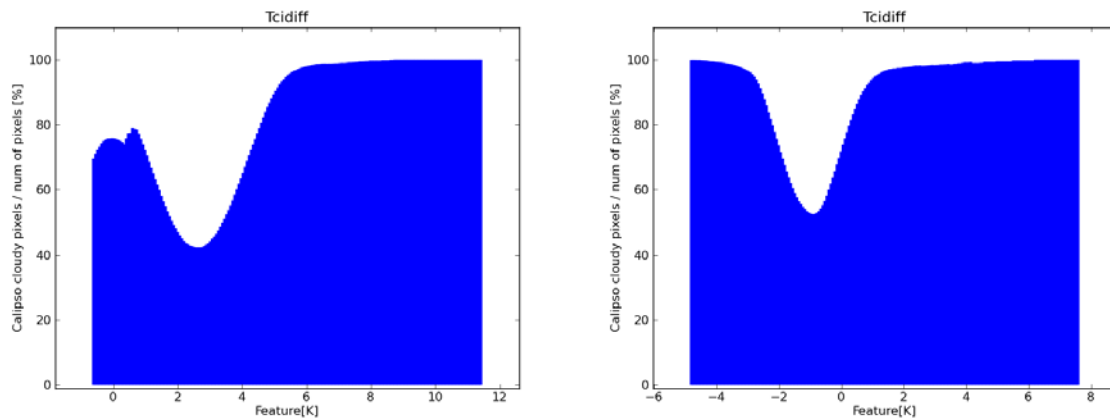


**Figure 3-2:** Cloud probabilities estimated from CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR temperature differences (denoted Tirdiff) between AVHRR 11  $\mu\text{m}$  brightness temperatures and the ERA-Interim (Dee et al., 2011) surface skin reference temperature over Low Latitude ocean surfaces during day (left) and over High Latitude mountain areas during night (right).

We conclude from Figure 3-1 and Figure 3-2 that conditions for efficient cloud screening may be drastically different depending on the geographic location and the prevailing illumination conditions (i.e., if it is day or night). This is one of the explanations for the very successful performance of simple bi-spectral VIS-IR cloud screening methods at low- to moderate latitudes (best exemplified by the results derived mainly from geostationary satellite data of the International Satellite Cloud Climatological Project – ISCCP – see Rossow et al., 1999). On the other hand, it also clearly illustrates the serious limitations for the same methods over high latitudes and over the Polar Regions.

### 3.5 Definition of a basic sub-set of constrained AVHRR image features

The Naïve Bayesian CMA-prob method utilises estimated conditional cloud probabilities (introduced in the previous section) for a sub-set of image features. However, rather than to define them in their purest form (as illustrated in Figure 3-1 and Figure 3-2) we have chosen to define them linked to pre-calculated dynamic image feature thresholds used by the Polar Platform System cloud software package (PPS, see Dybbroe et al, 2005a, 2005b) and in this particular case version PPS 2014 patch 1 [RD 1]. The reason for linking image features to pre-calculated thresholds is that the latter have been defined in a way that takes a wide range of environmental conditions into account (see Dybbroe et al., 2005a for more details). This concerns image feature variability due to the following factors: Solar and satellite geometry (direct angular dependence and dependence on scattering angles), prevailing atmospheric profiles of temperature and humidity, climatological ozone and aerosol amounts, topography and land cover and spectral surface emissivities. If not taking all these factors into account when training the probabilistic classifier, results would risk being imprecise and most likely misleading under certain conditions or at certain geographic locations. We claim that it is better to piggy-back ride on existing prepared threshold information, composed from knowledge built over many years of experience of cloud thresholding, than to try to train a classifier from scratch. The latter would require the creation of very large dimension Look-up Tables of statistical relations of cloudiness and image features and their respective dependencies on a wide range of environmental factors.



**Figure 3-3:** Cloud probabilities estimated from CALIPSO-CALIOP cloud data in the period 2006-2009 as a function of AVHRR temperature differences between AVHRR channel 4 and 5 (denoted Feature in the plots) over Low Latitude ocean surfaces during night. Left panel shows results in original form and right panel if plotting results as a function of temperature differences related to PPS thresholds (consisting of dynamic threshold plus a tuning offset value).

To illustrate the usefulness of this concept we consider one of the most commonly used AVHRR image features for detecting thin cirrus clouds (originally suggested by Inoue, 1987): The brightness temperature difference between AVHRR channels 4 and 5 at 11  $\mu\text{m}$  and 12  $\mu\text{m}$ , respectively. The main principle used for Cirrus detection is normally that the cloud transmissivity for thin ice clouds is higher in AVHRR channel 4 than in AVHRR channel 5, thus creating a positive brightness temperature difference between AVHRR channels 4 and 5. Figure 3-3 shows cloud probabilities as a function of this temperature difference (left part of Figure 3-3) but also as a function of the temperature difference relative to the corresponding PPS threshold (right part of Figure 3-3).

We notice that in its original form (left panel of Figure 3-3) we have two peaks in cloud occurrence where one is for differences close to zero K and the other for values exceeding approximately 4 K. The area between the peaks spans an interval of almost 4 K where cloud probabilities to a large extent are lower than 50 %. In the alternative formulation (Figure 3-3, right panel) results are much more distinctly organised and the range of probability values have been enlarged (which is favourable for the probabilistic classification process). The latter circumstance is especially true for the leftmost part of the distribution. We may interpret this as primarily an effect of being able to take into account the natural cloud-free contribution from atmospheric water vapour emission in the split-window channels. This emission is also able to create a discernible temperature difference in the absence of cirrus clouds explaining the broader and less decisive probability distribution in its original form for temperature differences below approximately 4 K. Resulting distributions after the coordinate transformation now clearly separates thin cirrus clouds to the right in the plot from the opaque clouds in the left part of the plot with cloud-free cases now concentrated around the transformed value of around  $-1$  K. The fact that this value is not 0 K might indicate that the currently used PPS threshold is not optimal (at least, if taking the currently used CALIPSO dataset as reference). However, this is of no importance here since the correct (CALIPSO-derived) distribution relative to the possibly biased PPS threshold will be the one used anyway.

**Table 3-1:** Spectral channels of the Advanced Very High Resolution Radiometer (AVHRR). Three different versions of the instrument are described as well as corresponding satellites.

Channel Number	Wavelength ( $\mu\text{m}$ ) <i>AVHRR/1</i> <i>Tiros-N,</i> <i>NOAA-6, 8, 10</i>	Wavelength ( $\mu\text{m}$ ) <i>AVHRR/2</i> <i>NOAA-7, 9, 11, 12, 14</i>	Wavelength ( $\mu\text{m}$ ) <i>AVHRR/3</i> <i>NOAA-15, 16, 17, 18</i> <i>NOAA-19, Metop-A</i> <i>Metop-B</i>
1	0.58-0.68	0.58-0.68	0.58-0.68
2	0.725-1.10	0.725-1.10	0.725-1.10
3A	-	-	1.58-1.64
3B	3.55-3.93	3.55-3.93	3.55-3.93
4	10.50-11.50	10.50-11.50	10.50-11.50
5	Channel 4 repeated	11.5-12.5	11.5-12.5

**Table 3-2:** Used transformed AVHRR image features for daytime probabilistic cloud masking.

Feature name	Definition	Main cloud detection ability
<b>Rvis</b>	<u>Over land:</u> AVHRR channel 1 TOA reflectances minus PPS thresholds  <u>Over ocean:</u> AVHRR channel 2 TOA reflectances minus PPS thresholds	Identification of bright clouds over dark Earth surfaces
<b>Tirdiff</b>	AVHRR channel 4 brightness temperatures minus ERA-Interim (Dee et al, 2011) surface skin temperatures minus PPS thresholds	Identification of clouds which are significantly colder than the Earth surface
<b>Rnir_3a</b> (morning orbit VHRR/3)	AVHRR channel 3a reflectances divided by AVHRR channel 1 reflectances	Identification of clouds with significant reflection in the visible near-infrared region (in particular water clouds and thick multi-layered ice clouds over snow-covered surfaces)
<b>Rswir_3b</b> (afternoon orbit all AVHRRs and morning orbit AVHRR/2)	(AVHRR channel 3b brightness temperatures minus AVHRR channel 5 brightness temperatures) minus PPS thresholds	Identification of clouds with significant reflection in the short-wave infrared region (water clouds and thick multi-layered ice clouds), alternatively, clouds with significantly higher transmissivity in channel 3b than in channel 5 (thin ice clouds)

Feature name	Definition	Main cloud detection ability
<b>Texture_day</b>	<u>Over land</u> : Not used (surface variability generally too large)!  <u>Over ocean</u> : (Sum of local 3x3 pixel variances for A VHRR channel 1 TOA reflectances, A VHRR channel 3b brightness temperatures (or A VHRR channel 3a reflectances), A VHRR channel 4 brightness temperatures and A VHRR channel 3b and 5 brightness temperature differences) minus PPS thresholds	Identification of fractional or broken clouds over ocean

With this background we now list in Table 3-2 and Table 3-3 a set of 8 transformed and constrained image features (related to original AVHRR channels described in Table 3-1) that will be used later for the definition of the CMA-prob probabilistic cloud mask estimates. Four of them are selected for daytime conditions (i.e., solar zenith angles below 90° - Table 3-2) and four of them for night-time conditions (Table 3-3). However, one feature (Tirdiff) is used both day and night. Finally, in order to account for geographical and topographical differences we define 9 geographical regions over which we will train the probabilistic classifiers. These regions are listed in Table 3-4.

**Table 3-3:** Used transformed AVHRR image features for night-time probabilistic cloud masking.

Feature name	Definition	Main cloud detection ability
<b>Tirdiff</b>	A VHRR channel 4 brightness temperatures minus ERA-Interim surface skin temperatures minus PPS thresholds	Identification of clouds which are significantly colder than the Earth surface
<b>Tcidiff</b>	A VHRR channel 4 brightness temperatures minus A VHRR channel 5 brightness temperatures minus PPS thresholds	Identification of thin cirrus clouds
<b>Twdiff</b>	(A VHRR channel 3b brightness temperatures minus A VHRR channel 4 brightness temperatures) minus PPS thresholds	Identification of water clouds
<b>Texture_night</b>	<u>Over land</u> : Not used (surface variability generally too large)!  <u>Over ocean</u> : (Sum of local 3x3 pixel variances for A VHRR channel 4 brightness temperatures and A VHRR channel 3b and 5 brightness temperature differences) minus PPS thresholds	Identification of fractional or broken clouds over ocean

Snow and land use information were taken from National Snow and Ice Data Center (NSIDC) analyses and International Geosphere Biosphere Programme (IGBP) analyses, both of them provided together with the used CALIPSO-CALIOP cloud product (denoted Cloud and Aerosol Layer Information product version 3.01).

**Table 3-4:** Geographical regions used when training the probabilistic classifiers.

Geographical region	Definition
<b>Polar ocean</b>	Ice-covered ocean at latitudes higher than 40°
<b>High-latitude ocean</b>	Ice-free ocean at latitudes higher than 40°
<b>Low-latitude ocean</b>	Ocean at latitudes lower than 40°
<b>High-latitude snow-covered mountains</b>	Mountain regions (topography exceeding 500 m) with snow-cover at latitudes higher than 40°
<b>High-latitude snow-free mountains</b>	Mountain regions (topography exceeding 500 m) without snow-cover at latitudes higher than 40°
<b>High-latitude snow-covered land</b>	Snow-covered land (topography below 500 m) at latitudes higher than 40°
<b>High-latitude snow-free land</b>	Snow-free land (topography below 500 m) at latitudes higher than 40°
<b>Desert regions</b>	Land areas without vegetation at latitudes lower than 40°
<b>Low-latitude vegetated regions</b>	Vegetated land areas at latitudes lower than 40°

### 3.6 Training the classifier using CALIPSO-CALIOP cloud data

For developing the CMA-prob classifier we have taken advantage of the previously collected dataset with optimally matched NOAA-18 and CALIPSO orbits described by Karlsson and Johansson (2013). This study and also several other studies (e.g., Stengel et al., 2012) have demonstrated that it is possible to collocate NOAA AVHRR data with CALIPSO data with comparable quality to what is achieved when matching with other internal datasets in the Aqua train (e.g. MODIS data). Some example results from this dataset have already been shown in Sections 3.4 and 3.5. However, some important and necessary restrictions to the utilised information have been applied during the training process.

A great asset of the CALIPSO-CALIOP cloud products is the superior sensitivity for cloud detection compared to corresponding conditions for passive data like data from the AVHRR sensor. But this is also a problem when using this information as the basis for a statistical training of a probabilistic cloud masking method. More clearly, there is a risk for “over-training”, i.e., that we force the method to try to detect clouds that are theoretically impossible to detect from AVHRR sensor data. As a result, the probabilistic cloud-screening method would then risk to systematically create artificial clouds in truly cloud-free areas since the cloud-free signal cannot be confidently separated from the cloudy signal for e.g. sub-visible cirrus clouds. Consequently, we need to find a way to restrict the used CALIOP-based cloud mask in the training process to include only those clouds which we believe are potentially discernible also in AVHRR images. In other words, we need to define as accurately as possible the AVHRR

cloud detection limit. On the other hand, applied training restrictions must not go too far so that they preclude detecting potentially detectable clouds which are not generally detected by today's cloud screening methods. We need to leave some margin for further improvement of cloud detection performance even if that margin probably is very small (when considering that the experience of AVHRR cloud detection is now based on more than 30 years of development).

We have again utilised the dataset collected by Karlsson and Johansson (2013) for finding the appropriate cloud detection limit. They concluded that the PPS method reached its optimal performance for clouds with optical thicknesses greater than 0.35. Below this value the method started to systematically miss clouds with increasing magnitude for smaller and smaller cloud optical depths. Further analysis of their data revealed that below a cloud optical thickness of approximately 0.2 the PPS loss of clouds exceeds 50 %, i.e., less than 50 % of the clouds with this optical thickness are detected. We have used this value (i.e., cloud optical thickness of 0.2) to represent the AVHRR cloud detection limit in the training of the probabilistic classifier. CALIPSO-CALIOP detected clouds below this threshold are treated as being non-existing and equivalent to cloud-free conditions. This compromise solution means that some clouds are still likely to be non-detectable by the probabilistic classifier but some of the currently non-detected clouds in the cloud optical thickness interval 0.2-0.35 may potentially be identified. As a consequence, our probabilistic classifier might still over predict cloud probability to some extent which may have some consequence for the final use of the results (for example, when creating new fixed cloud masks based on the probabilistic results).

The final training dataset consisted of the same matched global NOAA-18 and CALIPSO orbits (99 orbits in total) as being used by Karlsson and Johansson (2013). In addition, to cover also measurements from the 1.6  $\mu\text{m}$  channel (channel 3A in Table 3-1) not available for NOAA-18, the training dataset has been extended with 385 matched orbits from the NOAA-17 satellite. The large number of matched orbits in comparison to the NOAA-18 dataset is explained by the fact that matches with CALIPSO is only possible for morning orbit satellites close to the latitudes of around 70 degrees on both hemispheres. Also, instead of full global orbits being matched as for the NOAA-18 case, we have for every match only access to a limited part of an orbit (approximately equivalent to the AVHRR swath width since orbital tracks cross almost perpendicularly). Thus, to reach statistical significance of results we need to use a larger number of matched orbits. Notice also that the matched morning satellite dataset is only used for studies of AVHRR channel 3A and its related image features. For all other features, the statistics collected from the NOAA-18 dataset has been used.

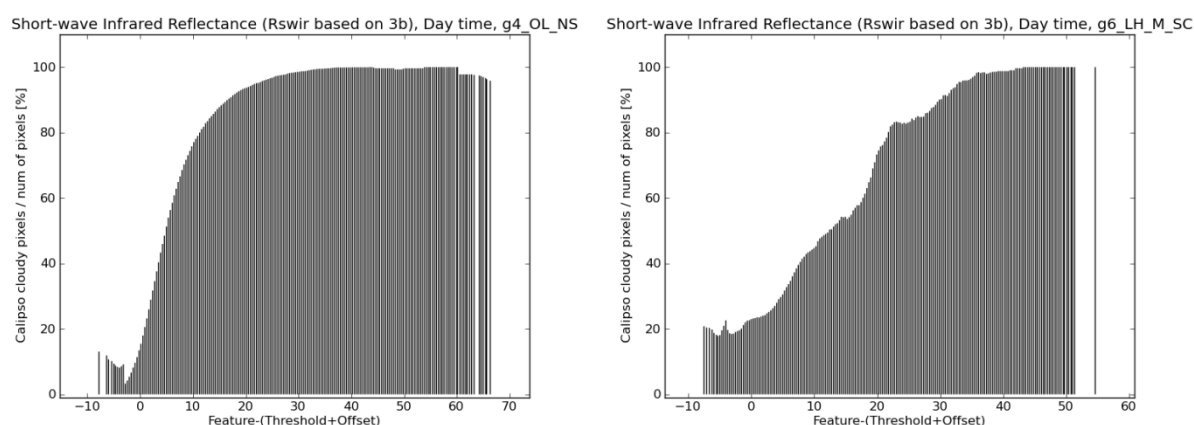
The collected training dataset spans the period 2006-2009 and provides a reasonable global coverage over all seasons during that period. All in all, it comprises almost 1 million matches of AVHRR Global Area Coverage (GAC) and CALIOP pixels/samples at approximately 5 km horizontal resolution. The constrained training (i.e., image features now being related to PPS threshold information) is based on results from the PPS software version 2014 patch 1 [RD 1]. This is a much-advanced PPS version compared to the original method described by Dybbroe et al. (2005). The main new features of the method concern adaptations to global processing (e.g., over desert and Polar Regions) and a systematic use of prescribed MODIS-derived surface emissivity information.

### 3.7 Achieved probability distributions

In this section we present some examples of the achieved probability distributions for most image features. In particular, we have chosen to display results over ice-free ocean at low latitudes together with results over snow-covered terrain at high altitudes (Greenland, Antarctica). These two different categories could be imagined as representing the least problematic (tropical ice-free ocean) and also the most challenging (snow-covered and very cold polar) of all categories. Thus, it gives some insight into what is possible and not possible as regards cloud screening of AVHRR imagery. Notice here that we show the conditional probabilities in the form  $P(f_i|cloudy)$  while the inverse form has to be used in the Naïve Bayesian formulation. However, as previously mentioned there is a mutual inter-relation between the two different conditional probabilities. Thus, they can be seen as just two different realizations of the same probability density distribution.

Figures Figure 3-1 and Figure 3-2 have already showed these probability distributions for image features **Rvis** and **Tirdiff** in Table 3-2 so we will not repeat them here.

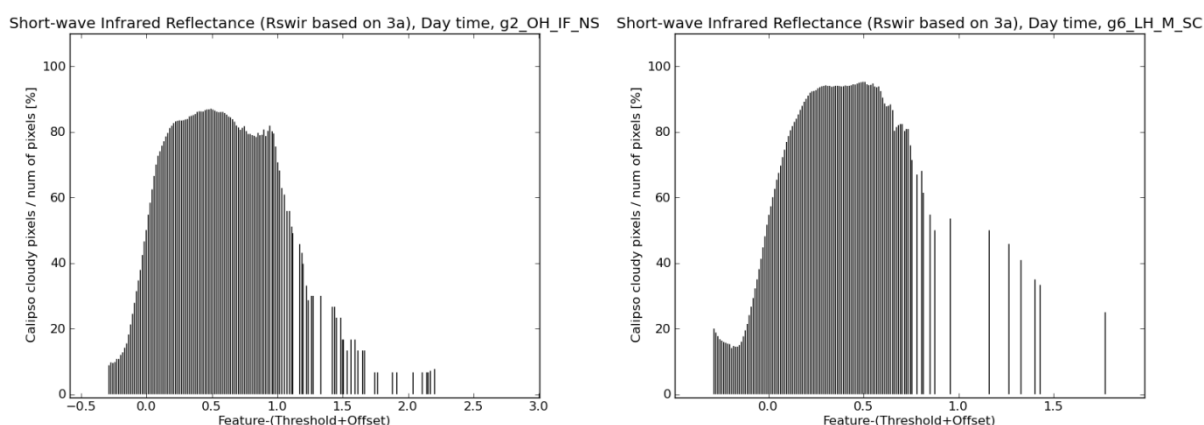
Figure 3-4 shows the distribution for the two different geographic categories for the reflected shortwave infrared feature **Rswir\_3b** (see Table 3-2). We notice the particularly high capability of separating cloudy from cloud-free cases over Tropical Ocean. It is obvious that cloud reflection is significantly more pronounced than any surface reflection in AVHRR channel 3b (if excluding sunglint like here). However, especially important is to notice the very good cloud separability offered here also over snow-covered cold Polar Regions. This differs from the performance of the two visible and infrared features illustrated earlier in Figure 3-1 and Figure 3-2, the latter showing a clearly decreasing capability for cloud screening at the polar location.



**Figure 3-4:** Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the **Rswir\_3b** feature (see Table 3-2) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009.

An interesting follow-on question here is to examine if the same cloud screening capability is retained over the polar region if replacing AVHRR channel 3b with channel 3a (defining the **Rnir\_3a** feature). The probability distributions for this case are illustrated in Figure 3-5 derived from CALIPSO matches with the NOAA-17 satellite. This also means that the ocean results are derived over ice-free high latitude ocean surfaces and not over tropical ocean surfaces. We conclude that also here we have a high capability for separating cloudy and clear cases

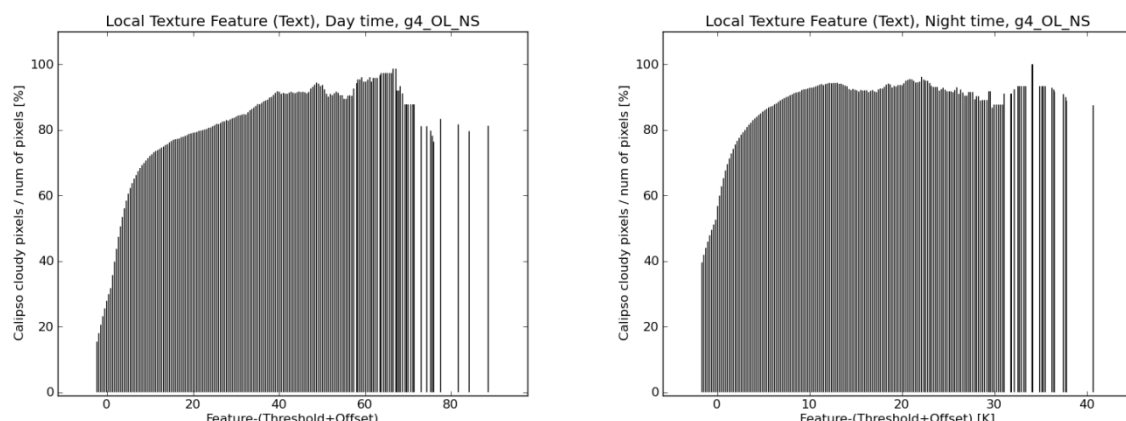
over both regions. The capability seems to be even higher for the polar case due to a higher contrast between cloud-free and cloudy cases over a smaller value range of the image feature.



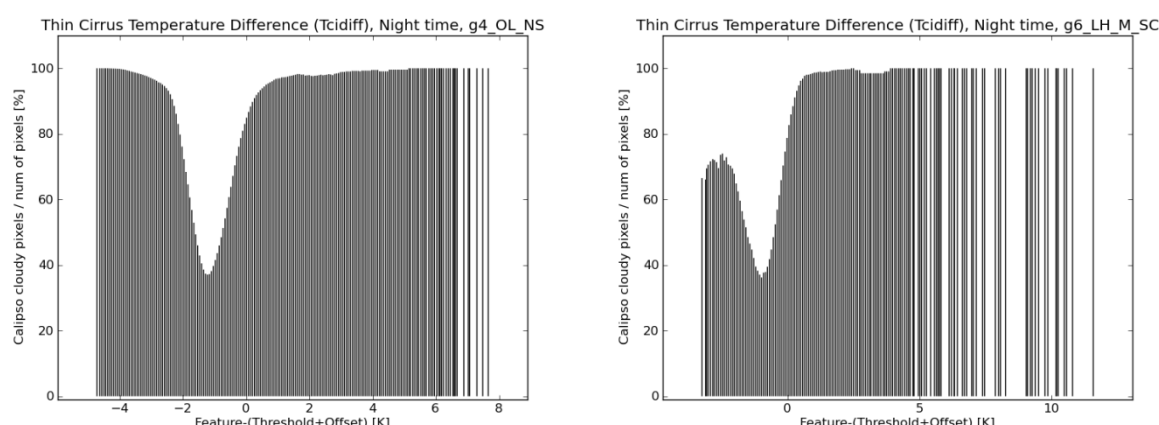
**Figure 3-5:** Frequency of cloud occurrence over High Latitude Ice-free ocean (left) and over Polar Snow-covered surfaces (right) as a function of the Rnir\_3a feature (see Table 3-2) derived from 385 globally matched NOAA-17/CALIPSO orbits in the period 2006-2009.

The remaining daytime feature to illustrate is the texture feature **Texture\_day**. However, since this feature is only used over ocean we compare results with the corresponding night version of this feature (**Texture\_night**) over Tropical Ocean in Figure 3-6. We conclude that this feature seems to give good contributions both day and night but that the usefulness appears to be still less than for the other features. This concerns especially the night-time version where probabilities are generally higher than 40 % everywhere, thus no value represents very clearly a large portion of cloud-free conditions.

We have two more night-time features to illustrate and Figure 3-7 shows the distribution for the **Tcidiff** feature (see also Figure 3-3). The two distributions are rather similar but it is clear that the cloud probability peak for lower differences (representing thick clouds) is less pronounced for the polar case. This indicates a decreasing capability of identifying opaque or multi-layered clouds over the Polar Region.

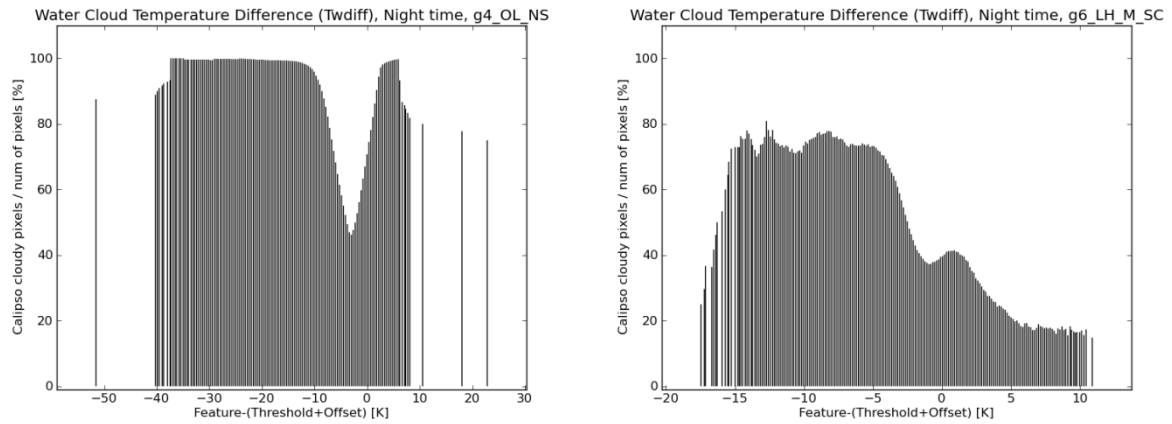


**Figure 3-6:** Frequency of cloud occurrence over Tropical Ocean as a function of the texture feature (see Table 3-2 and Table 3-3) during day (left) and night (right) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009.



**Figure 3-7:** Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the Tcidiff feature (see Table 3-3) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009.

Finally, Figure 3-8 shows the distributions for the **Twdiff** feature in Table 3-3. This feature is clearly able to separate water clouds (positive differences) from ice clouds (negative differences) over Tropical Ocean while this capability seems to be completely gone over the Polar Region. Here, only ice clouds seem to be discernible with some significance (however rather low probabilities).

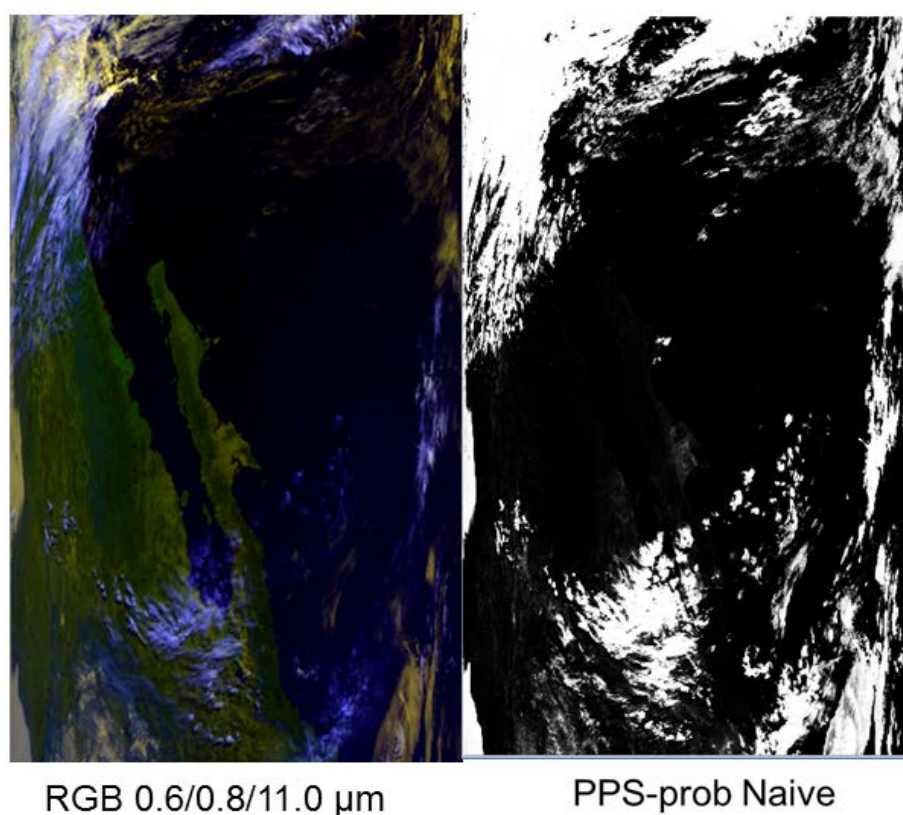


**Figure 3-8:** Frequency of cloud occurrence over Tropical Ocean (left) and over Polar Snow-covered Surfaces (right) as a function of the Twdiff feature (see Table 3-3) derived from 99 globally matched NOAA-18/CALIPSO orbits in the period 2006-2009.

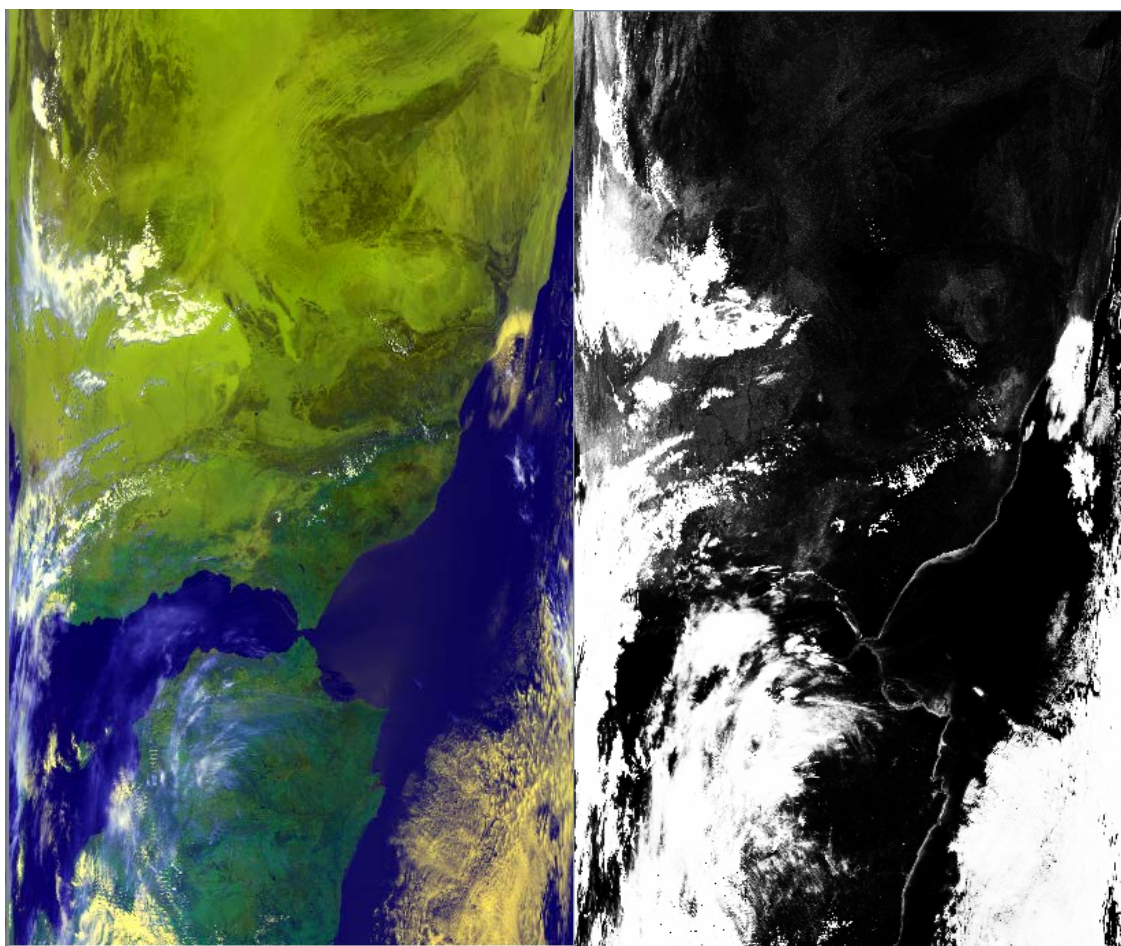
## 4 First results

### 4.1 Product demonstration

Figure 4-1 illustrates a case with the full cloud probability result displayed as a greyscale image together with a colour composite image of the original radiances. We notice from visual inspection that the areas with high CMA-prob cloud probabilities (white colours) correspond very well to cloud fields identified by visual inspection in the RGB composite for this particular case. However, noteworthy is that thin and broken cloud fields over the ocean surfaces are much more highlighted in the CMA-prob image than in the colour composite. This is mainly explained by the added cloud information coming from features **Rswir\_3b** and **Texture\_day** described earlier in Table 3-2. These features contain information from the 3.7 and 12  $\mu\text{m}$  channels which is information that is not displayed by the colour composite in the leftmost panel of Figure 4-1. Thus, the CMA-prob results are clearly based on more information than what is displayed in the RGB representation in Figure 4-1.

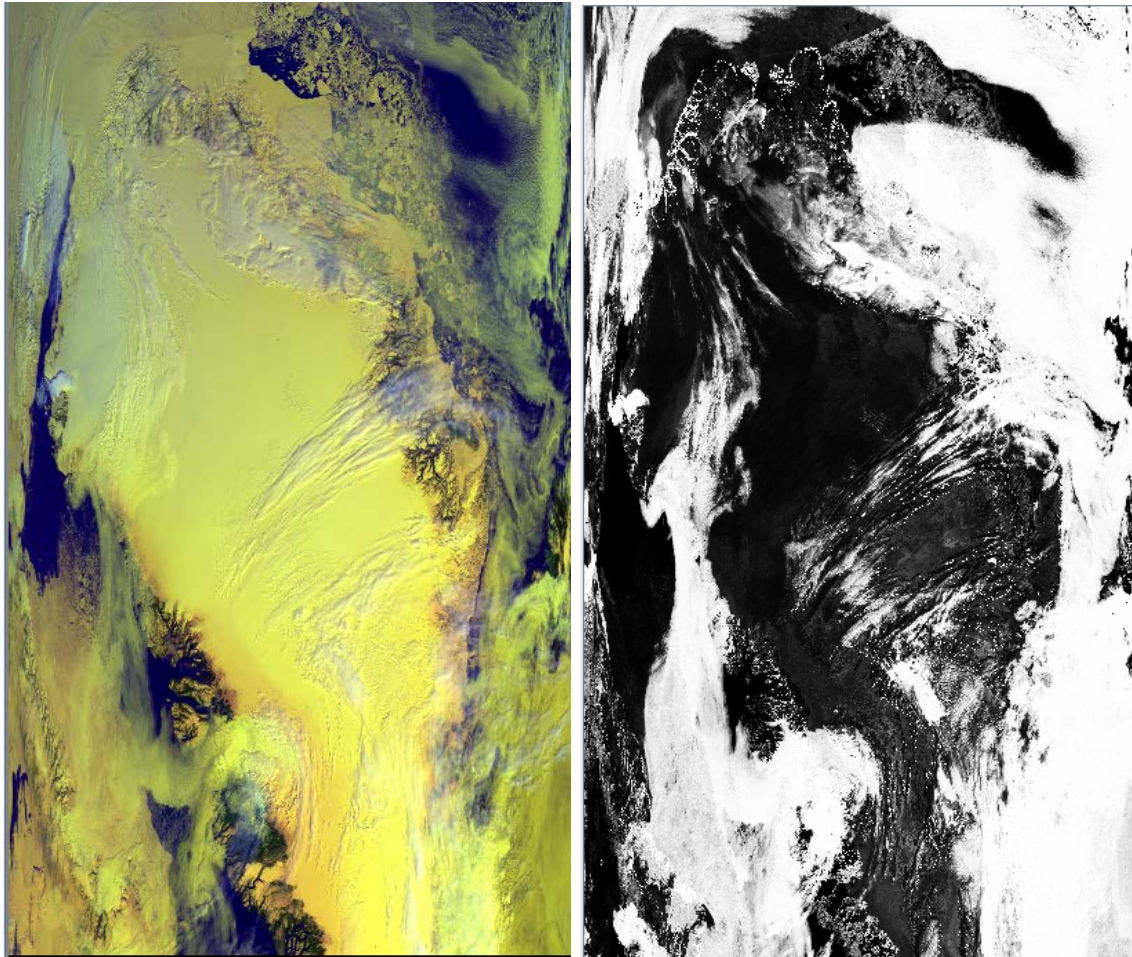


**Figure 4-1:** Part of an original NOAA-18 AVHRR GAC scene in satellite projection over the North American west coast (with Gulf of California and Baja California in the center) registered in ascending mode (i.e., North is down, South is up) from 26 January 2010. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100 %).



**Figure 4-2:** Part of an original NOAA-18 AVHRR GAC scene in satellite projection over Spain and northern Africa registered in ascending mode (i.e., North is down, South is up) from 16 May 2007. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100 %).

Figure 4-2 shows a NOAA-18 case including parts of the desert regions of Northern Africa. This example illustrates clearly how clouds and bright desert surfaces can be efficiently separated taking advantage of the full information content in all AVHRR channels. Especially, we notice that despite being over rather bright desert surfaces (i.e., surface reflectances in visible channels are here rather close to cloud reflectances) the resulting cloud probabilities are clearly at the zero level (black areas) for cloud free areas and close to 100 % (white areas) for cloudy areas. Notice again the high cloud probabilities for obviously rather thin cirrus cloud fields over Spain and the Mediterranean Sea. The most important information here comes mainly from AVHRR channels 3 and 5 which are not displayed in the colour composite image. The problem of using contextual (texture) information in applications like this results in high cloud probabilities on the sea side of the coast-lines (since texture features are only used over ocean surfaces). Consequently, also probabilistic approaches need a special coastal treatment which will be considered in the upcoming versions of CMA-prob.



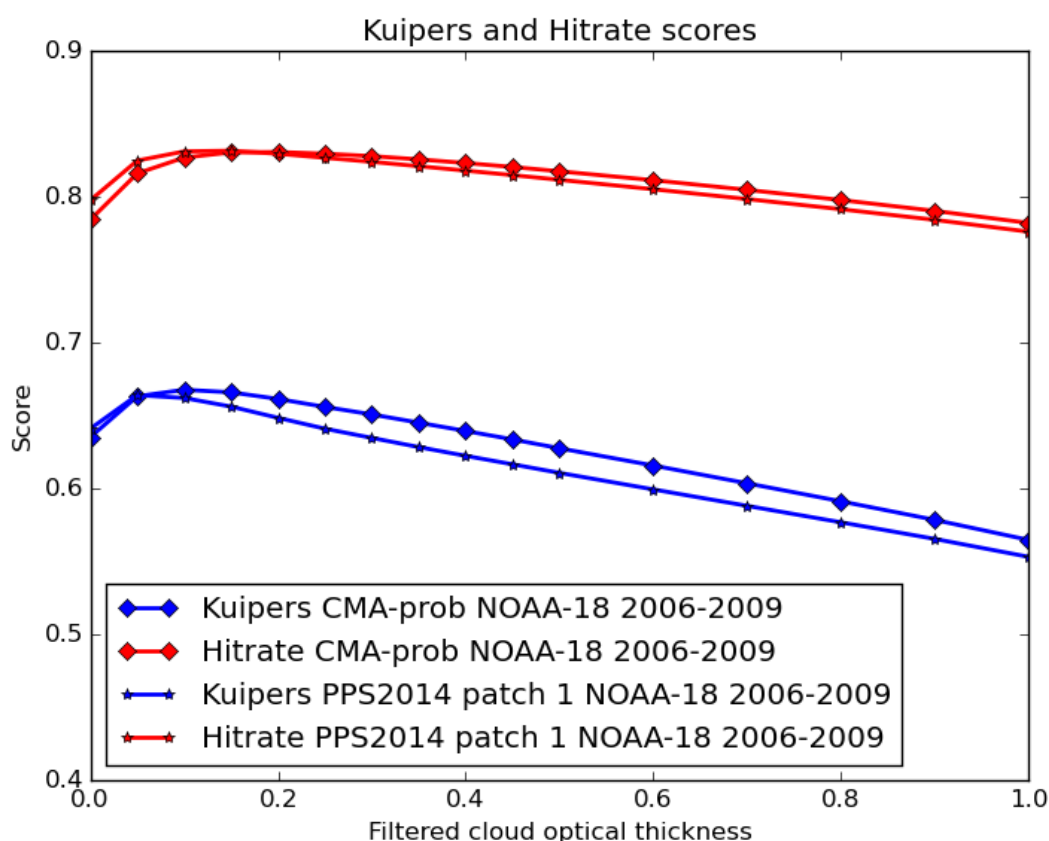
**Figure 4-3:** Part of an original NOAA-17 AVHRR GAC scene in satellite projection over Greenland registered in descending mode (i.e., North is up, South is down) from 4 June 2009. Left: Colour composite with AVHRR channel 1 (red), channel 2 (green) and channel 4 (blue). Right: Corresponding CMA-prob Naïve cloud probabilities (as greyscale image with range 0-100 %).

Finally, Figure 4-3 shows a NOAA-17 case over Greenland from 4 June 2009. This case clearly illustrates the strength of the **Rnir\_3a** feature in Table 3-2. The colour composite shows how snow-covered surfaces and cloud features are hard to separate (if not using shadow effects) but considering also the information in AVHRR channel 3a makes this distinction very efficient in the CMA-prob image.

## 4.2 Preliminary validation results

The new CMA-prob results have been evaluated using CALIPSO cloud mask for the 99 global orbits from the NOAA-18 satellite in the period 2006-2009 and results are shown in Figure 4-4. To realise this validation CMA-prob results were first converted into a binary cloud mask using a threshold value of 60 %. As has been explained previously, the training of CMA-prob from CALIPSO cloud masks included (despite special precautions) a small fraction of clouds not theoretically detectable in AVHRR data. Consequently, the realised CMA-prob probabilities are likely to be slightly too high. This motivates a higher threshold than 50 %. It was also shown by Karlsson et al. (2015) that the best validation results were found for a threshold close to 60 %.

Results in Figure 4-4 are visualised using the same plotting method as in the study of Karlsson and Johansson (2013). This method plots results as a function of thresholded cloud optical thicknesses which means that all CALIOP-detected clouds below the shown cloud optical thickness on the x-axis are treated as being cloud-free (i.e., thinner clouds filtered out). Consequently, original unfiltered results are seen for a cloud optical thickness value of 0.0.

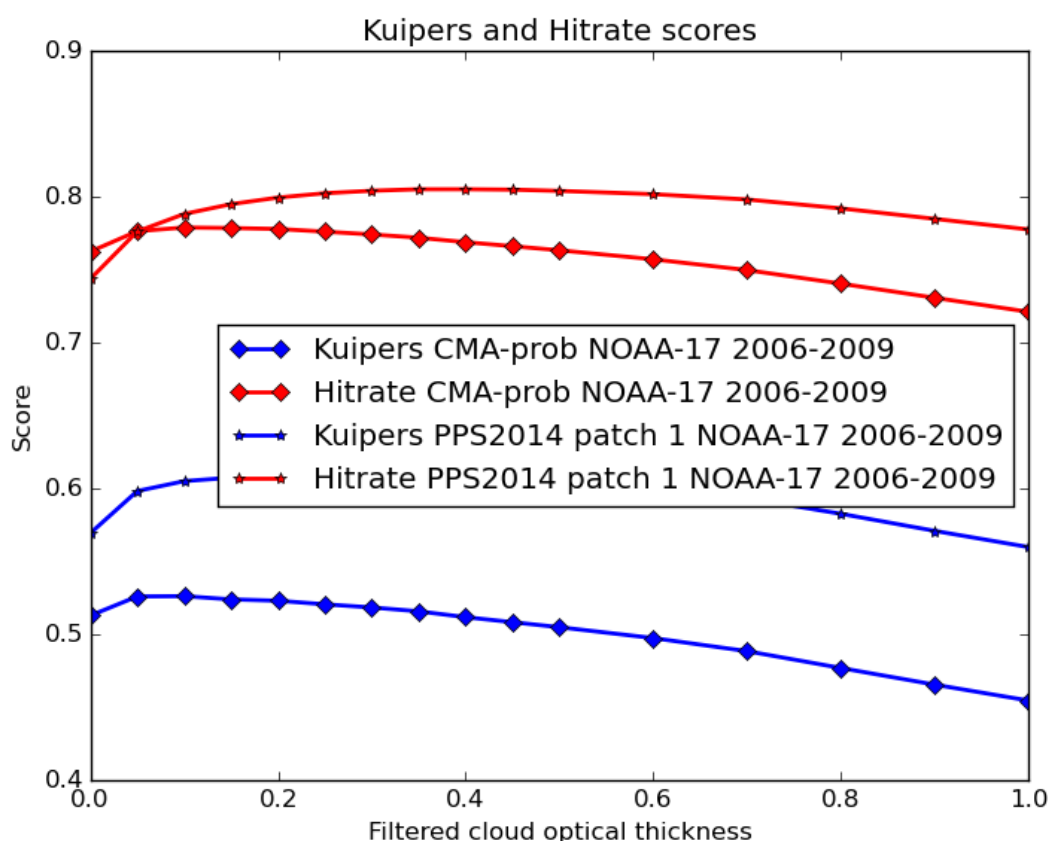


**Figure 4-4:** Hitrate and Kuipers Skill Scores plotted as a function of filtered CALIPSO-CALIOP cloud optical thickness (explained in text). Results for CMA-prob are compared to results of the official PPS version 2014 Patch 1 ([RD 1]). All results were derived for 99 NOAA-18 orbits in the period 2006-2009.

Results in Figure 4-4 clearly shows (e.g. at optical thickness value of 0.2 assumed to represent the lower detectability limit) that CMA-prob results are clearly comparable to the official PPS 2014 patch 1 results. This is very encouraging considering that PPS 2014 uses a rather complex setup of image features and threshold sequences adapted to specific conditions. Despite the use of a rather limited set of image features and conditions, CMA-prob is still capable of producing comparable results to PPS 2014.

Results from a similar validation of corresponding results from the 385 NOAA-17 orbits are shown in Figure 4-5. Here we can see that the official PPS method is still significantly better, especially as regards the separation of cloudy and clear cases (best reflected in the Kuipers' score). However, the deviation between the methods is to a large extent explained by special precautions taken in the official PPS method for very cold conditions over Greenland and Antarctica when the 3.7  $\mu\text{m}$  channel risk to become saturated due to poor radiometric resolution. The official PPS 2014 patch 1 method stops using brightness temperatures in that

channel for surface temperatures below 235 K due to these problems. This is not done for CMA-prob and this degrades results significantly, especially for night-time (Polar night) conditions. However, similar actions or restrictions could also be applied for CMA-prob in future realisations in order to improve results. In any case, day-time results for CMA-prob applied to NOAA-17 data are still comparable with official PPS 2014 patch 1 results.



**Figure 4-5:** Hitrate and Kuipers Skill Scores plotted as a function of filtered CALIPSO-CALIOP cloud optical thickness (explained in text). Results for CMA-prob are compared to results of the official PPS version 2014 Patch 1 ([RD 1]). All results were derived for 385 NOAA-17 orbits in the period 2006-2009.

Validation results have still to be extended with results based on an independent dataset (i.e., which has not been used for training). However, this has been done for an earlier prototype of the method based on PPS version 2010 and actually being the basis for the study by Karlsson et. al., 2010. Here, results were also evaluated on 78 independent NOAA-18 and NOAA-19 orbits in 2010 with very encouraging results.

## 5 Limitations and areas for future improvements

The CMA-prob method described in this document is the second prototype version based on dynamic threshold information from PPS version 2014 patch 1. The first version was described in the paper by Karlsson et al. (2015) and was based on dynamic threshold information from PPS version 2010. Thus, it is clear that this method should be seen as an extension of the official PPS software and it cannot be run independently from PPS. Consequently, if continuing with this approach the method needs to be updated (new training) for every new release of the PPS method.


The most prominent limitation of the current version of the method is that situations in twilight conditions are not treated separately as it is actually done in the official PPS scheme. Thus, training was made only for night-time and day-time conditions (separated by the solar zenith angle of 90 degrees) and the method has been implemented in the same manner. It is clear that results are therefore inferior to official PPS results in conditions near twilight. This will be taken care of in the next update of the method (linked to the planned PPS 2017 release).

Statistical methods are always limited by the amount of training data being used. It is clear from probability distributions shown in Section 3.7 that the current training dataset is not capable of providing very well-defined probability distributions. In that respect, it is very encouraging that results are still as good as being documented. However, it is very clear that an improved amount of training data would be beneficial for future versions of the method. Consequently, it is planned to extend the training material with considerably more data from CALIPSO in the period 2010 up to present date.

One particularly critical part of the training of the method was the inclusion of data from morning satellites using the 1.6  $\mu\text{m}$  channel. Here data from NOAA-17 was utilised and matched with CALIPSO. However, this is only possible at high latitudes. Thus, in the current method no training data has been achieved for lower latitudes and in particular over (sub-) tropical dry regions (deserts) where reflectivities in this channel are particularly high and often comparable to cloud reflectivities. Currently, statistics derived for snow-free land areas at high-latitude are also used for lower latitudes. This works surprisingly well but future efforts have to include ways of better characterising this image feature characteristics over low- and mid-latitudes. This can be done using other references than CALIPSO (e.g. MSG-SEVIRI data).

## 6 References

- Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars, L. van de Berg, J. Bidlot, N. Bormann, C. Delso, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J.-J. Morcrette, B.-K. Park, C. Peubey, P. de Rosnay, C. Tavalato, J.-N. Thépaut and F. Vitart, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quart. J. Roy. Meteor. Soc.*, **137**, 656, (doi = 10.1002/qj.828), 553—597.
- Derrien, M. and H. LeGléau, 2005: MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.*, **26**, 4707-4732.
- Dybbroe, A., A. Thoss and K.-G. Karlsson, 2005a: NWC SAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling - Part I: Algorithm description, *J. Appl. Meteor.*, **44**, pp. 39-54.
- Dybbroe, A., A. Thoss and K.-G. Karlsson, 2005b: NWC SAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling - Part II: Tuning and validation, *J. Appl. Meteor.*, **44**, 55-71.
- Frey, R. A., S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang, 2008: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5, *J. Atm. Ocean. Tech.*, **25**, 1057-1072.
- Heidinger, A.K., A. T. Evan, M. Foster and A. Walther, 2012: A Naïve Bayesian Cloud Detection Scheme Derived from CALIPSO and Applied within PATMOS-x. *J. Appl. Meteor. Climatol.*, **51**, 1129-1144.
- Inoue, T. 1987: A cloud type classification with NOAA 7 split-window measurements. *J. Geoph. Res.*, **92**, 3991–4000, DOI: 10.1029/JD092iD04p03991.
- Karlsson K.-G. and A. Dybbroe, 2010: Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations. *Atmos. Chem. Phys.*, **10**, 1789–1807, 2010.
- Karlsson, K.-G. and E. Johansson, 2013: On the optimal method for evaluating cloud products from passive satellite imagery using CALIPSO-CALIOP data: example investigating the CM SAF CLARA-A1 dataset. *Atmos. Meas. Tech.*, **6**, 1271–1286, [www.atmos-meas-tech.net/6/1271/2013/](http://www.atmos-meas-tech.net/6/1271/2013/), doi:10.5194/amt-6-1271-2013.
- Karlsson, K.-G., E. Johansson and A. Devasthale, 2015: Advancing the uncertainty characterisation of cloud masking in passive satellite imagery: Probabilistic formulations for NOAA AVHRR data, *Rem. Sens. Env.*, **158**, 126-139; doi:10.1016/j.rse.2014.10.028.
- Kossin, J. P. and M. Sitkowski, 2009: An objective model for identifying secondary eyewall formation in hurricanes. *Mon. Wea. Rev.*, **137**, 876-892.

	<p align="center"><b>Algorithm Theoretical Basis Document CLARA Edition 2 Probabilistic cloud masks CMA-prob</b></p>	<p>Doc. No: SAF/CM/SMH/ATBD/GAC/PBCM Issue: 1.3 Date: 15.10.2020</p>
-----------------------------------------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------

Kriebel, K. T., G. Gesell, M. Kästner and H. Mannstein, 2003: The Cloud Analysis Tool APOLLO: Improvements and Validations. *Int. J. Remote Sens.*, 01/2003; **24**(2003):2389-2408. DOI:10.1080/01431160210163065.

Merchant, C. J., A. R., Harris, E. Maturi and S. MacCallum, 2005: Probabilistic physically-based cloud screening of satellite infra-red imagery for operational sea surface temperature retrieval, *Quart. J. Royal Met. Soc.*, **131**, 2735-2755.

Musial, J. P., F. Hüsler, M. Sütterlin, C. Neuhaus and S. Wunderle, 2014: Probabilistic approach to cloud and snow detection on Advanced Very High Resolution Radiometer (AVHRR) imagery. *Atmos. Meas. Tech.*, **7**, 799-822, [www.atmos-meas-tech.net/7/799/2014/](http://www.atmos-meas-tech.net/7/799/2014/), doi:10.5194/amt-7-799-2014.

Pavolonis, M. J.; A. K. Heidinger and T. Uttal, 2005: Daytime global cloud typing from AVHRR and VIIRS: Algorithm description, validation, and comparisons. *Journal of Applied Meteorology*, **44**, Issue 6, 804-826.

Rossow, W. B., and R. A Schiffer, 1999: Advances in Understanding Clouds from ISCCP. *Bull. Amer. Meteor. Soc.*, **80**, 2261–2288.

Stengel, M., S. Mieruch, M. Jerg, K.-G. Karlsson, R. Scheirer, B. Maddux, J.F. Meirink, C. Poulsen, R. Siddans, A. Walther and R. Hollmann, 2013: The Clouds Climate Change Initiative: Assessment of state-of-the-art cloud property retrieval schemes applied to AVHRR heritage measurements, *Rem. Sens. Env.*, <http://dx.doi.org/10.1016/j.rse.2013.10.035>.

Winker, D.M., M.A.Vaughan, A. Omar, Y. Hu, K.A. Powell, Z. Liu, W.H. Hunt and S.A. Young, S.A, 2009: Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *J. Atm. Ocean. Tech.* **26**:11, 2310-2323.

## 7 Glossary

ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
CALIOP	Cloud-Aerosol Lidar with Orthogonal Polarization (CALIPSO)
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations
CDOP	Continuous Development and Operations Phase
CLARA	CM SAF cLoud, Albedo and surface RAdiation dataset
CMA-prob	Cloud Mask (probabilistic)
CM SAF	Satellite Application Facility on Climate Monitoring
CPP	Cloud Physical Properties
DRI	Delivery Readiness Inspection
DWD	Deutscher Wetterdienst (German MetService)
ECMWF	European Centre for Medium Range Forecast
ECV	Essential Climate Variable
EPS	European Polar System
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FOV	Field of view
GAC	Global Area Coverage (AVHRR)
GCOS	Global Climate Observing System
IOP	Initial Operations Phase
ITCZ	Inter-Tropical Convergence Zone
KNMI	Koninklijk Nederlands Meteorologisch Instituut
NASA	National Aeronautics and Space Administration
NDBC	National Data Buoy Center
NESDIS	National Environmental Satellite, Data, and Information System
NOAA	National Oceanic & Atmospheric Administration
NODC	National Oceanographic Data Center
NSIDC	National Snow and Ice Data Center

NWCSAF	Satellite Application Facility for Nowcasting
NWP	Numerical Weather Prediction
PPS	Polar Platform System
PRD	Product Requirement Document
PUM	Product User Manual
RMIB	Royal Meteorological Institute of Belgium
RMS	Root Mean Square
RSMAS	Rosenstiel School of Marine and Atmospheric Science
RSS	Remote Sensing Systems
SAF	Satellite Application Facility
SMHI	Swedish Meteorological and Hydrological Institute
SST	Sea Surface Temperature