Atmos. Meas. Tech. Discuss., 6, 8445–8507, 2013
www.atmos-meas-tech-discuss.net/6/8445/2013/
doi:10.5194/amtd-6-8445-2013

Atmospheric Measurement Techniques Discussions
Open Access

This discussion paper is/has been under review for the journal Atmospheric Measurement Techniques (AMT). Please refer to the corresponding final paper in AMT if available.

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial, F. Hüsler, M. Sütterlin, C. Neuhaus, and S. Wunderle

Geographisches Institut der Universität Bern (GIUB), 3012 Bern, Switzerland

Received: 15 May 2013 – Accepted: 2 August 2013 – Published: 25 September 2013

Correspondence to: J. P. Musial (jan.musial@giub.unibe.ch)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The derivation of probability estimates complementary to geophysical data sets has gained special attention over the last years. The information about a confidence level of provided physical quantities is required to construct an error budget of higher level products and to correctly interpret final results of a particular analysis. Regarding the generation of products based on satellite data the common input consists of a cloud mask which allows discrimination between surface and cloud signals. Further the surface information is divided between snow and snow-free components. At any step of this discrimination process a misclassification in a cloud/snow mask propagates to higher level products and may alter their usability. Within this scope a novel Probabilistic Cloud Mask (PCM) algorithm suited for the 1×1 km Advanced Very High Resolution Radiometer (AVHRR) data is proposed which provides three types of probability estimates between: cloudy/clear-sky, cloudy/snow and clear-sky/snow conditions. As opposed to the majority of available techniques which are usually based on a decision-tree approach in the PCM algorithm all spectral, angular and ancillary information is used in a single step to retrieve the probability estimates from the pre-computed Look Up Tables (LUTs). Moreover, the issue of derivation of a single threshold value for a spectral test was overcome by the concept of multidimensional information space which is divided into small bins by an extensive set of thresholds. The discrimination between snow and ice clouds and detection of broken, thin clouds was enhanced by means of the Invariant Coordinate System (ICS) transformation. The study area covers a wide range of environmental conditions spanning from Iceland through central Europe to northern parts of Africa which exhibit diverse difficulties for cloud/snow masking algorithms. The retrieved PCM cloud classification was compared to the PPSv2012 and MOD35 collection 6 cloud masks, SYNOP weather reports, CALIPSO vertical feature mask version 3 and to MOD10A1 collection 5 snow mask. The outcomes of conducted analyses proved fine detection skills of the PCM method with comparable or better results than the reference PPS algorithm.

AMTD
6, 8445–8507, 2013

1 Introduction

Cloud and snow detection on satellite imagery is a common part of a wide range of geophysical analysis. Therefore any misclassification introduced at this step have a direct effect on study results and may alter the final conclusions (Gómez-Chova et al., 2007). This issue has been widely discussed by a number of authors: Jones et al. (1996) described a significant diurnal bias in the 0.5° spatially averaged ATSR SST data product over the South Atlantic induced by the residual cloud contamination; Kaufman et al. (2005) found that misclassified clouds in MODIS imagery lead to 0.02 bias in the Aerosol Optical Thickness (AOT) estimates; Hall and Riggs (2007) analysed the improvements of cloud/snow discrimination in MODIS collection 5 data sets; Pincus et al. (2012) discussed the differences between cloud climatologies derived from MODIS and ISCPP data sets induced by different detection sensitivities and treatment of thin cirrus and partially cloudy pixels. Furthermore inconsistencies in satellite products employed by climate models increase their variability which mostly originate from the parametrisation of cloud radiative forcing (Houghton et al., 1996). In this respect accurate discrimination of cloud and snow covered areas supported by uncertainty estimations is required. Some of the existing approaches (Ackerman et al., 1998; Derrien and Le Gléau, 2005; Khlopenkov and Trishchenko, 2007; Vemury et al., 2001) separate classification results into few confidence categories (e.g. clear, probably clear, probably cloudy, cloudy). Nevertheless, implementation of this qualitative information into an error budget calculation of higher level products is not straightforward. The solution to this problem involves derivation of continuous probability estimates of each pixel belonging to clear, cloudy and preferably snow classes. There are few existing approaches which provide such a quantitative probability distribution together with classification results. Some of them are based on classical (Merchant et al., 2005; Uddstrom et al., 1999) or Naive (Heidinger et al., 2012) Bayesian theories which combine results of a single classifier (i.e. spectral/textural test) with a priori assumption on cloud condition in order to obtain posterior classification probability. The a priori knowledge originates from

AMTD
6, 8445–8507, 2013

additional data sets (e.g. climate model outputs) or collocated satellite observations (e.g. CALIPSO/CALIOP). Another algorithm proposed by Plummer (2008) expresses probability of cloud discrimination as a distance between tests results and threshold values. The final estimate is deemed as a maximum probability value across all performed tests. Other cloud masking method (Tian et al., 2000) involves probabilistic neural network classifiers employed to analyse temporal changes in a sequence of images. The clustering methods based on the Expectation Maximization (EM) technique were also found to be suitable for cloud probability retrieval (Gómez-Chova et al., 2007).

Derivation of probability estimates for snow discrimination on satellite imagery has been even less explored than in case of cloud detection. Recently, Hüsler et al. (2012) modified aggregated rating approach proposed by Khlopenkov and Trishchenko (2007) to suite European Alpine area and computed posterior snow classification probabilities employing logistic regression between ground data and numerical scores generated by spectral tests.

The main aim of this study was to develop a robust, meaning accurate and computationally inexpensive, algorithm that provides consistent probability estimates of a particular pixel in a satellite scene belonging to clear-sky, cloudy or snow classes. As the name "Probabilistic Cloud Mask" suggests the main focus of this study is on cloud coverage, however the validation of the snow component is presented as well. The PCM algorithm is suited for the 1 × 1 km AVHRR Local Area Coverage (LAC) data covering the extensive European region spanning from the northern parts of Africa to Iceland and the northernmost regions of Norway. The selected study area encompasses a wide range of ecosystems from desert to boreal vegetation and perenial snow together with broad illumination conditions including polar day and night. The variety of environmental conditions reflects different challenges occurring during the satellite cloud and snow discrimination.

The next section gives a short overview on existing algorithms for cloud and snow detection on AVHRR imagery with emphasis on required data sets and types of tests applied. Section 3 describes principals of the PCM method beginning with the rea-

AMTD
6, 8445–8507, 2013

soning for the spectral features enhancement by means of the Invariant Coordinate Transformation (ICS). Further the concept of multidimensional information space is discussed and the set of required input data together with thresholds values are introduced. Next the proposed methodology and its numerical implementation are discussed together with the theoretical algorithm limitations. Section 4 presents the comparison between the PCM results and the reference data consisting of PPS (Polar Platform System) and MOD35 cloud masks, MOD10A1 daily snow mask, CALIPSO/CALIOP vertical feature mask and SYNOP cloud observations. Section 5 contains the analysis of the acquired results which are concluded in Sect. 6.

2 Overview of existing multi-thresholding algorithms for satellite cloud and snow detection

The majority of existing algorithms for cloud and snow detection incorporate a series of spectral, textural and/or temporal tests (often called features) which are arranged in a decision-tree scheme. Singular tests are based on a comparison of radiances from a spectral channel or combination of channels with a threshold value which could be static (Bernstein, 1982) or may vary together with angular and atmospheric conditions (Ackerman et al., 1998; Dybbroe et al., 2005a; Di Vittorio and Emery, 2002; Yang et al., 2007). The former approach is more robust as the strong influence of atmospheric concentration of gases and aerosols, bidirectional reflectance, acquisition geometry on satellite measurements has been widely discussed in the literature (Fraser and Kaufman, 1985; Saunders et al., 1999; Schaaf et al., 2002; Vermote et al., 1997). Further surface properties such as albedo and emissivity modify spectral contrast and influence the cloud/snow detection sensitivity (Minnis et al., 2008; Yhann and Simpson, 1995). It is apparent over dark surfaces such as water in the Near Infrared (NIR) spectrum where discrimination of clouds is easier than over bright surfaces like deserts or snow (Ackerman et al., 1998; Di Vittorio and Emery, 2002). This implies that the thresholds should also vary across different land cover types. More difficult situations for the accu-

AMTD
6, 8445–8507, 2013

rate classification are related to pixels partially covered by snow or thin/broken clouds where there is a strong contribution from the surface in the measured satellite signal (Simpson et al., 2001). Furthermore, high aerosol loads in the atmosphere are often misclassified as cloud due to similar spectral signature (Martins et al., 2002). Threshold parametrization should account for all of the mentioned factors and either it can be derived empirically or optimised by means of RT modelling. Then it is stored as Look-Up Tables (LUTs) to safe computation time during the cloud/snow masking process (Dybbroe et al., 2005a).

Usually a single test provides a binary state of pixel such as cloudy/clear-sky or snow/snow-free. Further this information might be treated in a variety of ways. Some algorithms (Key and Barry, 1989) report cloud whenever one of the tests performed was successful. Other methods (Dybbroe et al., 2005a; Derrien and Le Gléau, 2005) arrange features in groups with decreasing detection sensitivity. In order to mark a pixel as cloudy all tests within a group have to be successful. Alternative approaches are based on a fuzzy logic where final confidence estimate of a pixel state is expressed as a product of singular estimates for each group of tests (Ackerman et al., 1998) or as a total sum of score values (Khlopenkov and Trishchenko, 2007; Hüsler et al., 2012). Nevertheless, in the further classification process these continuous confidence estimates are transformed into discrete classes using thresholds.

2.1 Ancillary data employed by multi-thresholding algorithms

Cloud and snow discrimination on satellite imagery is mainly based on multi-spectral measurements however additional ancillary data are required for the thresholds parametrization. This complementary information might be divided into meteorological and surface data sets, first of which feature high temporal variations whereas the latter ones usually are stable over time or change accordingly to well-known daily and annual cycles (Yhann and Simpson, 1995).

AMTD
6, 8445–8507, 2013

2.1.1 Ancillary meteorological data

An instantaneous atmospheric state can be estimated either by climate models or by rough approximations based on climatological mean values. Usually such simulations are of low spatial resolution yet an interpolation to a satellite grid results in significant bias over areas with rough topography (Zhao et al., 2008) or around zones with high temperature gradients such as coastlines. Another source of inaccuracies originates from temporal sampling of a climate model which may not correspond to satellite acquisition time thus data interpolation between two model steps is required (Khlopenkov and Trishchenko, 2007; Minnis et al., 2008). The atmospheric variable which significantly alters the radiative transfer is the concentration of water vapour often denoted as Total Column Water Vapour (TCWV). It expresses the integrated mass of water vapour per cross-sectional area unit of an atmospheric column ($kg\,m^{-2}$). Some of the sophisticated satellite cloud detection algorithms (Minnis et al., 2008) uses humidity, temperature and wind profiles instead of TCWV to more accurately resolve the RT processes within the atmosphere. The length of optical path related to the sensor viewing angle modifies the radiation absorption by water vapour thus it should be considered during the threshold derivation (Yhann and Simpson, 1995). Saunders and Kriebel (1988) proposed a split-window approach to account for this effect while discriminating cirrus clouds on a satellite image. The threshold for this spectral test is retrieved from the LUT where values are expressed as a function of the 10.8–12.0 µm Brightness Temperature Difference (BTD), secant of the sensor viewing angle and the 10.8 µm Brightness Temperature (BT).

The thermal contrast between surface and cloud tops can be a decisive factor when the spectral information is not sufficient for confident discrimination between cloudy or clear-sky conditions. It is derived as a difference between Skin Temperature (SKT) obtained from a climate model and the 10.8 µm BT. Moreover, the SKT data together with other atmospheric variables serve as inputs to RT calculations which are used to simulate satellite signals measured at specific spectral range for clear-sky conditions. These

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

AMTD

6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

expected values are then compared to real data acquired by sensor and if the deviations are significant a cloud presence might be assumed. Another useful information for the cloud detection on satellite imagery is the difference between air temperatures provided at two lowest altitude levels of the climate model. This indicates the presence of low level temperature inversion which reverses the expected thermal signature of clouds and may lead to misclassification (Dybbroe et al., 2005a).

2.1.2 Ancillary surface data

Surface characteristics are of great importance for the cloud/snow discrimination on satellite imagery, especially the land cover categorization. Therefore the availability of binary land/water mask is a minimum requirement for application of suitable threshold configurations within a classification algorithm. Sophisticated threshold parametrizations (Dybbroe et al., 2005a; Minnis et al., 2008) utilize more detailed land cover information together with Digital Elevation Model (DEM) data to enhance the detection accuracy over areas particularly ambiguous for correct classification (e.g. cloud detection over mountains, snow detection under the tree canopy). Additionally to land cover periodically updated snow/sea ice coverage data could be utilized for the further threshold refinements in cloud detection schemes (Dybbroe et al., 2005a; Minnis et al., 2008; Ackerman et al., 1998). In order to model the measured satellite signal, surface spectral properties such as albedo and emissivity have to be considered (Minnis et al., 2008). They change systematically over the coarse of a year therefore the RT simulation together with threshold parametrisation should feature temporal sampling (Dybbroe et al., 2005a).

2.2 Features used for satellite cloud/snow detection

Discrimination of cloud and/or snow coverage on satellite imagery is based on specific spectral properties at particular wavelengths which are utilised regardless the sensor

AMTD
6, 8445–8507, 2013

and algorithm employed. Next Subsections describe in details those of them which are applicable to the AVHRR instrument.

2.2.1 Reflectance tests in the 0.6 and 0.8 µm bands

In the Visible (VIS) 0.6 µm and Near Infrared (NIR) 0.8 µm spectral regions clouds and snow appear much brighter than underlying background. Furthermore, the spectral contrast of those surfaces over land is higher in the 0.6 µm channel whereas over water bodies it is more distinct in the 0.8 µm channel. Nonetheless, discrimination between snow and cloud cover in both wavelengths is impossible hence additional information either from the 1.6 or 3.7 µm channel is required. In order to diminish the influence of illumination conditions on a retrieved reflectance it is divided by a cosine of Sun Zenith Angle (SZA) and adjusted for the sun-earth distance variations. These corrections should be applied to any channel data within the reflective part of the electromagnetic spectrum. Some approaches (Dybbroe et al., 2005a) utilise reflectance in the channel 0.6 µm with and without the SZA normalisation, claiming that the latter one is particularly useful for cloud and snow detection over dark background at low sun elevations ($< 4^{\circ}$).

2.2.2 Reflectance tests in the 1.6 and 3.7 µm bands

In the 1.6 and 3.7 µm spectra there is a significant reflectance contrast between water clouds and snow. This spectral property together with the reflectance at 0.6 µm is employed by the Normalized Difference Snow Index (NDSI) in order to standardise and enhance snow detection. Although high NDSI values are usually associated with snow covered areas they may refer to ice clouds as well. This ambiguity is difficult to resolve using only spectral information thus some ancillary data such as Skin Temperature (SKT) should be utilized. Despite the equality of snow and ice clouds spectral response, one may assume the presence of an ice cloud when the difference between SKT and the 10.8 µm BT significantly deviates from 0 K. This situation often occurs

AMTD
6, 8445–8507, 2013

due to strong convection during the summertime when cloud tops (e.g. Cumulonimbus) consist of ice particles while the Earth's surface remains warm. Even during the cold part of the year this thermal contrast might be decisive depending on accuracy and temporal sampling of the SKT data.

2.2.3 Brightness temperature difference test between 3.7 and 10.8/12.0 µm bands

Contamination of the 3.7 µm signal by the solar component complicates utilization of this channel during the daytime. Thus usually it is applied during the night to detect clouds which have lower emissivities and scatters radiation more efficiently at this wavelength than at 10.8 or 12.0 µm. For optically thick clouds this results in the negative BTD between 3.7 µm and the IR channels while for thin clouds, where a lot of surface radiation is transmitted, the difference is positive. Furthermore, during night the 3.7–10.8 µm BTD was found to be more useful for detection of warm clouds and low stratus/fog layers (Eyre et al., 1984) whereas the difference between 3.7 and 12.0 µm has high sensibility to thin cirrus (Dybbroe et al., 2005a). For low radiative temperatures measured by early generations of the AVHRR instrument (prior to NOAA15) some cloud detection inconsistencies may occur due to periodic noise in the 3.7 µm channel (Warren, 1989).

2.2.4 Brightness temperature difference test between 10.8 and 12.0 µm bands

The 10.8–12.0 µm BTD is particularly useful for detection of cirrus clouds which are not apparent at other wavelengths. It is positive for thin clouds due to higher atmospheric transmittance at 10.8 µm than at 12.0 µm (Inoue, 1985). Moreover, it heavily depends on atmospheric water vapour concentration and sensor viewing angle thus a threshold value for the cirrus detection test should be derived dynamically using radiative transfer modelling or a robust parametrization (Saunders and Kriebel, 1988).

AMTD
6, 8445–8507, 2013

2.2.5 Temperature difference test between Earth's surface and the 10.8 µm band

The spectral region around 10.8 µm is slightly affected by the absorption of atmospheric gases (so called atmospheric window) thus it well approximates the surface temperature. Therefore, it can be used together with SKT data derived from climatological records or from climate models. The cloud presence can be assumed if the difference between these two variables is sufficiently high. This feature is particularly useful over ocean during night when water temperature is usually well approximated by the climate models. However, special attention is required in case of the temperature inversion when the expected positive difference between SKT and 10.8 µm cloud top temperature becomes negative. Over barren or sparsely vegetated areas such as deserts strong diurnal surface temperature cycle might be poorly represented in the SKT data (Pavolonis, 2010) which may lead to erroneous test results. Therefore, considering all of the mentioned aspects the threshold values for this test should be rather conservative and limited to the detection of relatively cold clouds.

2.2.6 Spatial uniformity tests

Regardless spectral properties of clouds their appearance on a satellite image is distinct especially over homogeneous surfaces (Saunders and Kriebel, 1988; Ackerman et al., 1998). If a texture variation analysed within a small image window (e.g. 5 × 5 pixels) is significant then the cloud presence may be assumed over the central pixel. Due to high surface heterogeneity over land this test is usually applied only over water bodies using 0.8 µm channel during the day and 10.8 µm channel for the night. When dealing with polar region the sea ice cover has to be considered as it may exhibit similar texture variations as clouds. In such a case for the night-time conditions the 3.7 µm BT or 3.7–12 µm BTD are used to detect cloud edges and to filter out leads (cracks in ice filled with water) which could be of sub-pixel size (Dybbroe et al., 2005a).

2.2.7 Temporal consistency tests

Most of the high frequency positive anomalies of surface albedo or negative thermal anomalies in a sequence of satellite images can be attributed to clouds. The temporal consistency tests require derivation and periodical update of a reference clear-sky radiance statistics accumulated over some time period for various spectral channels (Rossow and Garder, 1993). These calculations are usually performed over small spatial domains (e.g. $30 \times 30\,\mathrm{km}^2$) for similar land cover types. More sophisticated approaches (Lyapustin et al., 2008) utilize a spatio-temporal covariance matrix in order to detect cloud induced changes in spectral and textural signatures at a single pixel level. While comparing the clear-sky radiances with different acquisition times special attention has to be paid to angular and atmospheric effects which alter the surface spectral response. Some possible solutions to these issues involve: threshold relaxation in order to account for a wide range of conditions; comparison of measurements taken at similar solar times from corresponding orbits; atmospheric and bidirectional effects corrections of the reference clear sky composite. Additional data processing employed by the temporal consistency tests increases computational demands, yet they feature valuable detection skills especially for the geostationary satellites.

2.3 Limitations of multi-spectral thresholding algorithms

Detection of clouds and snow on satellite imagery by means of the multi-spectral thresholding algorithms usually requires estimation of a single threshold value for each test which accounts for a wide range of atmospheric, angular and surface conditions (Yang et al., 2007; Yhann and Simpson, 1995). This complexity results either in reduced classification accuracy due to the threshold relaxation or in extensive analysis of a spectral and ancillary information. However, in each situation any misplacement of a single threshold value leads to errors which could be depreciated by utilization of a fuzzy approach (see score functions in Khlopenkov and Trishchenko, 2007). Furthermore, parametrizations of the multi-spectral thresholding algorithms are mostly derived

AMTD
6, 8445–8507, 2013

empirically thus their application to other satellite sensors is not straightforward. The test sequence and design of a decision-tree approach have a crucial significance because the results of previous steps have an influence on the following ones. Thus any misclassification at a single test/group of tests level may mislead the algorithm and lead to wrong results.

3 Probabilistic Cloud Mask (PCM) algorithm description

The main scientific motivation for the PCM algorithm development was to create a robust (meaning accurate and fast) classification method which would diminish the main sources of errors originating from the commonly used multi-spectral thresholding approach (see Sect. 2.3). Moreover it was supposed to detect snow cover and clouds at the same time and to provide classification probability between those categories. The PCM method stems from the multidimensional analysis of spectral features and ancillary data. Its parametrization is derived on the basis of training data sets composed of binary cloud and snow masks. In this way the results of two discrimination algorithms are combined and supplemented with probability estimates without the lost of classification accuracy (see Sect. 4). The described features of the PCM method are unique amongst other available techniques which substantiate the need for its development.

In the next subsections the description of the PCM algorithm will be presented in the following order: input data utilised in the study, principals of the spectral features, concept of the multidimensional information space, methodology of classification, numerical implementation and post-classification with cloud shadow estimation.

3.1 Required input data

The PCM algorithm was suited for the AVHRR instrument which has been operating aboard the suite of National Oceanic and Atmospheric Administration (NOAA) and Meteorological Operating (MetOp) polar orbiting satellites. The selection of this sensor was related to its long data record (30 yr) which serves as a valuable input to short-range climatological studies conducted by the Remote Sensing Group of the University of Bern (Hüsler et al., 2011). The proposed methodology is applicable to the AVHRR-2 (NOAA7-14) and the AVHRR-3 (NOAA15 and later, MetOp series) instruments with the spectral bands centred around: 0.6, 0.8, 1.6 or 3.7, 10.8 and 12.0 µm. For the AVHRR-3 sensors channel switching occurs at the illumination transition zone such as the 3.7 µm measurements are taken during the night while for the sunlit portion of an orbit channel 1.6 is optionally activated (for some satellites e.g. NOAA18, NOAA19 over Europe this channel is always deactivated). On the contrary, the AVHRR-2 instruments are equipped only with the 3.7 µm channel. Radiances at 1.6 µm include only the solar component of the electromagnetic spectrum whereas at 3.7 µm most of the radiation originates from the Earth's surface with a small contribution of the solar signal. Nevertheless, this reflective part was retrieved from the 3.7 µm channel by subtracting thermal component approximated by the 10.8 µm BT under the assumption of unit emissivity (Allen et al., 1990; Khlopenkov and Trishchenko, 2007).

In this study the 1 × 1 km Local Area Coverage (LAC) AVHRR measurements covering the extensive European subset (–34° W to 46° E, 28° N to 71° N) were utilized. During the algorithm training phase more than 2000 scenes with the 3.7 µm channel configuration acquired by the NOAA16, 17, 18, 19 satellites throughout the years 2009–2011 were employed whereas for the 1.6 µm channel configuration around 400 NOAA17 scenes from the year 2009 were used. As a result the set of LUTs was derived which was further utilized by the PCM procedure to classify the collection of NOAA16 images from the year 2011 and the collection of NOAA17, 18 images from the year 2008. These data sets were taken as an input to all presented analyses in this study. Due to the data availability issues time of the training dataset for the NOAA16 satellite overlays with the analysis period. Nevertheless, the results for all of the selected NOAA platforms stays in a good agreement which indicates that the overlapping period did not have a significant influence on the computed statistics.

AMTD
6, 8445–8507, 2013

Apart from the satellite measurements the following ancillary data were used: land cover obtained from the Global Land Cover 2000 Project (Mayaux et al., 2004; Bartholomé and Belward, 2005), DEM provided by the United States Geological Survey (USGS), and the Skin Temperature (SKT) derived from the ECMWF deterministic forecast with the 3 h step. All data sets were remapped to the Lambert Equal Area projection which maintains the pixel size and thus is more suitable for spatial statistics calculations.

3.2 Spectral features employed in the PCM method

In the majority of cloud/snow masking techniques the multi-spectral information is selectively exploited through the sequence of independent tests. However, when few tests give opposite results the final decision can be hard to make. In the PCM algorithm this issue is resolved by the concept of multidimensional LUT (see Sect. 3.3) which holds spectral and ancillary information together. Thus, for particular data combination there is only one possible solution which eliminates the ambiguity between different test results. The size of the LUT is a limiting factor therefore to reduce its dimensionality the Invariant Coordinate System (ICS) transformation is applied (Nordhausen et al., 2008). It utilises the Principal Component Analyses (PCA) (Mardia and JM, 1979) and two scatter matrices in order to construct independent components which do not rely on a distribution mean. The first scatter matrix is a regular covariance matrix while the second one is a matrix of the fourth moment which describes data rotation within the PCA. They are derived on the basis of a randomly selected winter satellite scene with vast snow cover and utilised throughout the rest of transformations. The ICS technique is performed for the daytime data to combine reflectances with the thermal contrast between SKT and the 10.8 µm BT. It is applied selectively only to pixels with probable cloud contamination (high thermal contrast) which fulfil specific criteria. Thus over water bodies (areas further than 8 km from the shoreline) it is performed for pixels with the SKT-10.8 µm greater than 8 K to account for warm ocean currents not included in the SKT data. For the shoreline zones it is applied to pixels with the 0.6 µm reflectance higher than 0.3 to account for mixed land/water pixels. Furthermore, regions below 1200 m are only considered if the SKT-10.8 µm is greater than 8 K. For higher altitudes this threshold is set to 16 K to account for the local thermal variations which cannot be resolved by a coarse resolution climate model. Over land, pixels which are unlikely to be overcasted with the reflectance lower than 0.15 at 0.6 µm or with the 10.8 µm BT greater than 290 K are not considered. These restriction are meant to improve ice cloud detection over snow and broken cloud discrimination where the spectral information is ambiguous but thermal contrast with surface is significant. After the ICS transformation the size of the LUT is reduced by the SKT dimension. Moreover, if it is not available the PCM algorithm can still proceed with not enhanced reflectances. A short overview on the spectral features employed in the PCM is presented in the next Subsections.

3.2.1 First enhanced spectral feature

The first enhanced spectral feature consists of two components depending on the time of the day. For the sunlit portion of a scene limited by the Sun Zenith Angle (SZA) of 89° it is composed of the second invariant component of the ICS transformation based on the SKT-10.8 µm temperature difference and on the reflectance at 0.6 µm over land and at 0.8 µm over water. This results in the enhanced spectral contrast because thin cirrus or cold sub-pixel clouds modify thermal signal more efficiently than the short-wave radiation (see Fig. 1). For the night part of a scene (SZA $\geq$ 89°) this feature consists of the scaled SKT-10.8 µm difference.

3.2.2 Second enhanced spectral feature

The content of the second enhanced spectral feature depends on the channel configuration of the AVHRR sensor and on the time of the day.

- channel 1.6 µm: whenever channel 1.6 µm is activated over the sunlit portion of a scene this feature consists of the second invariant component of the ICS transformation based on the SKT-10.8 µm temperature difference combined with the

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
AMTD
6, 8445–8507, 2013
Probabilistic approach to cloud and snow detection on AVHRR imagery
J. P. Musial et al.

reflectance at 1.6 µm. In this situation the thermal difference improves discrimination between ice cloud and snow which have very similar spectral signature at this wavelength. However, due to the fact that ice clouds have usually lower temperatures than the snow covered land the SKT-10.8 µm thermal contrast is often higher for clouds than for cloud-free snow covered areas. For the night time part of a scene (SZA ≥ 89°) this feature does not contain any information.

– channel 3.7 µm: whenever the 3.7 µm channel is activated, the second enhanced spectral feature is divided into two parts. The first one is computed for the SZA below 85° and it consists of the second invariant component of the ICS transformation based on the reflective part of the 3.7 µm spectrum and the SKT-10.8 µm difference. A combination of the reflectance with the thermal difference information enhances the ice cloud/snow discrimination in the same manner as described before (see Fig. 1). The 3.7 µm reflectance for the SZA > 85° has a weak signal and contains a lot of noise. Thus the scaled 10.8–3.7 µm BTD is used for the twilight conditions (85° ≥ SZA < 89°) and for the night-time.

3.2.3 Third spectral feature

The third spectral feature is not incorporated in the ICS transformation and it is composed of the 10.8–12.0 µm BTD which is useful for the thin cirrus cloud detection.

3.3 Multidimensional information space

The main concept of the PCM algorithm is based on the LUT with precomputed classification probability estimates in a form of an array composed of multiple dimensions corresponding to different spectral features, land cover classes and angular conditions. This array is called multidimensional information space because it holds spectral and ancillary information together as opposed to the spectral space which contains only satellite measurements. In order to include continuous data such as spectral features or viewing angles in the array they are binned by an extensive set of thresholds. Such

AMTD
6, 8445–8507, 2013

an approach allows specifying multiple, irregularly distributed threshold values per feature which resolves the problem of a single threshold estimation existing in the multi-spectral thresholding methods. Another advantage of the PCM algorithm is the direct derivation of probabilities from the LUT by means of all available spectral and ancillary information. Therefore all data are analysed at once in contrast to the decision-tree algorithms where features/tests are organised in a sequence.

The information space is composed of 8 dimensions:

- *Time of the day* divided into: day (SZA $< 85^{\circ}$), twilight (SZA $\geq 85^{\circ}$ and SZA $< 89^{\circ}$) and night (SZA $\geq 89^{\circ}$).
- *First enhanced spectral feature* divided by the following thresholds: 0, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25, 0.275, 0.30, 0.35, 0.40, 0.45, 0.50, 0.60, 0.80, 1.0.
- *Second enhanced spectral feature*:
  - 1.6 µm channel activated: this feature is divided by the following thresholds: 0, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25, 0.30, 0.40, 0.50, 0.6, 0.7.
  - 3.7 µm channel activated: this feature is divided by the following thresholds: 0, 0.025, 0.05, 0.075, 0.1, 0.125, 0.15, 0.175, 0.2, 0.225, 0.25, 0.30, 0.40, 0.50.
- *Third spectral feature* divided by the following thresholds: −0.8, −0.6, −0.4, −0.2, 0.0, 0.1, 0.4, 0.6, 0.8, 1.0, 1.25, 1.5, 1.6, 1.8, 2.0, 2.5, 3.0, 3.5.
- *Texture feature* which is derived only over open waters by convolution of the 0.8 µm reflectance for the sunlit areas and the 10.8 µm BT during the night-time with a 3×3 kernel (Eq. 1). The convolution results are scaled to 0–1 range where high values characterise pixels located at the cloud edges. This feature is divided by the following thresholds: 0.05, 0.1, 0.25, 0.5, 1.

$$\mathbf{K} = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} \tag{1}$$

- *Sensor viewing sectors* which are derived by dividing the satellite zenith angle with the following thresholds: 0, 15, 30, 45, 55, 70.
- *Relative azimuth sectors* which are derived by dividing the relative azimuth angle, defined as a difference between sun and sensor azimuth angles (180° = forward scattering), with the following thresholds: 0, 45, 90, 135, 180.
- *Land cover/use* developed within the scope of the Global Land Cover 2000 (GLC2000) project (Bartholomé and Belward, 2005). There are three additional surface categories which are derived internally by the PCM algorithm: coastline water defined as a 8 km buffer zone from the shore including inland waters, sun glint over water and sun glint over desert which are discriminated by the simple sun/sensor angular dependency described by Ackerman et al. (1998). All these areas feature significantly higher reflectances due to: specific angular conditions (sun glint); higher concentration of non-maritime aerosols (Wang and Shi, 2006), sediments or algae (Wang and Shi, 2005) and shoreline variation induced naturally by tides or artificially by satellite geolocation problems.

The positions of values within the information space depends on diurnal and annual cycles. First one is driven by the illumination conditions and alters mainly the reflectance due to bidirectional effects. Therefore separate information spaces were developed for different satellite overpass times. For morning satellites with the 1.6 µm channel activated the acquisition time was divided into two ranges: 00:00–12:00, 13:00–00:00 h UTC whereas for the satellites with the 3.7 µm channel activated the division was set to: 07:00–10:00, 11:00–14:00, 15:00–18:00, 19:00–06:00h UTC. The annual cycle is related to changes in albedo and surface emissivity properties induced

AMTD
6, 8445–8507, 2013

by the vegetation development and soil moisture variations. Thus the informations spaces were further divided between different seasons: winter (November–January), spring (February–April), summer (May–July), autumn (August–October). Each possible daytime/season combination was stored separately for the 1.6 and 3.7 µm channels configurations which results in 48 LUTs.

3.4 Methodology of classification

In order to discriminate snow, clear-sky and cloudy pixels on a satellite image with the PCM algorithm the information space has to be filled with classification probability estimates. This process is divided into 2 steps. First, a temporary information space is created which contains one more dimension where the categories: clear-sky, cloudy and snow originating from the training dataset are stored. Second, the frequencies of occurrence of those classes within each bin are transformed into the probability estimates. The development of the PCM method consists of the following steps:

1. *Composition of training dataset* which involves collocation of the AVHRR measurements with the snow/clear-sky/cloudy classification which may originate from ground data, supplementary satellite data source (e.g. CALIPSO, MODIS products) or from results of another classification algorithm applied to the AVHRR data. Nevertheless, the data quantity should be large enough to sufficiently sample the multidimensional array with information acquired under a wide range of environmental conditions. In this study the training data set was composed of the PPS version 2012 cloud masks (Dybbroe et al., 2005a) retrieved from AVHRR data and merged with the MOD10A1 daily snow masks (Hall et al., 2002) based on MODIS data. Those algorithms were selected due to their high classification accuracy (Dybbroe et al., 2005b; Hall and Riggs, 2007) as well as the data (MOD10A1) and source code (PPS) availability. More comprehensive comparison of available cloud/snow classification techniques is beyond the scope of this study. However, it has to be emphasized that the mean PCM classification accuracy

AMTD
6, 8445–8507, 2013

cannot exceed the one associated with the training dataset. To improve the quality of the training dataset it was first visually inspected and low quality scenes were removed. Then significant misclassifications occurring in the remaining images were corrected by means of the supervised classification. This procedure was applied to the extensive data set (∼2400 scenes) acquired over Europe between 2009–2011 by different AVHRR sensors mounted on board NOAA satellites denoted with numbers: 16, 17, 18, 19.

2. *Formation of temporary information spaces* which involves computation of all spectral or textural features for the selected AVHRR training images. Further the frequencies of occurrence of clear-sky, snow and cloudy classes originating from the training data for each combination of dimensions (i.e. features, angular and ancillary data) are inserted into the information space. Therefore a single value within the temporary information space corresponds to numerical count of all pixels embedded into a 9-dimensional bin composed of the following elements: time of the day, three spectral features, texture feature, viewing and azimuth sectors, land cover, PPS/MOD10A classification. In other words this array might be treated as a 9-dimensional histogram. This procedure was repeated 48 times for each information space characterised by the different channel configuration (1.6/3.7 µm), acquisition hour and season.

3. *Derivation of classification probability* which is based on numerical counts from the temporary information spaces which are first rearranged in the descending order for every multidimensional bin. Further the probability is computed as a simple ratio between counts of the most frequent classification category within a bin and the total number of counts in this bin. The retrieved value is described as a classification probability between the two most frequent classes within the bin. Similar analysis in a two-dimensional space is presented in Fig. 2. The PCM classification contains probability estimates for the following combination of categories: snow-free/snow, clear-sky/cloudy, snow/cloudy. For the sake of visualisation obtained

AMTD
6, 8445–8507, 2013

values are recoded according to Eq. (2). For the bins which were not filled by the training dataset probability estimates are retrieved by the nearest neighbour interpolation of existing values within the array. At the last stage all information spaces containing probability estimates are compressed and stored in the NetCDF4 format as the LUTs.

$$P \in \begin{cases} 0 \geq P \leq 100 \text{ probability of 0–100 \% between clear-sky and snow conditions} \\ 100 \geq P \leq 200 \text{ probability of 0–100 \% between snow and cloudy conditions} \\ 200 \geq P \leq 300 \text{ probability of 0–100 \% between cloudy and clear-sky conditions} \end{cases} \quad (2)$$

4. *Classification of a satellite image* is a procedure very similar to the construction of temporary information spaces because it requires preparation of the same spectral and ancillary data (binned values of spectral features, angles, etc.). However, further this information is directly used to retrieve probability estimates from a LUT. The bottleneck of this process is related to localisation of all input data associated with a large satellite scene within a LUT which itself has more than 60 million values. This issue is resolved by the fast Approximate Near Neighbour (ANN) searching method (Arya et al., 1998) which is performed for each dimension separately.

3.5 PCM numerical implementation

The PCM algorithm has been implemented in the R statistical language (Team, 2012) and the source code with help files and test cases are available on the web: http://pcm.r-forge.r-project.org/. The flow-chart of the method is depicted in Fig. 3. The input data sets to the PCM method require few pre-processing steps which consist of: discrimination of shallow water defined as a 8 km buffer zone from a shore, Earth distance correction of reflectances, and up-scaling of the SKT estimates. The last process involves bi-linear and temporal interpolation of two SKT estimates closest to the satellite overpass time. They are usually of much coarser spatial resolution than the AVHRR grid and correspond to certain hours. However, the up-scaled data do not resolve well the temperature variation in the rough topography terrains such as high mountains. To account for this effect a constant, temperature correction is derived by multiplying the lapse rate of 0.6 K/100 m with the altitude difference between the 1 × 1 km Digital Elevation Model (DEM) provided by the USGS and the bi-linearly up-scaled 0.2° × 0.2° DEM (geopotential at surface) originating from the ECMWF model. This correction is added to the final SKT estimates.

At the first stage of the PCM algorithm the required data: AVHRR reflectances and brightness temperatures, View Zenith Angle (VZA), Sun Zenith Angle (SZA), Relative Azimuth (RAZ), land cover, upscaled Skin Tempaerature (SKT) and the Digital Elevation Model (DEM) are ingested and no-data values are removed. Further, several processes are performed:

- division between day, twilight and night (see Sect. 3.3).
- derivation of reflectance from the 3.7 µm BT.
- identification of sun glint areas (after Ackerman et al., 1998) over water and desert which are incorporated into the land cover data.
- kernel convolution and derivation of the texture feature from the 0.6 and 10.8 µm channels.

Further reflectances and the thermal difference between SKT and 10.8 µm channel are fetched into the Invariant Coordinate System (ICS) transformation separately suited for the 1.6/3.7 µm channels. For the daytime data its results consist of two enhanced spectral features where first one together with the 10.8–12.0 µm BTD are the same for both channel configurations. For the night-time the first two features contain scaled SKT-10.8 µm and 3.7–10.8 µm thermal differences respectively. Such processed spectral information, together with the texture feature, acquisition angles and expanded land cover data are located within the threshold values (see Sect. 3.3) by means of the ANN procedure. This leads to determination of indexes along each of the eight dimensions of the information space which are merged into two separate tables depending on the

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

3a/3b channel availability. Finally the index tables are used to retrieve classification probability estimates from the LUTs for every pixel in the satellite scene.

3.6 Post-classification and cloud shadow estimation

The majority of geophysical analyses require a binary cloud and/or snow classification thus the continuous PCM probability estimates are recoded to the discrete classes assuming clear-sky pixels for the value ranges from 0 to < 50 and from > 250 to 300; cloudy pixels for the > 150 to ≤ 250 range and snow pixels from ≥ 50 to ≤ 150. Complementary to the acquired results the cloud shadow mask is computed on a basis of simple geometrical relationships (Fig. 4). This procedure consists of several steps: first, a rough approximation of the cloud height h is estimated from the SKT-10.8 µm thermal contrast assuming the constant temperature lapse rate of 0.6 K/100 m. Second, the acquired altitude values are processed by the maximum value filter with a 5×5 window to enhance reliability of the estimates at the cloud edges where the thermal contrast is smaller due to fractional cloud cover. Further, the computations are performed only for pixels located at cloud edges and initially involve calculation of a shadow length L expressed as cloud height divided by the tangent of sun elevation angle (90°-SZA). Next the Sun Azimuth angle (SAZ) is reduced to 0–90° range (SAZ_{red}) and the X_{scale}, Y_{scale} are derived according to convention described by Eq. (3). Finally the end of the shadow (x_2, y_2) is estimated on the basis of known coordinates of the cloud edge pixel centre (x_1, y_1) and displacement vectors along x and y axis computed as a ratio of shadow length and sine and cosine of SAZ_{red} respectively (Eq. 4).

After the derivation of a cloud shadow mask it is incorporated together with the land cover data into the PCM binary output to compose a discrete classification which consists of the following categories: no data, clear-sky water, clear-sky land, clear-sky snow, pixel adjacent to cloud over water, pixel adjacent to cloud over land, pixel adjacent to cloud over snow, cloud shadow over water, cloud shadow over land, cloud shadow over snow, cloud. An example of the PCM output (Fig. 5) presents the NOAA17

AMTD
6, 8445–8507, 2013

AVHRR scene acquired over the Alps on the 1st January 2008 together with the probability estimates and discrete classification mask.

$$\mathrm{SAZ_{red}} = \begin{cases} \mathrm{SAZ} & \in 0 \geq \mathrm{SAZ} \leq 90; \quad X_{\mathrm{scale}} = -1;\ Y_{\mathrm{scale}} = -1 \\ \mathrm{SAZ} - 90 & \in 90 > \mathrm{SAZ} \leq 180; \quad X_{\mathrm{scale}} = -1;\ Y_{\mathrm{scale}} = 1 \\ \mathrm{SAZ} - 180 & \in 180 > \mathrm{SAZ} \leq 270; \quad X_{\mathrm{scale}} = 1;\ Y_{\mathrm{scale}} = 1 \\ \mathrm{SAZ} - 270 & \in 270 > \mathrm{SAZ} \leq 360; \quad X_{\mathrm{scale}} = 1;\ Y_{\mathrm{scale}} = -1 \end{cases} \tag{3}$$

$$\begin{aligned} x_2 &= L \times \sin\left(\mathrm{SAZ_{red}}\right) \times X_{\mathrm{scale}} + x_1 \\ y_2 &= L \times \cos\left(\mathrm{SAZ_{red}}\right) \times Y_{\mathrm{scale}} + y_1 \end{aligned} \tag{4}$$

3.7 Limitations of the PCM algorithm

Regardless the advantages of the PCM method such as: determination of the classification probability for clear-sky, snow and cloudy conditions; automatic algorithm training phase there are some limitations which originate from the proposed methodology of probability derivation. Although, the algorithm provides a great flexibility over other classification methods in terms of values and number of thresholds their choice is still matter of subjective decision. Nevertheless, it is possible to determine an arbitrary number of regularly distributed thresholds for each feature and still the algorithm will feature considerable detection skills. The approach applied in this study assumes the higher density of thresholds around the value ranges associated with uncertain pixels (e.g. low reflectance values related to broken/cirrus clouds). Furthermore threshold quantity modifies the probability values as wider bins within the information space are more likely to contain a mixture of classes in comparison to smaller bins. Thus a high threshold density assures better classification accuracy at the price of bigger LUT size which itself is a limiting factor.

The quality of the PCM results mostly depends on the accuracy of clear-sky/snow/cloud mask used during the algorithm training step. Moreover the quantity of this data set should be large enough to sufficiently sample the multidimensional array with information acquired under a wide range of environmental conditions. Although,

AMTD
6, 8445–8507, 2013

in this study a broad set of ecosystems was considered (from deserts to tundra and perennial ice) the performance of the method over other regions than Europe is still to be determined as well as the applicability to older AVHRR sensor (prior to NOAA16).

4 PCM cloud and snow classification results

The accuracy assessment of the PCM algorithm was divided into 4 parts where first the results were compared to the original PPS cloud mask version 2012 in order to verify the difference between training data set and the acquired output. Then the MODIS cloud mask (MOD/MYD35 collection 6 products) and daily snow cover composite (MOD10A1 collection 5 product) were used as a reference to compare the PCM results with the other robust classification algorithms suited for modern sensor with higher radiometric resolution than AVHRR. Further, the PCM output was validated against ground-truth data originating from more than 4000 SYNOP stations (both automatically and manually operated) which provide detailed cloudiness observations. Finally, it was validated against the vertical feature mask version 3 derived from the CALIPSO/CALIOP (Cloud-Aerosol Lidar with Orthogonal Polarization) lidar measurements. All of the mentioned data sets feature a discrete form therefore the continuous PCM probability estimates coded according to the convention described in Sect. 3.4 were transformed to a binary form. This was achieved assuming the clear-sky pixels for the coded probability values between 0–50 and 250–300; cloudy pixels for the 150–250 range and snow pixels for the 50–100 range.

4.1 Inter-comparison with the PPS cloud mask

In order to evaluate discrepancies between the PCM and PPS algorithms the total cloud cover estimates expressed as $100\,\% \times \frac{\text{cloudy pixels}}{\text{total pixel count}}$ over each satellite scene were computed. The acquired results for both methods were then linearly regressed and the basic statistics of the distribution of differences were derived (Fig. 6). It was

AMTD
6, 8445–8507, 2013

found that both methods are in good agreement regardless the image acquisition time being morning satellites NOAA16, 17 and the afternoon one NOAA18 as well as the channel configuration: 1.6 µm for NOAA17 and 3.7 µm for NOAA16,18. The total cloud cover differences between PCM and PPS outputs are low −0.5–1.0 % at average with standard deviations not exceeding 3.6 % and with correlations ≥ 0.93.

Further relationships between the PCM-PPS differences and angular/thermal conditions (Fig. 7) as well as land cover (Fig. 8) were investigated. This involved derivation of annual histograms (grey shading in Fig. 7) for the selected variables from all of the available images and computation of total cloud cover within each bin. As far as the sun zenith angle is concerned PCM reports more clouds for higher angles than PPS. This relation starts to decrease again above the angle of 85° which determines the twilight conditions. For the satellite zenith angle the PCM-PPS difference distribution features characteristic "saw-shape" pattern which originates from the dependency of the probability estimates on the angular sectors. The variation of discrepancies increases towards oblique angles but the values stay within −2–2 % range. On the contrary, the relative azimuth angle between Sun and a satellite does not influence the difference in a systematic way. It seems that most of the variation can be associated with the number of observations acquired at specific angular conditions. The relationship between the PCM-PPS total cloud cover differences and the thermal contrast between SKT and the 10.8 µm channel features a distinct pattern. For negative values (SKT lower than 10.8 µm) associated with a temperature inversion or with underestimated SKT prediction the PCM method tends to report more clouds than PPS (up to 8 %). This situation often occur during the night over sparsely vegetated areas (Sahara desert, Spain) where climate models poorly represent rapid radiative cooling which lead to erroneous temperature inversions. This changes the PPS parametrization and probably result in lower total cloud cover (Fig. 10). For small positive thermal contrast 0–5 K PPS reports more clouds but just beyond this range the relation changes rapidly and around 10 K there is a positive peak where PCM detects more clouds. Above this value the total cloud cover differences are getting smaller as high thermal contrast between the

AMTD
6, 8445–8507, 2013

surface and satellite measurements is usually associated with the overcasted conditions. Considering the land cover classes (Fig. 8) PPS method reports more clouds over water bodies (up to 2 %) and desert (up to 5 %) while PCM detects more clouds over land (up to 2.5 %) especially over perennial snow/ice (up to 5 %) but only for satellites without the 1.6 µm channel activated. Over the vegetated areas the differences are the most pronounced (up to 4 %) for the needle-leave forest class which covers high latitudes zones and mountainous regions. For the broad-leave forest the discrepancies are slightly lower however they are still 1–1.5 % higher than for crop-lands or shrubs. This may indicate that the PCM method is more sensitive to cloud cover over dark dense vegetation in comparison to PPS.

The next step of conducted analysis involved computation of annual total cloud cover composites over Europe separately for the day and night-time conditions in order to investigate distinct spatial patterns. The results presented on Figs. 9 and 10 for the NOAA18 satellite reveal some artefacts on the PPS image i.e. areas with unreliably high cloud-coverage over desert or Spain which are significantly less pronounced on the PCM composition. These areas are related to land cover categories distinguished by the USGS classification (Anderson, 1976) and utilized by the PPS which do not reflect well the local spectral properties of the surface. On the other hand PCM takes as input GLC2000 classification (Bartholomé and Belward, 2005) which discriminates more categories with more uniform spectral signature and therefore does not lead to cloud overestimation. Another problematic regions are associated with mountains where PPS reports less clouds especially during the night-time due to the cloud conservative threshold (–22 K) in the 10.8 µm-SKT test. This effect is particularly noticeable over Norway and the Alps. Nevertheless, most of the spatial cloud coverage patterns visible on the PPS image are also present in the PCM composite which confirms a good agreement between these two methods.

Finally the spatio-temporal aspect of the PCM-PPS total cloud cover discrepancies was analysed as a function of latitude and time (Fig. 11). For all selected satellites and both years 2008, 2011 it could be seen that the temporal pattern of differences does not

AMTD
6, 8445–8507, 2013

change over the coarse of the year. In comparison to the PPS method PCM tends to detect more clouds for the latitudes between 30–40° and 50–60° while for the latitudes between 40–50° and 60–70° it reports less of them. This pattern corresponds to the spatial distribution of land and water over Europe where PCM reports more clouds over land whereas PPS reports more of them over water (Fig. 8).

4.2 Inter-comparison with MODIS cloud and snow mask

The PCM classification consists of the cloud and snow components thus to assess its accuracy two types of products derived from the MODIS data were used as reference. First one – MOD35/MYD35 – consists of a cloud mask (Ackerman et al., 1998) generated from a 5 min swath segment recorded by the TERRA and AQUA satellites. In order to compare the cloud masks derived from the MODIS and AVHRR data, image acquisition times had to be collocated as the cloud cover dynamics requires a small interval between the observations. In this study the maximum time difference between observations was set to 15 min. Out of the three selected NOAA platforms only the one labelled with number 18 has a sufficiently close orbit to the TERRA and AQUA satellites to perform such a comparison. Over Europe for the year 2008 after the image collocations the joined AVHRR/MODIS data set consisted of 118 TERRA scenes and 238 AQUA scenes for both day and night conditions. Furthermore, the MODIS product was modified to exclude the uncertain pixel category and to reclassify the confident/probably clear categories to clear-sky class as well as the confident/probably cloudy categories to overcasted class.

Regarding the MOD/MYD35 comparison, the analyses were performed analogically to the ones described in the previous subsection. First the total cloud cover over each scene was computed for both cloud masking algorithms to compose a scatter plot (Fig. 12a). It occurs that on average the MODIS product reports 4.4 % more clouds than PCM remaining at the same time with the high correlation of 0.95. Further the PCM-MODIS total cloud cover discrepancies were analysed as a function of land cover, sun zenith angle as well as the SKT-10.8 µm thermal contrast (Fig. 12b, c and d). The results

AMTD
6, 8445–8507, 2013

show that the PCM constantly reports less clouds regardless the land cover class with the highest difference for the perennial snow/ice areas (up to 12 %). As far as the sun zenith angle is concerned the absolute differences increase almost linearly and the best agreement between the methods (±2.5 %) is reported for the SZA < 55°. On the other hand for the twilight conditions (SZA > 85°) total cloud cover discrepancies between PPS and MOD35 reach 10 %. Their relationship with the SKT-10.8 µm thermal contrast (Fig. 12d) resembles the one from the PPS comparison. When the surface temperature is lower than satellite measurements PCM reports more clouds (up to 5 %). For the thermal difference between 0–5 K range the PCM-MOD10A absolute discrepancies increases (up to 18 %), then they diminish rapidly and above the 10 K both algorithms reports the same amount of clouds.

The MOD10A1 product (Hall et al., 2002) utilised as a reference for the PCM classification consists of daily snow cover composites retrieved from the TERRA satellite data. As it is one of the few globally available snow cover data with high spatial and temporal resolutions in this study it was regarded as a ground-truth information. On the basis of the MOD10A1 and PCM products the contingency table was constructed (Eq. 5) which allows computation of the following classification quality indicators (Eqs. 6–11): Probability Of Detection (POD), False Alarm Rate (FAR), Hit Rate (HR) and Kuiper's Skill Score (KSS). The acquired results indicates that the difference in estimated snow cover between the PCM and MOD10A algorithms is significant (Fig. 14). Considering the PCM snow detection skills the highest POD ≈ 0.6 is reported for the NOAA17 which has the 1.6 µm channel activated. On the other hand the AVHRR sensors with the 3.7 µm channel constantly operating aboard NOAA16, 18 satellites feature lower snow detection skills POD ≈ 0.4. The extremely high POD and low FAR for the clear conditions as well as high HR are related to the fact that clear conditions are predominant over the year in comparison to the snow cover. As a result the proportion between these categories in the contingency table is uneven and voids the usability of these indicator. In such a case the KSS indicator allows estimating the objective overall algorithm performance which clearly indicates higher snow detection accuracy for the

AMTD
6, 8445–8507, 2013

1.6 µm channel configuration. Further the total snow cover scatter plots, derived analogically to the ones from the cloud coverage analyses, were computed (Fig. 13d, h and l). The obtained results show that PCM reports significantly less snow cover (up to 9.5 % at average) than MOD10A1 with a moderate correlation coefficient (0.5–0.7) and a high standard deviation (up to 17 %). The absolute PCM-MOD10A1 total snow cover difference rises with the sun zenith angle up to 85° where it slightly decreases due to separate treatment of the twilight conditions (Fig. 13a, e and i). In turn the relationship with the satellite zenith angle indicates that PCM reports more snow at high angles possibly due to increasing instrument Field Of View (FOV) but only for the NOAA16, 17 platforms. The opposite trend for the NOAA18 satellite is not easily explainable and it might be related to geographic distribution of snow in Europe which is the most persistent over the Alps and Scandinavia. These regions are located around the centre of a swath of the NOAA18 satellite overpass thus the amount of snow pixels across the scan is not even. The PCM-MOD10A1 total snow cover differences vary with the relative azimuth angle according to a pattern mostly related to quantity of the available observations. Nevertheless, for the NOAA18 there is a significant increase of discrepancies around the azimuth angle of 60° (towards forward scattering) which again might be associated with combined geographical and angular conditions rather than the algorithm's configuration.

counts	PCM clear	PCM snow	
MODIS clear	a	b	(5)
MODIS snow	c	d	

$$\mathrm{POD_{snow}} = \frac{d}{c+d} \tag{6}$$

$$\mathrm{POD_{clear}} = \frac{a}{a+b} \tag{7}$$

$$\mathrm{FAR_{snow}} = \frac{b}{b+d} \tag{8}$$

AMTD
6, 8445–8507, 2013

$$\text{FAR}_{\text{clear}} = \frac{c}{a+c} \tag{9}$$

$$\text{HR} = \frac{a+d}{a+b+c+d} \text{ where } 0 \geq \text{HR} \leq 1 \tag{10}$$

$$\text{KSS} = \frac{a \times d - c \times b}{(a+b) \times (c+d)} \text{ where } -1 \geq \text{KSS} \leq 1 \tag{11}$$

4.3 Validation against SYNOP weather reports

The global network of SYNOP stations (Fig. 16a) provides detailed cloud cover observations acquired in automatic or manual way at high temporal resolution (up to one hour). They are collected in a consistent manner during day and night for a long period which makes them an excellent validation source for the satellite based products. In this study SYNOP total cloud cover observations expressed in octants were compared to the PCM cloud probability which was recoded to 0–100 % range where 0 % denotes confident clear/snow class and 100 % indicates confident cloudy category. In order to maintain the coherence between these two data sets the PCM outputs in binary form (clear/cloudy) as well as the cloud probability were averaged over the 30 km buffer zone around each station. This area should correspond to the sky extent usually visible to an observer who performs the measurements (Karlsson, 1995). Analogically to the comparison with the MOD/MYD35 data, the satellite and ground cloud observations were collocated within the 15 min time interval. Next the mean PCM cloud probability was computed as functions of the SYNOP cloud amount and daytime from all of the collocated stations and satellites in the selected years. (Fig. 16b–d). The acquired probability distributions span from around 12 % for completely clear-sky SYNOP observations to 90 % for fully overcasted conditions regardless the daytime. Furthermore, the histograms of annual normalised probability distributions for clear, cloudy and snow conditions were computed (Fig. 15). Their shape very well corresponds to other probabilistic cloud mask distributions enclosed in the studies of Gómez-Chova et al. (2007) and Heidinger et al. (2012).

AMTD
6, 8445–8507, 2013

In order to compute the quality indicators (Eqs. 5–11 with $_{\text{cloudy}}$ instead of $_{\text{snow}}$) the PCM and PPS cloud fractions derived over the 30 km buffer zone together with the SYNOP observations were transformed to a binary form assuming overcasted conditions over 25 %/2 oktants cloud amount. The obtained results proved fine cloud detection skills of both algorithms (high POD, HR, KSS and low FAR) without any substantial differences between them. The overall accuracy during the night is a bit lower than during the day which is to be expected provided the availability of the thermal information only. Considering the selected satellites the NOAA17 platform with 1.6 µm configuration provides the most reliable cloud coverage data during the day whereas for the night time the best results are obtained by the NOAA18 platform. Nevertheless, the differences in cloud detection skills between different satellites are small for both PCM and PPS algorithms.

4.4 Validation against CALIPSO/CALIOP lidar feature mask

The last part of the performed analysis involved validation of the PCM and PPS cloud masks against CALIOP vertical feature mask (CAL_LID_L2_VFM). The CALIOP lidar is an active system which is able to detect even very thin cloud layers with high horizontal resolution of 333 m. This makes it an extremely valuable source for validation of any medium resolution satellite cloud mask (Karlsson and Dybbroe, 2010). Nevertheless, before such a comparison is possible satellite observations have to be collocated in the space and in the time domains. In this study the approach applied to this issue consists of three steps. At first, CALIOP observations with a time interval lower than 15 min as compared to the AVHRR acquisition time were selected. Next, the vertical profile containing the feature mask was transformed to a vector with a binary flag present whenever a cloud layer was reported within an atmospheric column. Finally, the location of acquired samples within the 1 × 1 km AVHRR grid was determined by means of the nearest neighbour technique. The CALIPSO satellite share the same orbit as AQUA platform with the MODIS instrument thus according to the arguments given in Sect. 4.2 only comparison with the NOAA18 satellite was possible.

AMTD
6, 8445–8507, 2013

From the collocated CALIOP/NOAA18 observations the quality indicators of the PCM and PPS cloud detection skills were computed. The acquired results (Fig. 18) confirm the good performance of both algorithms with slightly better indicators in case of the PCM algorithm. The absolute values of computed quality indicators slightly differ from those presented in the AVHRR/SYNOP validation, however the relationships between them remain the same. The POD for cloudy conditions is lower by around 0.05 for the nigh-time acquisitions in comparison to the day ones while the FAR features a reverse trend. This in turn results in higher FAR for the clear conditions during night. Nevertheless, all of those differences are small thus overall cloud detection accuracy for the PCM and PPS algorithms expressed by the KSS indicator is almost the same for day and night conditions.

5 Discussion

The performed analyses aimed at assessing the accuracy of the PCM algorithm from the perspective of cloud and snow detection skills. Considering that all of the reference data sets feature discrete form, the provided probability estimates between: snow-free/snow, snow/cloudy and cloudy/clear-sky conditions were transformed to binary classes using thresholds described in the Sect. 4. Such a prepared data set was then compared to the PPS and MODIS cloud masks, MOD10A1 snow mask, CALIOP vertical feature mask and SYNOP total cloud amount observations. As far as the classification probability is concerned the lack of an appropriate reference data set allows only a basic comparison of the PCM cloud probability against the SYNOP total cloud amount measurements. Moreover, computed normalised probability distribution for cloudy conditions (Fig. 15) is consistent with the findings of Gómez-Chova et al. (2007) and Heidinger et al. (2012).

The acquired results prove high cloud detection skills of the PCM method by reporting a good agreement with the reference data sets. On average PCM detects 4.4 % less clouds than MOD/MYD35 products (Fig. 12a) while in case of the PPS cloud mask the

AMTD
6, 8445–8507, 2013

differences are even lower from −0.5 to 0.9 % (Fig. 6). Similar results were reported by Heidinger et al. (2002) who found that for the 3 selected satellite scenes MODIS cloud masks contained 1–3 % more clouds than the CLAVR (Stowe et al., 1999) cloud mask derived from the LAC AVHRR data. These negative discrepancies decrease with the SZA for the MOD35 data (Fig. 12c) whereas the opposite trend is associated with the PCM-PPS differences which eventually become positive (Fig. 7a, e and i). This inconsistency might be partially related to the fact that MODIS cloud mask reports more clouds over snow than PCM (Fig. 12b) while PPS reports less (Fig. 8). Therefore, during the cold season associated with high SZA the absolute PCM-MOD35 difference is increasing. Regarding the sensor viewing geometry the comparison against MODIS data is not conclusive as both instruments (AVHRR and MODIS) measure the same area under different angles. Therefore such an analysis was performed only with the PPS algorithm which relies on the same AVHRR input data as the PCM. The distribution of total cloud cover differences as a function of satellite zenith angle features a characteristic "saw" shape pattern (Fig. 7b, f and j) which originates from the division of angles values into sectors. It can be presumed that without the consideration of this variable the differences between PCM and PPS methods would rise continuously and the overall discrepancies between the algorithms would be higher. The relationship of the differences with the sun-satellite relative azimuth angle does not reveal any significant pattern and the noticeable variations might be attributed to quantity of the available observations (Fig. 7c, g and k). On the contrary, the distribution of total cloud cover discrepancies related to the thermal contrast between SKT and the 10.8 µm BT has a distinct drop of values within the 0–10 K range for both PPS and MODIS comparisons (Figs. 7b, f, j and 12d). This effect could be partially associated with the different approach towards image texture analysis over water bodies where PPS and MOD35 mark a pixel as cloudy whenever the spectral variance in its vicinity (e.g. 5 × 5 window) is sufficiently high (Di Vittorio and Emery, 2002). However, this implies that some of the clear-sky pixels around cloud edges with small thermal contrast are misclassified In the PCM method the textural analysis is performed differently by means of

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

AMTD

6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

the kernel convolution with a small window (Eq. 1) which assures that only the the pixels at cloud edges are detected. Therefore, the PPS and MOD35 products report more clouds over water bodies than PCM (Figs. 8 and 12b) which in turn modifies the overall results significantly as it is the most extensive land cover class. This is also confirmed by the spatio-temporal analysis of the total cloud cover differences between the PCM and PPS methods (Fig. 11). For the latitude ranges between 40°–50° and 60°–70° where the water class is predominant over the study area the PPS algorithm reports more clouds than PCM regardless the time of the year. For other surfaces the total cloud cover difference remains negative in case of comparison with the MODIS data whereas for the PPS comparison PCM tends to report more clouds over the land surfaces excluding deserts. The same relationships could be observed on the annual total cloud cover composites for the NOAA18 satellite (Figs. 9 and 10). Moreover, the day-time PPS composite contains visible artifacts associated with areas characterised by the unreliably frequent cloud cover (up to 100 %) over Africa and Spain. They are related to land cover classes described by Anderson (1976) which do not resolve well the local spectral characteristics and lead to overestimation of cloud amount. These areas are less apparent on the PCM total cloud cover composite because the PCM algorithm utilises more detailed land cover categorisation (Bartholomé and Belward, 2005). Furthermore, the PPS-based training data set was improved by means of the supervised classification before being utlized during the PCM development phase. Another noticeable spatial differences are associated with mountains (e.g. Alps) where PCM reports more clouds than PPS especially during the night-time. This is related to the fact that in the PPS method the SKT-10.8 µm thermal contrast test features very cloud conservative threshold (22 K at night) over the rough topography areas. This was introduced to account for the temperature variability induced by the comparison of the 1 × 1 km AVHRR data with low resolution SKT estimates (Dybbroe et al., 2005a). Nevertheless, due to this relaxed threshold the SKT-10.8 µm test often indicates clear-sky conditions which misleads the PPS algorithm and results in lower cloud coverage over the mountainous regions.

AMTD
6, 8445–8507, 2013

In order to evaluate the accuracy of the PCM and PPS algorithms the total cloud cover observations provided by the SYNOP stations (Fig. 16a) were compared to the PCM and PPS cloud fractions derived over the 30 km buffer zone around each station. This analysis involved creation of the contingency table and computation of classification quality indicators (Eqs. 5–11). They are the same for both methods or slightly better in case of the PCM results (Fig. 17). The small improvement over the PPS algorithm was gained by the supervised classification correction applied to the PPS/MOD10A1 training data set. High probability of detection (> 0.8) and low false alarm rate (< 0.1) for the cloudy conditions remain at the same level for both algorithms regardless the time of the day. Lower detection skills for the clear-sky cases can be attributed to the fact that an observer at a SYNOP station may report clouds which are further than the 30 km distance (for more comprehensive explanation see Karlsson, 1995). Nevertheless, the overall performance of both algorithms represented by the Kuiper's Skill Score is high and the differences between the selected satellites are inconsiderably small. Furthermore, these results are consistent with the CALIOP/AVHRR validation (see Sect. 4.4) from the qualitative perspective, however their absolute values are higher. This is due to the fact that computation of fractional cloud cover within the 30 km buffer zone increases the probability of cloud detection. On the other hand the CALIOP/AVHRR validation was performed on the single pixel basis which yields more accurate results. The sensitivity of a lidar system to detection of thin clouds which are not visible for the AVHRR sensor is well apparent as a relatively high FAR (> 0.4) for the clear conditions. This is also the reason for the higher cloud amount in the cloud mask derived from the MODIS instrument which has 1.38 and 8.55 µm channels that improve the detection of mid and high clouds (Ackerman et al., 1998). The acquired classification quality criteria for the CALIOP/PPS cloud mask comparison remain in a good agreement (within the 10 % range) with the study of Karlsson and Dybbroe (2010) who performed the same analysis over the Artic region.

Within the 30 km buffer zones around each station the mean PCM cloud probability was analysed as a function of the SYNOP cloud amount (Fig. 16b–d). The acquired

AMTD
6, 8445–8507, 2013

distribution feature a reasonable spread of values from ∼ 12 % for the clear-sky conditions to ∼ 90 % for the fully overcasted sky regardless the time of the day and the AVHRR channel configuration. The acquired probability distribution remains in a good consistency with the SYNOP validation results for the POLarization and Directionality of the Earth's Reflectances (POLDER) cloud mask (Bréon and Colzy, 1999). Although, in that case the distribution expressed the percentage of cloudy pixels as a function of the SYNOP cloud amount, the PCM probability estimates feature almost a binary distribution (Fig. 15) which fits well to the binary output of the POLDER cloud mask. The same characteristics of the cloud probability distribution were reported in other studies (Gómez-Chova et al., 2007; Heidinger et al., 2012). Its specific "U-shape" where most of the pixels are classified as confident clear 0 % or cloudy 100 % is related to the high spectral contrast of clouds on the satellite image as compared to other surfaces (see red dots on Fig. 2). Most of the classification probability estimates significantly different from 0 or 100 % represent: pixels with fractional cloud and/or snow coverage, ice clouds overlaying a cold surface, thin cirrus clouds. These cases are far less frequent than fully overcasted or clear pixels which results in highly polarised probability distribution (Fig. 15).

The snow component of the PCM output was validated against the MOD10A1 daily composite which is one of the few existing snow cover data sets matching the spatial and temporal resolution of the AVHRR data. Although, the high accuracy of this product was proven (Hall and Riggs, 2007) some misclassification especially between snow and clouds as well as overestimation of snow in forests in some cases may influence the quality indicators for the PCM method (Eqs. 5–11). Moreover it has to be taken into account while interpreting the results that NOAA16 data were collected during the year 2011 while NOAA17 and NOAA18 data originate from the year 2008. Thus, different snow cover conditions between these periods influence the computed statistics. In this analysis only the cloud free areas were considered therefore some pixels at snow cover edges (Fig. 5), which are often misclassified as clouds (Hall and Riggs, 2007), were not taken into account. The best agreement between PCM and MOD10A1 total snow cover

AMTD
6, 8445–8507, 2013

estimates is reported for the NOAA17 satellite which has the 1.6 µm channel activated. For the platforms with the 3.7 µm channel constantly operating the discrepancies are more significant. Although the POD of snow is considerably higher for the NOAA17 platform (Fig. 14) the FAR indicator is on the same level for the NOAA18 satellite. Therefore it can be concluded that the PCM classification results for the NOAA17, 18 satellites are valuable in terms of the snow composite derivation. Nevertheless, the accuracy of the PCM snow cover classification is strongly related to the illumination conditions. It was found that for the NOAA17, 18 satellites the absolute PCM-MOD10A1 differences increase significantly above the sun zenith angle of 70° (Fig. 13a, e and i). Similar value was reported by Solberg et al. (2010). This effect is not that apparent for the NOAA16 satellite possibly due to different snow conditions in 2011 as compared to 2008.

The comparison and validation of the PCM results against space-borne and ground data proved the high cloud detection skills. As far as the snow classification is concerned the overall accuracy of the algorithm is reasonable but the special precaution has to be taken for data generated without the 1.6 µm channel especially for the high sun zenith angle acquisitions. The lack of reference probability estimates for cloudy/clear-sky, cloudy/snow or snow/clear-sky conditions allows only the comparison with the SYNOP cloud amount which confirms the expected value spread with low probabilities (~ 15 %) for the clear-sky conditions and high ones (~ 90 %) for the fully overcasted sky. Finally, the PCM cloud and snow classification was proven to be valid for a broad range of environmental conditions across Europe and northern parts of Africa.

6 Conclusions

This study presents a robust algorithm for cloud and snow detection on AVHRR imagery which provides complementary probability estimates. Its unique design is based on the concept of a multidimensional information space implemented as an array with

AMTD
6, 8445–8507, 2013

precomputed probability estimates where each dimension corresponds to different spectral, angular and ancillary data combination. In order to construct this array first the continuous data such as spectral measurements have to be binned using a set of threshold values which do not have to be equally distributed. Such an approach resolves the problem of a single threshold value derivation which should feature the highest discrimination skills for the particular environmental conditions. On the contrary to the decision-tree approaches PCM algorithm considers all of the available information in a single step to extract the probability estimates between clear-sky/cloudy, clear-sky/snow and snow/cloudy conditions from the multidimensional array. Another novel feature of PCM is the employment of the Invariant Coordinate System (ICS) transformation which utilises Principle Component Analysis (PCA) to combine reflectance in AVHRR channel 0.6/0.8 µm or 1.6 µm or reflective part of 3.7 µm with the thermal difference between 10.8 µm channel and skin surface temperature provided by the climate model. This enhances the detection of thin/broken clouds as well as the separation between snow and ice clouds at the same time reducing the dimensionality of the analysis.

The PCM classification accuracy was assessed by the comparison with an extensive set of a reference data consisting of the PPS and MOD35 cloud masks, CALIPSO/CALIOP vertical feature mask, SYNOP cloud observations and MOD10A1 daily snow composite. The reported cloud detection skills were high ($POD > 0.8$ and $FAR < 0.1$) and remained at the same level or better than the reference PPS algorithm (due to initial correction of the training data set). Furthermore some artefacts related to areas with implausibly high cloud cover present in the PPS data were almost not apparent in the PCM classification. The snow detection skills derived on a basis of comparison with MOD10A1 product are moderate ($0.42 < POD < 0.62$ and $0.08 < FAR < 0.17$) with significantly better results for the NOAA17 with the 1.6 µm channel activated and for the low sun zenith angle acquisitions. The probability estimates for cloudy conditions were found to be directly proportional to the SYNOP cloud amount and feature a reasonable spread (from ~ 12 % for clear-sky conditions to ~ 90 % for overcasted sky).

AMTD
6, 8445–8507, 2013

Nevertheless, to fully validate the PCM product the inter-comparison with other probabilistic classification algorithms such as the one included in the PATMOS-x package as well as the modified version of the PPS algorithm will be performed. The source code of the PCM method implemented in R together with the test cases and annual LUTs are provided on the web page: http://pcm.r-forge.r-project.org/.

Acknowledgements. The authors gratefully acknowledges the Science and Technology S+T division of armasuisse for financial support of the project; the Satellite Application Facility on Support to Nowcasting & Very Short Range Forecasting (NWC SAF), Satellite Application Facility on Numerical Weather Prediction (NWP SAF), Pascal Brunel (Météo-France) for providing the AVHRR processing software; European Centre for Medium-Range Weather Forecasts (ECMWF) for providing the NWP and SYNOP data; Joint Research Centre of the European Commission for providing land cover data; USGS for providing DEM data and NASA for providing MODIS and CALIPSO/CALIOP data.

References

Ackerman, S., Strabala, K., Menzel, W., Frey, R., Moeller, C., and Gumley, L.: Discriminating clear sky from clouds with MODIS, J. Geophys. Res., 103, 32–141, 1998. 8447, 8449, 8450, 8452, 8455, 8463, 8467, 8473, 8481

Allen, R., Durkee, P., and Wash, C.: Snow/Cloud Discrimination with Multispectral Satellite Measurements, J. A. Meteorol., 29, 994–1004, doi:10.1175/1520-0450(1990)029<0994:SDWMSM>2.0.CO;2, 1990. 8458

Anderson, J. R.: A land use and land cover classification system for use with remote sensor data, vol. 964, United State Printing Office, Washington, 1976. 8472, 8480

Arya, S., Mount, D., Netanyahu, N., Silverman, R., and Wu, A.: An optimal algorithm for approximate nearest neighbor searching fixed dimensions, J. ACM, 45, 891–923, 1998. 8466

Bartholomé, E. and Belward, A.: GLC2000: a new approach to global land cover mapping from Earth observation data, Int. J. Remote Sens., 26, 1959–1977, 2005. 8459, 8463, 8472, 8480

Bernstein, R.: Sea surface temperature estimation using the NOAA 6 satellite advanced very high resolution radiometer, J. Geophys. Res.-Oceans, 87, 9455–9465, 1982. 8449

Bréon, F.-M. and Colzy, S.: Cloud detection from the spaceborne POLDER instrument and validation against surface synoptic observations, J. Appl. Meteorol., 38, 777–785, 1999. 8482
Derrien, M. and Le Gléau, H.: MSG/SEVIRI cloud mask and type from SAFNWC, Int. J. Remote Sens., 26, 4707–4732, 2005. 8447, 8450
Di Vittorio, A. V. and Emery, W. J.: An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land, IEEE T. Geosci. Remote, 40, 1682–1694, 2002. 8449
Dybbroe, A., Karlsson, K., and Thoss, A.: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling, Part I: Algorithm description, J. Appl. Meteorol., 44, 39–54, 2005a. 8449, 8450, 8452, 8453, 8454, 8455, 8464, 8480
Dybbroe, A., Karlsson, K., and Thoss, A.: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling, Part II: Tuning and validation, J. Appl. Meteorol., 44, 55–71, 2005b. 8464
Eyre, J., Brownscombe, J., and Allam, R.: Detection of fog at night using Advanced Very High Resolution Radiometer (AVHRR) imagery, Meteorol. Mag., 113, 266–271, 1984. 8454
Fraser, R. and Kaufman, Y.: The relative importance of aerosol scattering and absorption in remote sensing, IEEE T. Geosci. Remote, 23, 625–633, 1985. 8449
Gómez-Chova, L., Camps-Valls, G., Calpe-Maravilla, J., Guanter, L., and Moreno, J.: Cloud-screening algorithm for ENVISAT/MERIS multispectral images, IEEE T. Geosci. Remote, 45, 4105–4118, 2007. 8447, 8448, 8476, 8478, 8482
Hall, D. and Riggs, G.: Accuracy assessment of the MODIS snow products, Hydrol, Process., 21, 1534–1547, 2007. 8447, 8464, 8482
Hall, D., Riggs, G., Salomonson, V., DiGirolamo, N., and Bayr, K.: MODIS snow-cover products, Remote Sens. Environ., 83, 181–194, 2002. 8464, 8474
Heidinger, A. K., Anne, V. R., and Dean, C.: Using MODIS to estimate cloud contamination of the AVHRR data record, J. Atmos. Ocean. Tech., 19, 586–601, 2002. 8479
Heidinger, A. K., Evan, A., Foster, M., and Walther, A.: A naive Bayesian cloud-detection scheme derived from CALIPSO and applied within PATMOS-x, J. Appl. Meteorol. Clim., 51, 1129–1144, 2012. 8447, 8476, 8478, 8482
Houghton, J., Meiro Filho, L., Callander, B., Harris, N., Kattenburg, A., and Maskell, K.: Climate change 1995: The science of climate change: contribution of working group I to the second assessment report of the Intergovernmental Panel on Climate Change, vol. 19390, Cambridge University Press, 1996. 8447

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Hüsler, F., Fontana, F., Neuhaus, C., Riffler, M., Musial, J., and Wunderle, S.: AVHRR Archive and Processing Facility at the University of Bern: A comprehensive 1-km satellite data set for climate change studies, EARSeL eProceedings, 10, 83–101, 2011. 8458

Hüsler, F., Jonas, T., Wunderle, S., and Albrecht, S.: Validation of a modified snow cover retrieval algorithm from historical 1-km AVHRR data over the European Alps, Remote Sens. Environ., 121, 497–515, 2012. 8448, 8450

Inoue, T.: On the temperature and effective emissivity determination of semi-transparent cirrus clouds by bi-spectral measurements in the 10 µm window region, J. Meteorol. Soc. Jpn., 63, 88–99, 1985. 8454

Jones, M., Saunders, M., and Guymer, T.: Reducing cloud contamination in ATSR averaged sea surface temperature data, J. Atmos. Ocean. Tech., 13, 492–506, 1996. 8447

Karlsson, K.-G.: Estimation of cloudiness at high latitudes from multispectral satellite measurements, Ambio, 24, 33–40, 1995. 8476, 8481

Karlsson, K.-G. and Dybbroe, A.: Evaluation of Arctic cloud products from the EUMETSAT Climate Monitoring Satellite Application Facility based on CALIPSO-CALIOP observations, Atmos. Chem. Phys., 10, 1789–1807, doi:10.5194/acp-10-1789-2010, 2010. 8477, 8481

Kaufman, Y. J., Remer, L. A., Tanre, D., Li, R.-R., Kleidman, R., Mattoo, S., Levy, R. C., Eck, T. F., Holben, B. N., Ichoku, C., Martins, J. V., and Koren, I.: A critical examination of the residual cloud contamination and diurnal sampling effects on MODIS estimates of aerosol over ocean, IEEE T. Geosci. Remote, 43, 2886–2897, 2005. 8447

Key, J. and Barry, R.: Cloud cover analysis with Arctic AVHRR data: 1. Cloud detection, J. Geophys. Res.-Atmos., 94, 18521–18535, 1989. 8450

Khlopenkov, K. and Trishchenko, A.: SPARC: New cloud, snow, and cloud shadow detection scheme for historical 1-km AVHHR data over Canada, J. Atmos. Ocean. Tech., 24, 322–343, 2007. 8447, 8448, 8450, 8451, 8456, 8458

Lyapustin, A., Wang, Y., and Frey, R.: An automatic cloud mask algorithm based on time series of MODIS measurements, J. Geophys. Res., 113, D16207, doi:10.1029/2007JD009641, 2008. 8456

Mardia, K. V., Kent, J. T., and Bibby, J.: Multivariate AnalysisAcademic Press, New York, London, 1979. 8459

Martins, J. V., Tanré, D., Remer, L., Kaufman, Y., Mattoo, S., and Levy, R.: MODIS cloud screening for remote sensing of aerosols over oceans using spatial variability, Geophys. Res. Lett., 29, 8009, doi:10.1029/2001GL013252, 2002. 8450

Mayaux, P., Bartholomé, E., Fritz, S., and Belward, A.: A new land-cover map of Africa for the year 2000, J. Biogeogr., 31, 861–877, 2004. 8459
Merchant, C., Harris, A., Maturi, E., and MacCallum, S.: Probabilistic physically based cloud screening of satellite infrared imagery for operational sea surface temperature retrieval, Q. J. Roy. Meteorol. Soc., 131, 2735–2755, 2005. 8447
Minnis, P., Trepte, Q. Z., Szedung, S. M., Chen, Y., Doelling, D. R., Young, D. F., Spangenberg, D. A., Miller, W. F., Wielicki, B. A., Brown, R. R., Gibson, S. C., and Geier, E. B.: Cloud detection in nonpolar regions for CERES using TRMM VIRS and Terra and Aqua MODIS data, IEEE. T. Geosci. Remote, 46, 3857–3884, 2008. 8449, 8451, 8452
Nordhausen, K., Oja, H., and Tyler, D.: Tools for Exploring Multivariate Data: The Package ICS, J. Stat. Softw., 28, 1–31, 2008. 8459
Pavolonis, M. J.: Advances in extracting cloud composition information from spaceborne infrared radiances-A robust alternative to brightness temperatures, Part I: Theory, J. Appl. Meteorol. Clim., 49, 1992–2012, 2010. 8455
Pincus, R., Platnick, S., Ackerman, S., Hemler, R., and Patrick Hofmann, R.: Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators, J. Climate, 25, 4699–4720, 2012. 8447
Plummer, S.: The GLOBCARBON Cloud Detection System for the Along-Track Scanning Radiometer (ATSR) Sensor Series, IEEE T. Geosci. Remote, 46, 1718–1727, 2008. 8448
Rossow, W. and Garder, L.: Cloud detection using satellite measurements of infrared and visible radiances for ISCCP, J. Climate, 6, 2341–2369, 1993. 8456
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. 8451, 8454, 8455
Saunders, R., Matricardi, M., and Brunel, P.: A fast radiative transfer model for assimilation of satellite radiance observations-RTTOV-5, European Centre for Medium-Range Weather Forecasts, Reading, UK, 1999. 8449
Schaaf, C. B., Gaoa, F., Strahlera, A. H., Luchtb, W., Lia, X., Tsanga, T., Strugnella, N. C., Zhanga, X., Jina, Y., Mullerc, J.-P., Lewisd, P., Barnsleye, M., Hobsone, P., Disneyd, M., Robertsd, G., Dunderdalec, M., Dollc, C., d'Entremontf, R. P., Hug, B., Liangh, S., Privettei, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sens. Environ., 83, 135–148, 2002. 8449
Simpson, J., McIntire, T., Stitt, J., and Hufford, G.: Improved cloud detection in AVHRR daytime and night-time scenes over the ocean, Int. J. Remote Sens., 22, 2585–2615, 2001. 8450

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

Solberg, R., Wangensteen, B., Metsämäki, S., Nagler, T., Sandner, R., Rott, H., Wiesmann, A., Luojus, K., Kangwa, M., and Pulliainen, J.: GlobSnow Snow Extent Product Guide Product Version 1.0, European Space Angency, Finland, 2010. 8483

Stowe, L., Davis, P., and McClain, E.: Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the Advanced Very High Resolution Radiometer, J. Atmos. Ocean. Tech., 16, 656–681, 1999. 8479

Team, R. C.: R: A Language and Environment for Statistical Computing, http://www.R-project.org/, R Foundation for Statistical Computing, Vienna, Austria, 2012. 8466

Tian, B., Azimi-Sadjadi, M., Vonder Haar, T., and Reinke, D.: Temporal updating scheme for probabilistic neural network with application to satellite cloud classification, IEEE T. Neural Netw., 11, 903–920, 2000. 8448

Uddstrom, M., Gray, W., Murphy, R., Oien, N., and Murray, T.: A Bayesian cloud mask for sea surface temperature retrieval, J. Atmos. Ocean. Tech., 16, 117–132, 1999. 8447

Vemury, S., Stowe, L., and Anne, V.: AVHRR pixel level clear-sky classification using dynamic thresholds (CLAVR-3), J. Atmos. Ocean. Tech., 18, 169–186, 2001. 8447

Vermote, E., Tanré, D., Deuze, J., Herman, M., and Morcette, J.: Second simulation of the satellite signal in the solar spectrum, 6S: An overview, IEEE T. Geosci. Remote, 35, 675–686, 1997. 8449

Wang, M. and Shi, W.: Estimation of ocean contribution at the MODIS near-infrared wavelengths along the east coast of the US: Two case studies, Geophys. Res. Lett., 44, 3196–3205, doi:10.1109/TGRS.2006.876293, 2005. 8463

Wang, M. and Shi, W.: Cloud masking for ocean color data processing in the coastal regions, IEEE T. Geosci. Remote, 44, 3196–3105, 2006. 8463

Warren, D.: AVHRR channel-3 noise and methods for its removal, Int. J. Remote Sens., 10, 645–651, 1989. 8454

Yang, Y., Di Girolamo, L., and Mazzoni, D.: Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land, Remote Sens. Environ., 107, 159–171, 2007. 8449, 8456

Yhann, S. R. and Simpson, J. J.: Application of neural networks to AVHRR cloud segmentation, IEEE T. Geosci. Remote, 33, 590–604, 1995. 8449, 8450, 8451, 8456

Zhao, T., Guo, W., and Fu, C.: Calibrating and evaluating reanalysis surface temperature error by topographic correction, J. Climate, 21, 1440–1446, 2008. 8451

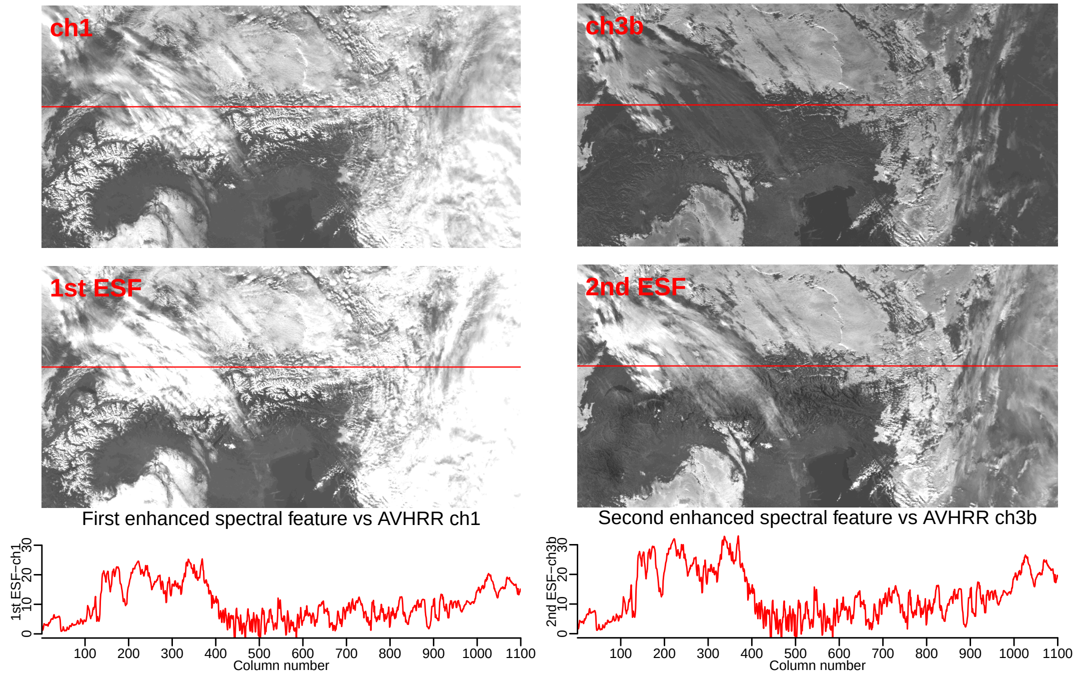

Fig. 1. Differences between Enhanced Spectral Features (ESF) generated by the Invariant Coordinate Sytem (ICS) transformation and the AVHRR channel 1 (0.6 µm) and channel 3b (3.7 µm). Bottom panels depict the differences between the ESFs and the AVHRR reflectances retrieved from the spectral profiles marked with the red lines. For details see Sects. 3.2.2 and 3.2.1.

AMTD
6, 8445–8507, 2013

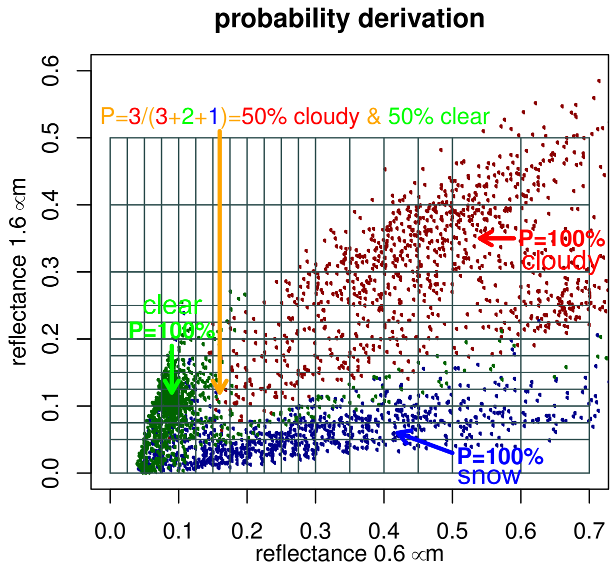

Fig. 2. Schematic graph presenting the concept of probability derivation in the PCM algorithm based on the two dimensional spectral space composed of reflectances at 0.6 and 1.6 µm. In the equation *P* denotes probability and numbers denote the counts of cloudy, clear and snow pixels within the bin. The samples were derived from the PPS and MOD10A1 classifications of the NOAA17 satellite scene acquired over the Alps on the 1 January 2008.

AMTD
6, 8445–8507, 2013

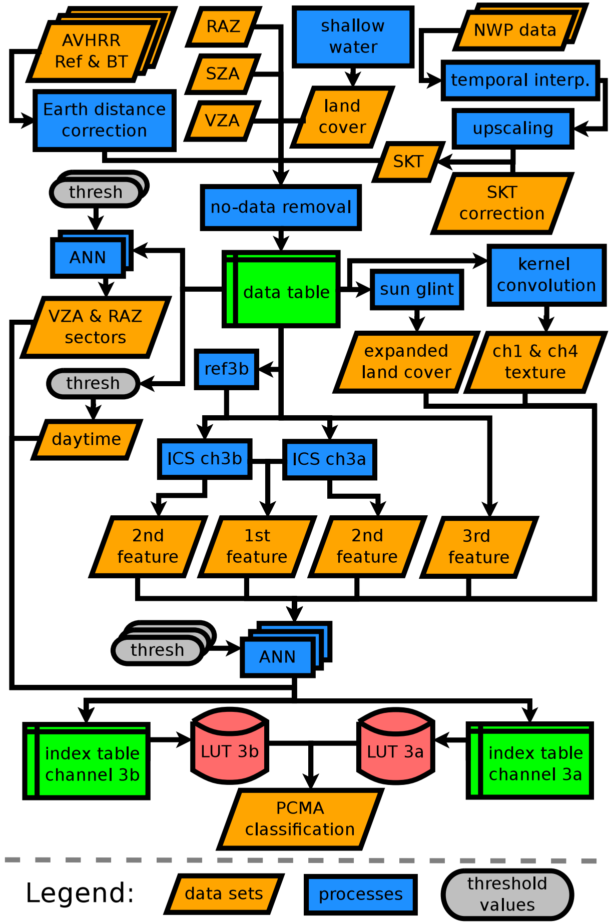

Fig. 3. Flow-chart of the PCM algorithm. For details see Sect. 3.5.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

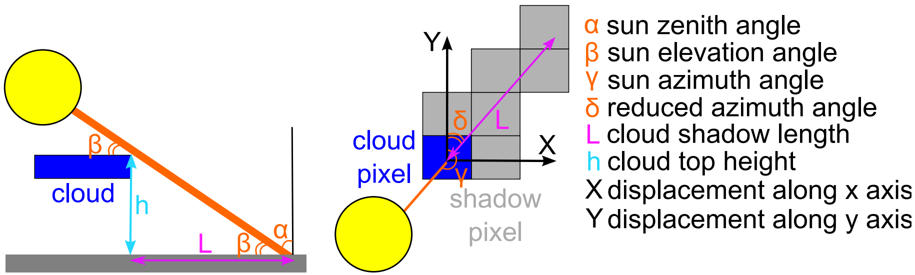

Fig. 4. Derivation of cloud shadow. For details see Sect. 3.6.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

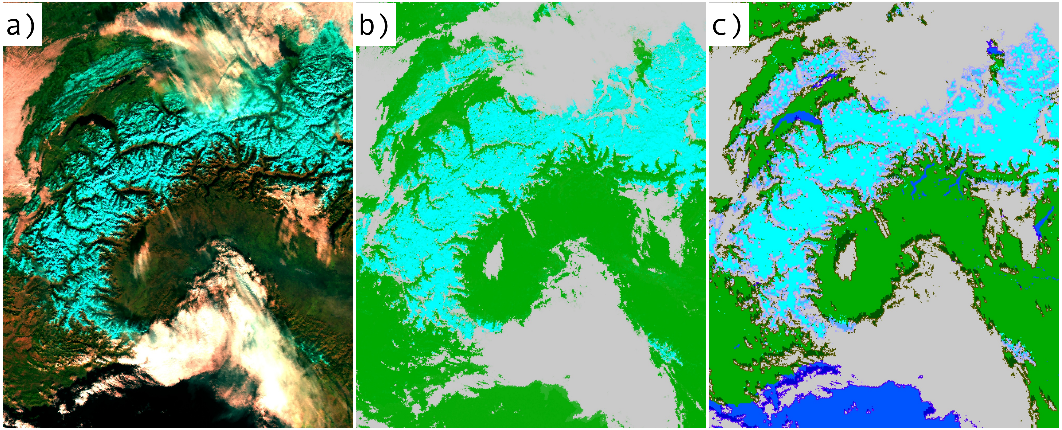

Fig. 5. PCM classification example of the NOAA17 scene acquired over the Alps on 1 January 2008 at 10:00 UTC. **(a)** false color composite ($R = 1.6\,\mu m$, $G = 0.8\,\mu m$, $B = 0.6\,\mu m$), **(b)** probabilistic cloud and snow mask, **(c)** binary cloud/cloud shadow/snow/land-water mask with classes described in Sect. 3.6. In **(b)** and **(c)** grey colour depicts clouds, green depicts cloud- and snow-free areas, blue depicts water and light blue depicts snow.

AMTD
6, 8445–8507, 2013

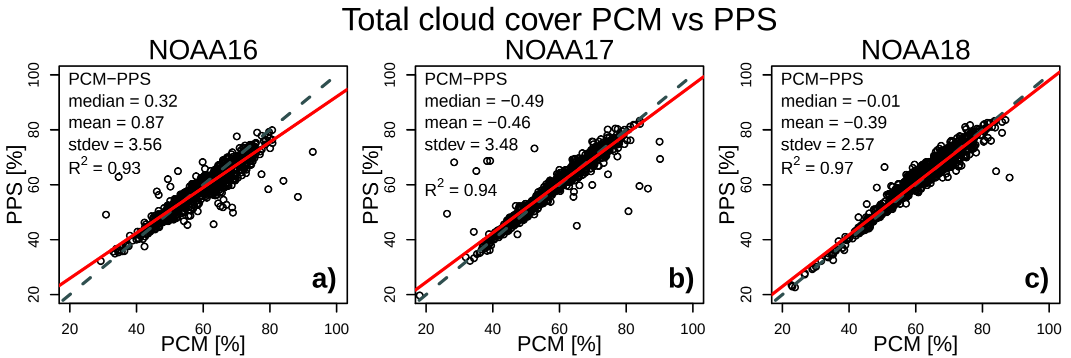

Fig. 6. Scatter plots of the total cloud cover estimates computed over each of NOAA16, 17, 18 satellites scenes separately for the PCM and PPS algorithms. Red line denotes linear trend between these two data sets. Some statistics of the PCM-PPS total cloud cover differences distribution are reported. For details see Sect. 4.1.

AMTD
6, 8445–8507, 2013

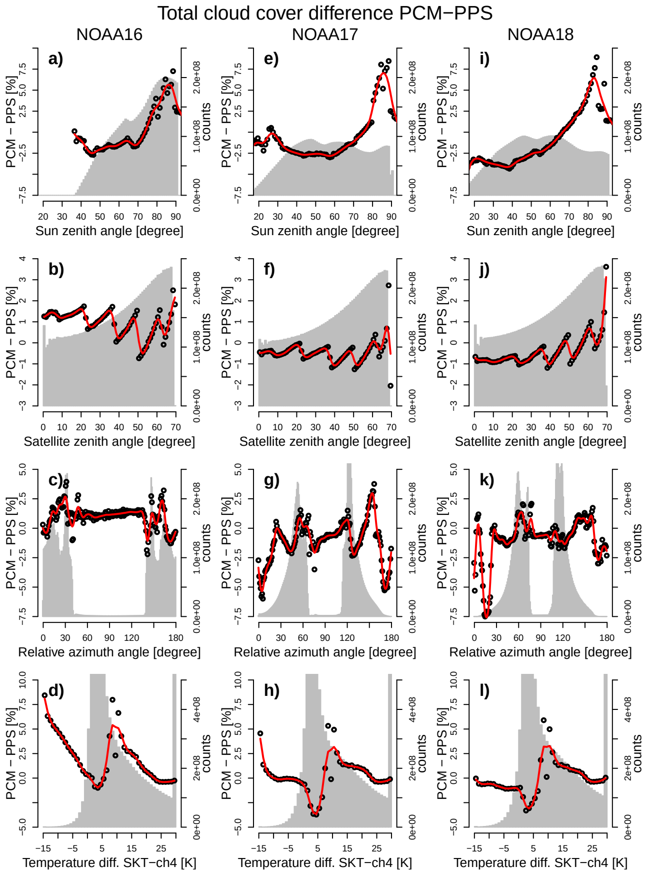

Fig. 7. Total cloud cover differences PCM–PPS as a function of selected variables derived from the annual pixel counts for the years 2011 (NOAA16) and 2008 (NOAA17, 18). The data quantity is presented as grey-shaded histograms. Red line denotes trend computed by the smoothing spline method. For details see Sect. 4.1.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

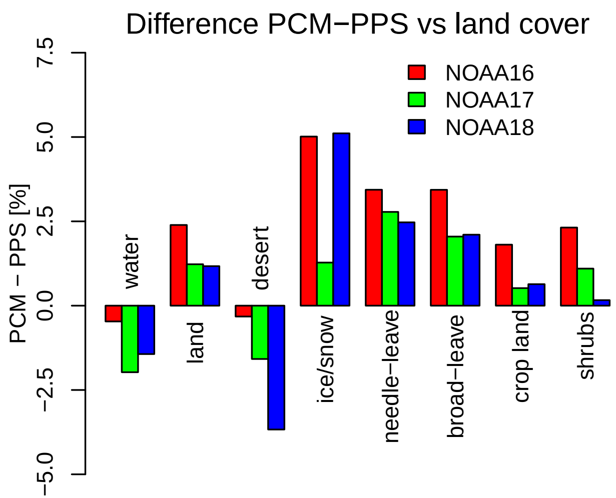

Fig. 8. Total cloud cover differences PCM–PPS as a function of land cover derived from the annual pixel counts for the years 2011 (NOAA16) and 2008 (NOAA17, 18). For details see Sect. 4.1.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

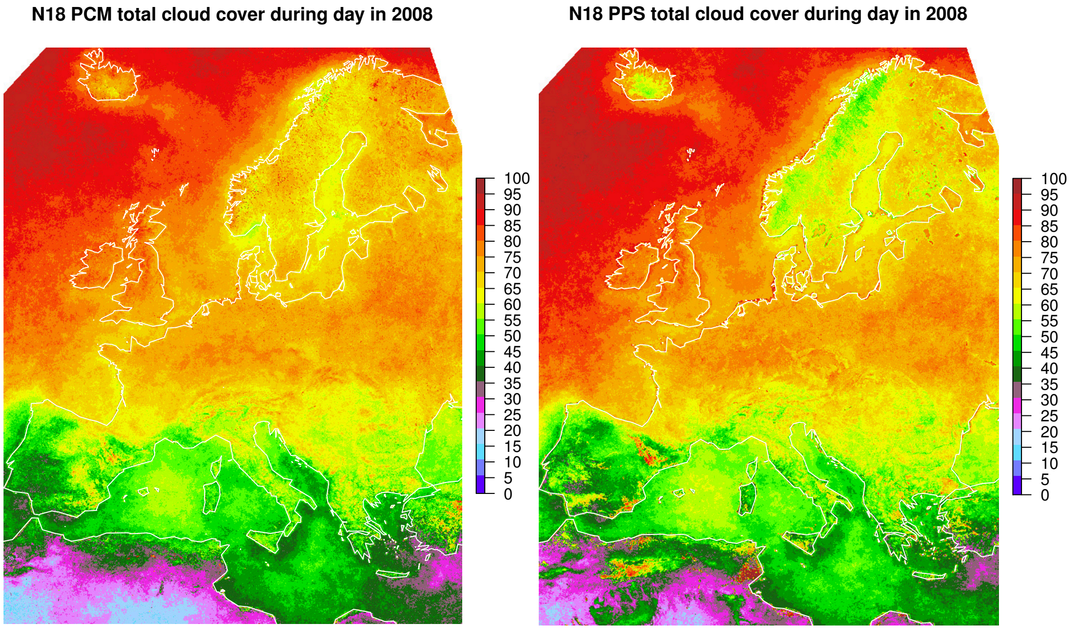

Fig. 9. PCM and PPS annual total cloud cover composites in 2008 derived from the collection of day-time NOAA18 scenes.

AMTD
6, 8445–8507, 2013

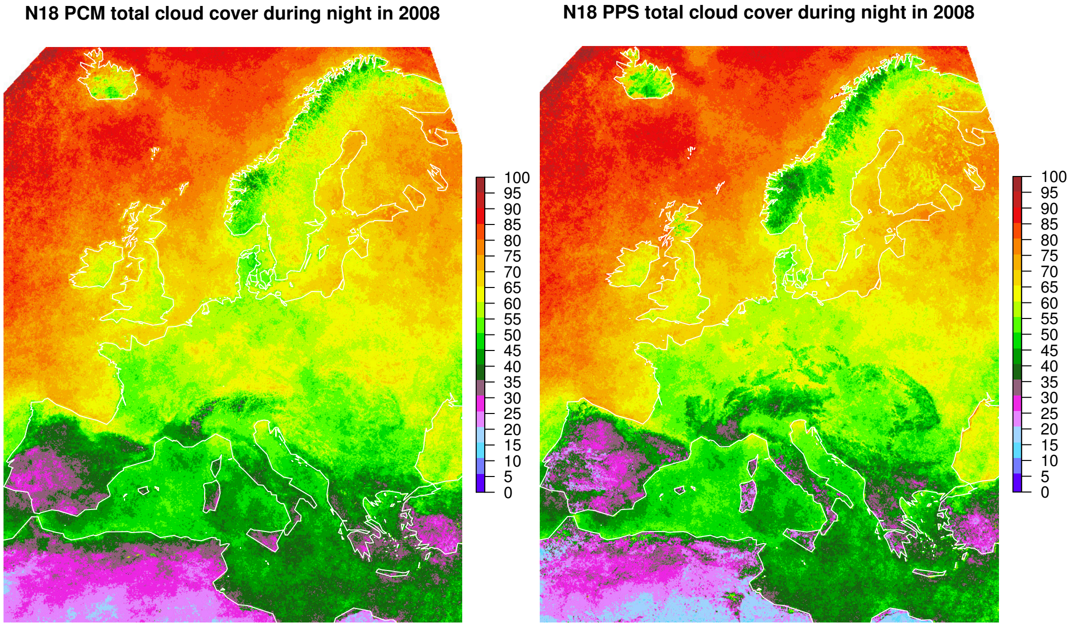

Fig. 10. PCM and PPS annual total cloud cover composites in 2008 derived from the collection of night-time NOAA18 scenes.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

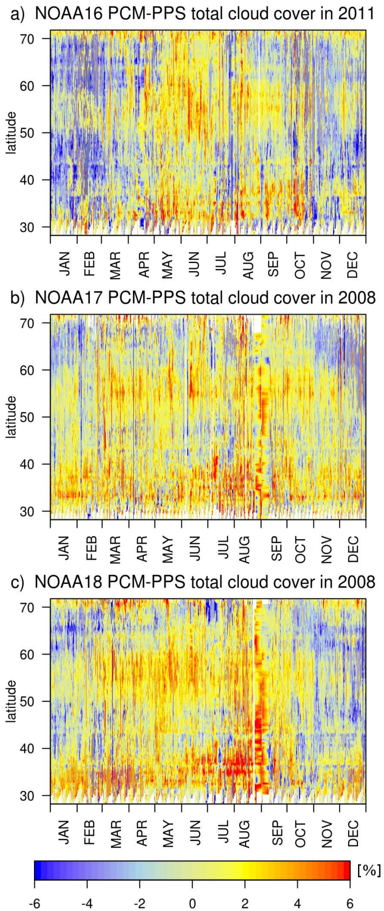

Fig. 11. Distribution of latitudinal total cloud cover differences between the PCM and PPS algorithms throughout the year 2011/2008 for the selected satellites. For details see Sect. 4.1.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

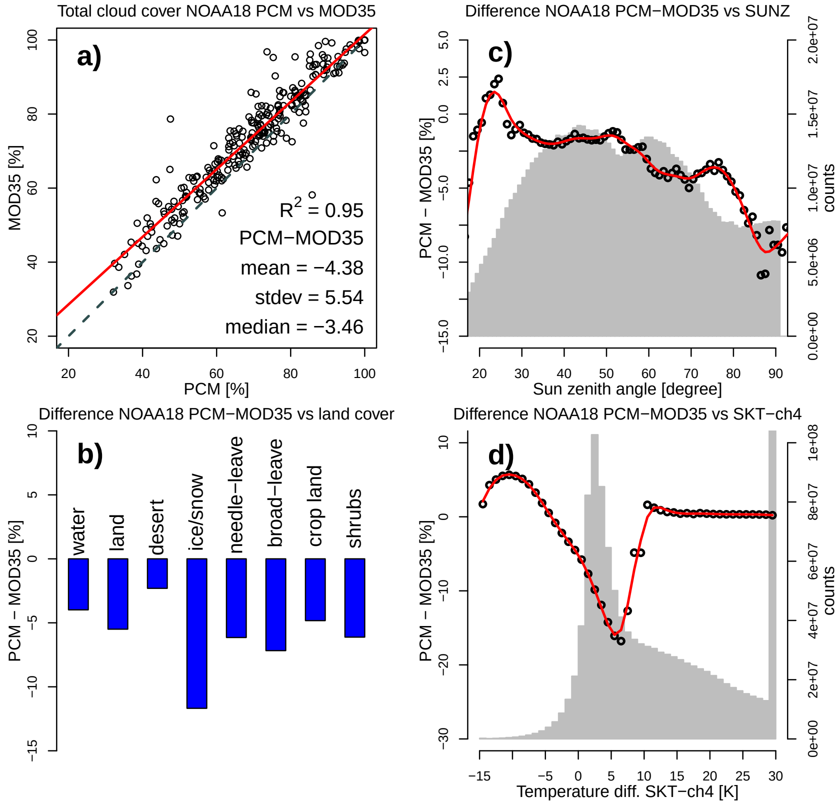

Fig. 12. Differences in total cloud cover between PCM and MOD/MYD35 products derived from annual pixel counts of the NOAA18 satellite for the year 2008: **(a)** computed over each scene, **(b)** as a function of land cover, **(c)** as a function of sun zenith angle, **(d)** as a function of thermal differences between SKT and 10.8 µm channel. Red line denotes trend. For details see Sect. 4.2.

AMTD
6, 8445–8507, 2013

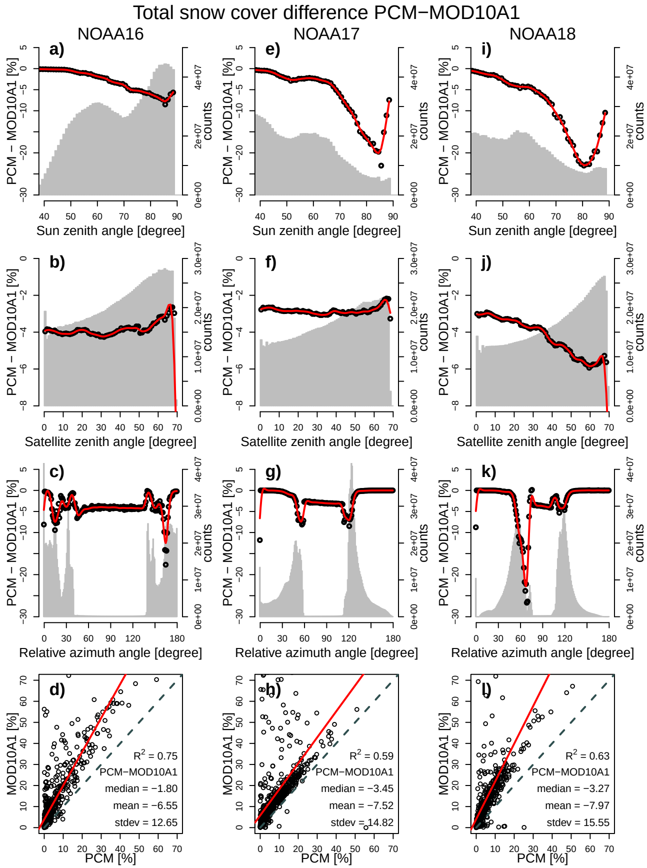

Fig. 13. Total snow cover difference PCM-MOD10A1 as a function of selected variables derived from annual pixel counts for the years 2011 (NOAA16) and 2008 (NOAA17, 18). The data quantity is presented as grey-shaded histograms. Bottom panels present the total snow cover scatter plots derived from estimates computed over each scene. Red line denotes trend. For details see Sect. 4.2.

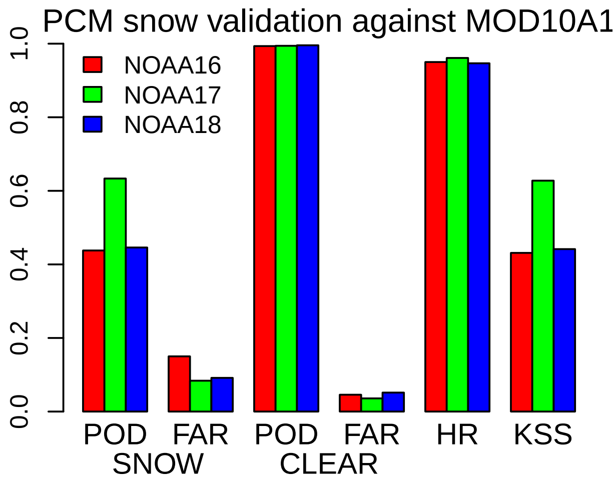

Fig. 14. Accuracy indicators of the PCM snow detection skills derived from the comparison with the MOD10A1 product. For details see Sect. 4.2.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

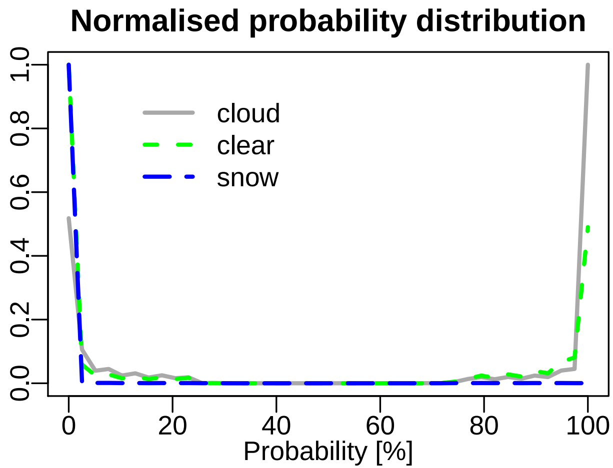

Fig. 15. PCM normalised probability distribution for clear, cloudy and snow conditions derived from annual counts of the NOAA18 satellite data. For details see Sect. 4.3.

AMTD
6, 8445–8507, 2013

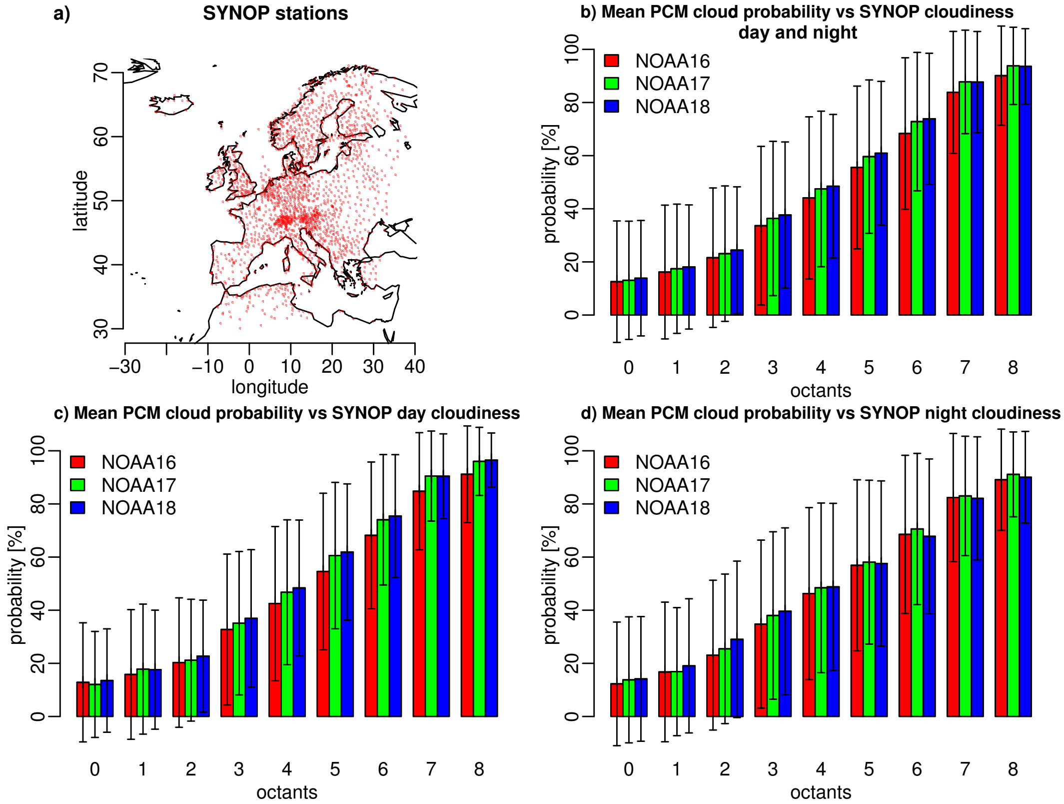

Fig. 16. (a) Geographic distribution of SYNOP stations. **(b)–(d)** Distribution of mean PCM cloud probability as a function of SYNOP cloud amount expressed in octants for different daytimes. Error bars denote standard deviations. For details see Sect. 4.3.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.

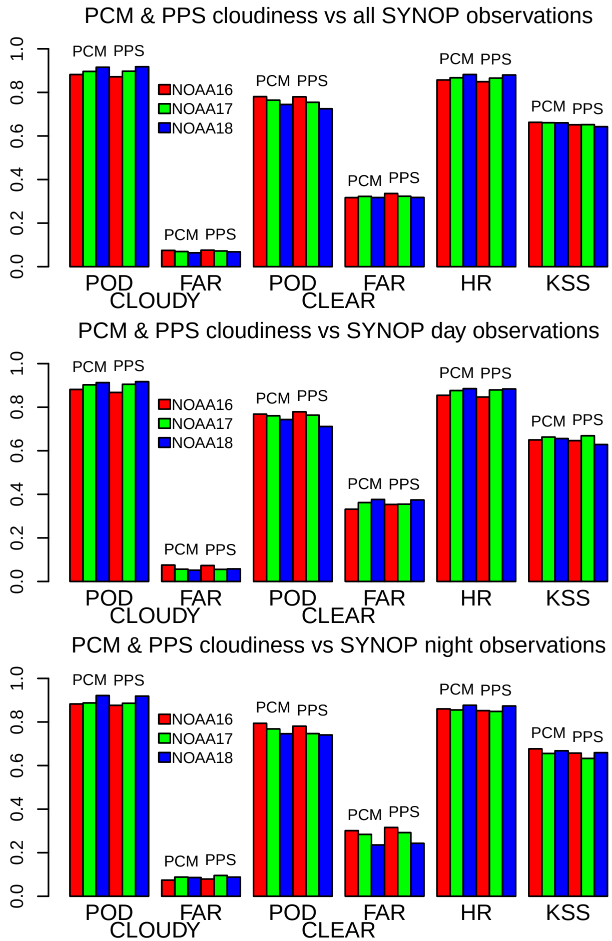

Fig. 17. Accuracy indicators of the PCM and PPS cloud detection skills derived from the comparison with SYNOP cloud observations. For details see Sect. 4.3.

AMTD
6, 8445–8507, 2013

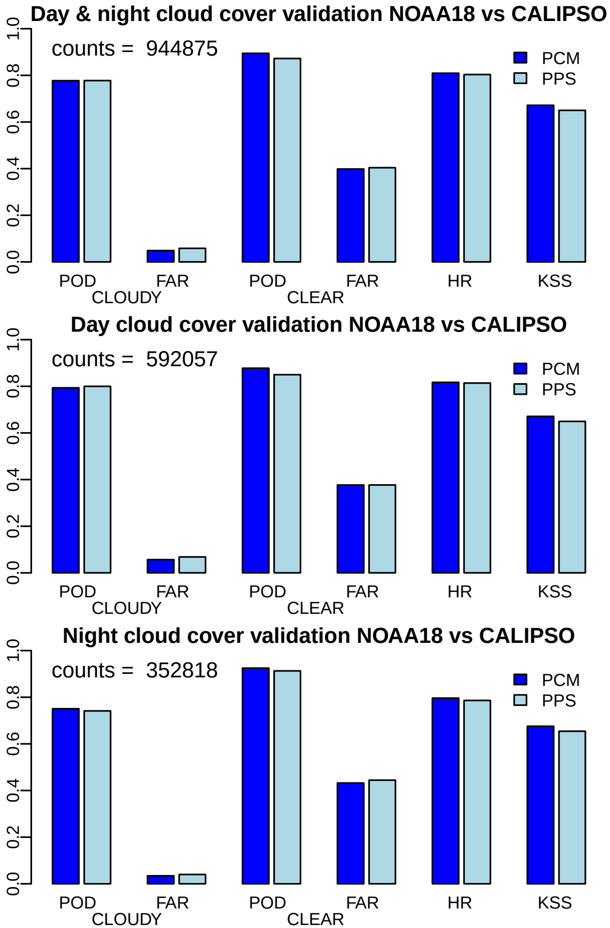

Fig. 18. Accuracy indicators of the PCM and PPS cloud detection skills derived from the comparison with CALIOP vertical feature mask. For details see Sect. 4.4.

AMTD
6, 8445–8507, 2013

Probabilistic approach to cloud and snow detection on AVHRR imagery

J. P. Musial et al.