INT. J. REMOTE SENSING, 2003, VOL. 24, NO. 3, 437–456

Remote sensing of cloud cover in the Arctic region from AVHRR data during the ARTIST experiment

C. ANANASSO†, R SANTOLERI†, S. MARULLO‡ and F. D'ORTENZIO†

†Istituto di Fisica dell'Atmosfera, Consiglio Nazionale delle Ricerche, Roma, Italy
‡ENEA, Centro Ricerche Casaccia, Roma, Italy

(*Received 10 October 2000; in final form 3 September 2001*)

Abstract. In this paper we present a cloud detection algorithm developed for the Arctic region using Advanced Very High Resolution Radiometer (AVHRR) data. Our approach is a simplified version of the Ebert method to discriminate between clouds, ice and open water in the Arctic Sea. The algorithm is tuned to work on an AVHRR scene typical of the winter to spring transition period. The algorithm has been applied to 1 month (154 scenes) of NOAA-14 AVHRR images (from 16 March to 15 April 1998) covering the region of the Arctic Sea near the Svalbard Islands. The cloud detection results are analysed using various check procedures. The algorithm's pixel classification performance was verified by a satellite image expert. The misclassified pixels were digitalized on the image and counted by the expert in order to quantify the algorithm's accuracy. The cloud classification results are quite accurate: 70% of the images (109) have an error less than 5% and only 11% of the image results have an error greater than 10%. The method's performance is also tested against independent cloud and ice observations obtained, respectively, from the Ny-Ålesund meteorological base and from the Special Sensor Microwave/Imager (SSM/I) dataset. The comparison with these independent sources of data confirms the algorithm's good performance.

1. Introduction

In recent years the increasing importance of problems related to climatic changes moved research towards a deeper study of the possible causes and probable effects. Some regions of our planet seem to be more sensitive than others to an increase in surface temperature, or to a decrease in precipitation, or to a variation in the atmosphere's composition. Two of these sensitive areas are the polar regions (Tsay *et al.* 1989). The Arctic and Antarctic areas have very particular climatic regimes: a variation in global surface temperatures is expected to be amplified in these regions. The most important climatic processes are greatly influenced by cloud cover and by the interactions between clouds and radiative fluxes. Particularly, polar cloud cover variations affect sea-ice conditions and consequently attenuate the short-wave radiation reaching the surface, increase downwelling long-wave radiation and modify the albedo feedback (Welch *et al.* 1992).

In the last decades, the lack of data due to the difficulty to obtain direct observations in these remote regions is compensated by the use of meteorological

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160110114510

satellites. They allow a long term and quite continuous sensing of the areas of interest and provide an enormous quantity of data. But to interpret polar satellite images is often very difficult due to the similarity of cloud and ice or snow surface spectral radiances (Welch *et al.* 1992). Cloud detection algorithms from satellite sensor data have been developed using visible, near-infrared and thermal infrared measurements and they have been based on threshold methods, radiative transfer models and statistical classification schemes (Key and Barry 1989). Most of the cloud detection methods were developed and successfully used for low and middle latitude data, but they are not applicable in the polar regions. A historical review of cloud detection algorithms is given by Rossow *et al.* (1989). The failure of the application of these methods to polar data is caused by a number of reasons including: snow and ice-covered surfaces having the same radiative properties and the same temperatures as clouds; the darkness during the polar night makes data collected in the visible channels unusable; the thermal structure of the troposphere is characterized by frequent isothermal and inversion layers; satellite radiometers operate near the limits of their performance range due to extremely low surface temperatures and solar illuminations; rapid small-scale variations in cloud cover can cause a possible change in snow and ice concentration (Key and Barry 1989).

Specific cloud detection algorithms have been developed for the polar regions by modifying the middle latitude schemes or by developing '*ad hoc*' algorithms. Key and Barry (1989) developed a method specifically for Arctic Advanced Very High Resolution Radiometer (AVHRR) data, based on ideas of the International Satellite Cloud Climatology Project (ISCCP). They also used the Scanning Multichannel Microwave Radiometer (SMMR) in order to detect the ice edge and produce sea-ice masks. For each pixel, nine spectral features are analysed and four surface (snow-free land, snow-covered land, open water and sea ice) and three cloud classes are defined. The cloud classes are discriminated by the brightness temperature of channel 4, which is assumed as representative of the cloud top temperature. In a successive paper, Key (1990) used statistical parameters that consider the structural properties of clouds and improved the preceding algorithm by identifying several types of clouds.

Lubin and Morrow (1998) tested the method developed by Ebert (1987) using summer AVHRR data and ship meteorological observations. The cloud classification method was applied to the 32×32 pixel cell centred about the ship's position during each satellite overpass. They defined 18 categories of surfaces and cloud types and they used eight spectral or textural features from the Ebert (1987) cloud classification method. The cloud classifications were compared with surface weather observations to determine whether the automated satellite cloud classifier could actually produce realistic descriptions of the scene. They concluded that a carefully applied pattern recognition technique could be used to produce a good cloud climatology for the Arctic Sea.

In this paper we describe an AVHRR cloud detection algorithm developed for the Arctic region in the framework of the European Community ARTIST (Arctic Radiation and Turbulence Interaction Study) Project. The main goal of ARTIST is the study of the effects of clouds and Arctic haze on surface and columnar radiative fluxes. The experimental phase of the project was carried out over the Svalbard Islands from 16 March to 15 April 1998. In this framework the development of a method to estimate the radiation budget at the air–sea interface using AVHRR data and its validation against the experimental field measurements is one of the objectives of the project. The first step of the satellite radiation budget implementation is a

correct identification of cloud cover in the AVHRR data for the ARTIST field experiment period.

A first attempt to apply Ebert's algorithm to ARTIST data was a failure due to the different period of the year, i.e. we were in winter–spring instead of summer, and to inherent inability of AVHRR to discriminate so many categories of surface and cloud types in this period of the year.

We therefore present a new cloud detection algorithm developed for the ARTIST experiment (§3). Our approach is a simplified version of the Ebert (1987) method to discriminate between clouds, ice and open water in the Arctic Sea during the winter to spring transition period. The algorithm has been applied to all the NOAA-14 AVHRR data available between 16 March 1998 and 15 April 1998, which cover the region of the Arctic Sea near the Svalbard Islands. In §4 the cloud detection results are analysed using various check procedures. In §4.1 we use an expert's visual control to test the algorithm. Moreover, the performance of the method is tested against independent cloud and ice observations obtained, respectively, from the Ny-Ålesund meteorological base (§4.2) and from the Special Sensor Microwave/Imager (SSM/I) dataset (§4.3). Conclusions are given in §5.

2. The data

2.1. *AVHRR data*

The data used in this work consist in 154 satellite passes of NOAA-14 AVHRR passages covering of the Arctic during ARTIST. The data were acquired from the Tromsö station and processed using DSP (Display Software Package RSMAS, University of Miami) at the Istituto di Fisica dell'Atmosfera of Rome.

The AVHRR measures radiation from the Earth–atmosphere system in five channels of the electromagnetic spectrum (visible, near-infrared and thermal infrared range). The satellite orbits the Earth about 14 times each day from 833 km above its surface and in the polar regions it allows a frequent sensing of the areas (up to five or six satellite overpasses each day in the same area) (Lubin and Morrow 1998).

Using DSP the raw data (at full spatial resolution, 1.1 km at nadir) are subjected to a sequence of processing steps: ingestion, navigation, geophysical conversion and remapping (Kilpatrick *et al.* 2001). Ingestion is the process that reads the raw data and converted them in a standard image format; the image is then adjusted from line/pixel space to latitude/longitude space (navigation). This step is partially automatic (from the orbital elements of the spacecraft, time and attitude), and partially manual, adjusting the coastlines on the image. Next, the geophysical conversion converts AVHRR counts to calibrated radiance values: planetary albedo for channel 1 and channel 2, and brightness temperature for channel 3 and channel 4. Finally, the data are mapped in geographical projection with a resolution of 1 km over an the area centred on the Svalbard Islands. We selected the area between 68° N and 83° N and 0° and 30° E. An equirectangular projection was used in order to have constant resolution over the entire region.

The measured radiances from channel 1 and 2 are due to solar radiance reflected by the Earth surface and atmosphere and were obtained by calibrating satellite counts:

$$R_i = A_i C_i + B_i \tag{1}$$

where i can be 1 or 2, R is a radiometrically calibrated pseudo-albedo in percent (Lubin and Morrow 1998) that can be considered as the planetary albedo (hereafter,

albedo) in the electromagnetic range of the sensor channel; C_i are the satellite counts. Slope and bias (A_i and B_i) can vary with time. There is, in fact, evidence that the visible (channel 1) and the near-infrared (channel 2) channels, which have no on-board calibration devices, deteriorate in orbit. It was therefore necessary to develop some correction algorithms to evaluate the post-launch degradation of these two channels and to correct the measured signals of the upwelling radiance at the top of the atmosphere (Rao and Chen 1995). To ensure the quality of AVHRR data, NASA currently provides updated coefficients to take into account the degradation rate of the radiometer and the Earth–Sun distance for the date of measurements. The coefficients for the period of our data are $A_1 = 0.136\,144$, $B_1 = -5.581\,904$; $A_2 = 0.172\,641$, $B_2 = -7.078\,281$ (http://edcwww.cr.usgs.gov/glis/hyper/guide/avhrr).

The brightness temperatures of the near and thermal infrared channels were computed using in-flight calibration coefficients. The channel 3 radiance is the sum of contributions from the solar reflected radiance and surface emission (see §3.1). Channel 4 radiance is the terrestrial emission in the thermal infrared region (Robinson 1985). An example of the channel 1 albedo maps is reported in figure 1.

2.2. *SSM/I data*

SSM/I (Special Sensor Microwave/Imager) is a passive microwave radiometer on board of the DMSP (Defense Meteorological Satellite Program) satellite series. The SSM/I orbit is near-circular, sun-synchronous and near-polar, with an altitude of 860 km and the orbital period is 102 min. SSM/I measures brightness temperature in the microwave range of the Earth–atmosphere system in four spectral bands between 19 and 86 GHz.

The sea-ice detection utilizes low frequency measurements: sea-ice has a higher emissivity than water and hence the two can be separated. In this study we used directly the sea-ice concentration maps elaborated and provided by NSIDC (National Snow and Ice Data Center) at $0.25 \times 0.25^\circ$ spatial resolution for the period and the area of interest.

The algorithm used at NSIDC was developed and tested by a NASA (National Aeronautics and Space Administration) team. The three SSM/I channels used in calculating sea-ice concentration are the 19.4 GHz horizontally (H) and vertically (V) polarized channels and the vertically polarized 37.0 GHz channel. (http://nsidc.org/NASA/GUIDE/SSMI/nasa.html). The SSM/I radiances from each of the three channels are first mapped onto polar stereographic grid (the so-called SSMI grid). The gridded radiances are then used to calculate grids for the two independent variables used in the algorithm to evaluate the first-year ice concentration and the multiyear ice concentration. Summing the last ones they evaluate the total ice concentration. The sensitivity of the calculated ice concentrations to ice temperature variations is describe in Cavalieri *et al.* (1984) and Swift and Cavalieri (1985). Errors in the derived sea-ice concentrations are due to the inability of the algorithm to discriminate among more than two radiometrically different sea-ice types, seasonal and non-seasonal variations in sea-ice emissivity and random or systematic instrument errors. (http://nsidc.org/NASA/GUIDE/SSMI/nasa.html).

2.3. *Meteorological data*

Since 1992 a meteorological observatory program is in progress at Koldewey Station (78°9′N, 11°9′E) near Ny-Ålesund. All activities are carried out in cooperation with the Norsk Polar Institutt, who provides the routine synoptic

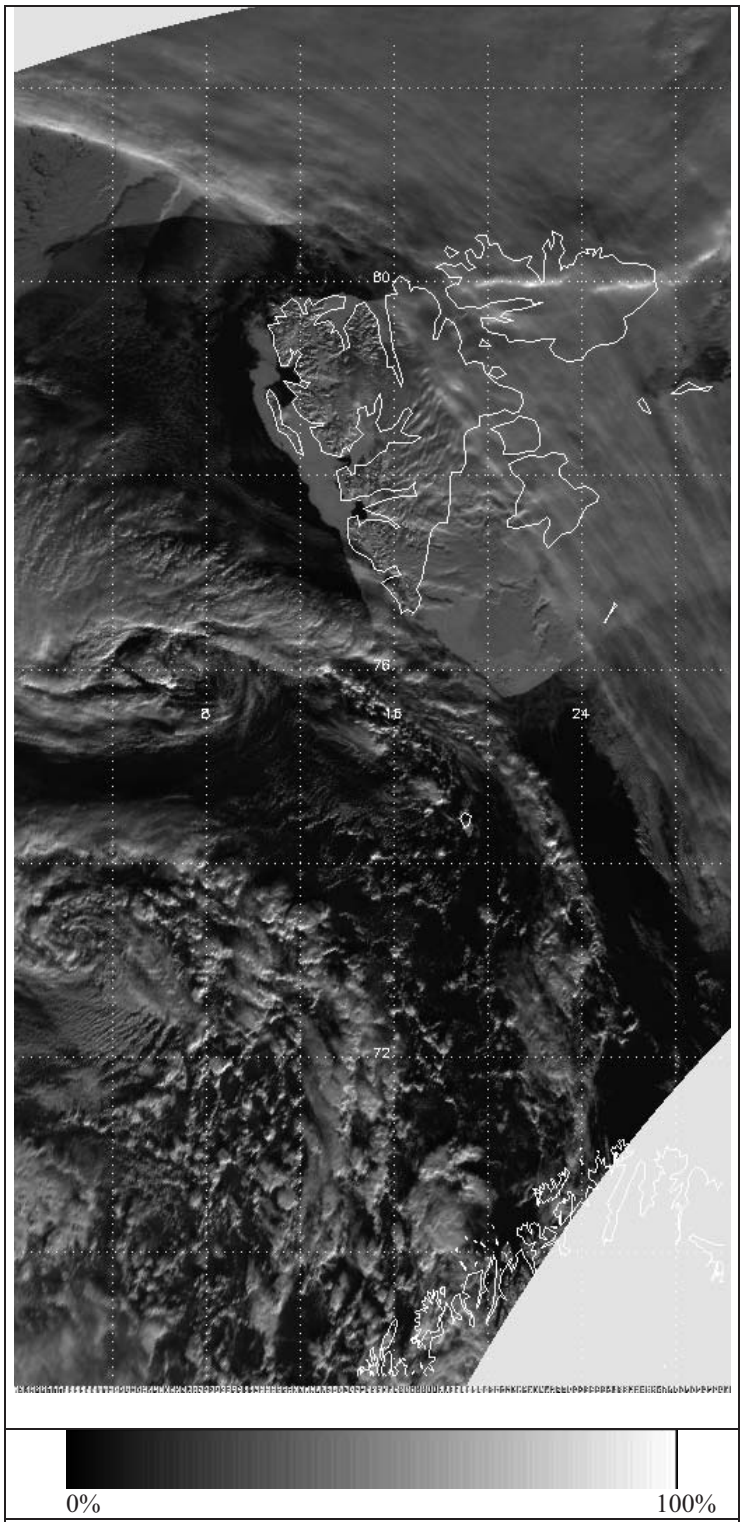

Figure 1. Example of AVHRR final data processing products: channel 1 albedo map (in percent) of the Svalbard Islands (lat. 68–83 N, long. 0–30 E).

observations. These observations include measurements of air temperature, air pressure, wind vector, dew point temperature, clouds, horizontal visibility and some additional observations. The full observation program is carried out at 6, 12, 18 UTC. We have downloaded the total sky cloud cover for the entire period of interest directly from their web site (www.awi-bremerhaven.de/MET/NyAlesund).

3. Cloud detection algorithm

The proposed algorithm is a simplified version of Ebert's (1987) method and was developed to work on AVHRR scenes, typical of the transition seasons in the Arctic region. The algorithm classifies each pixel of the AVHRR scene in three classes: ice over sea or over land (ICE in the following), open water (OPEN WATER) and clouds (CLOUD). In this study, the class for land not covered by ice/snow is not considered because this situation never occurs in any of the available images.

As in every threshold method, the algorithm is based on the definition of the limit values for some selected parameters for each defined class. Ten representative satellite passes are selected to define these thresholds. These images are characterized either by the presence of all three classes or are completely covered by clouds. The threshold values have been defined after several trials in analysing the effect of the chosen limits on the pixel classification of the 10 selected images.

The developed method consists of three steps and it is illustrated by the flowchart in figure 2. The first step is devoted to the classification of open water and to the discrimination of the pixels that can be clearly identified as cloud or ice. At this step a further class of unclassified pixels is defined. In the second and third step the unclassified pixels are subjected to further analysis that allow the final separation between ice and cloudy pixels.

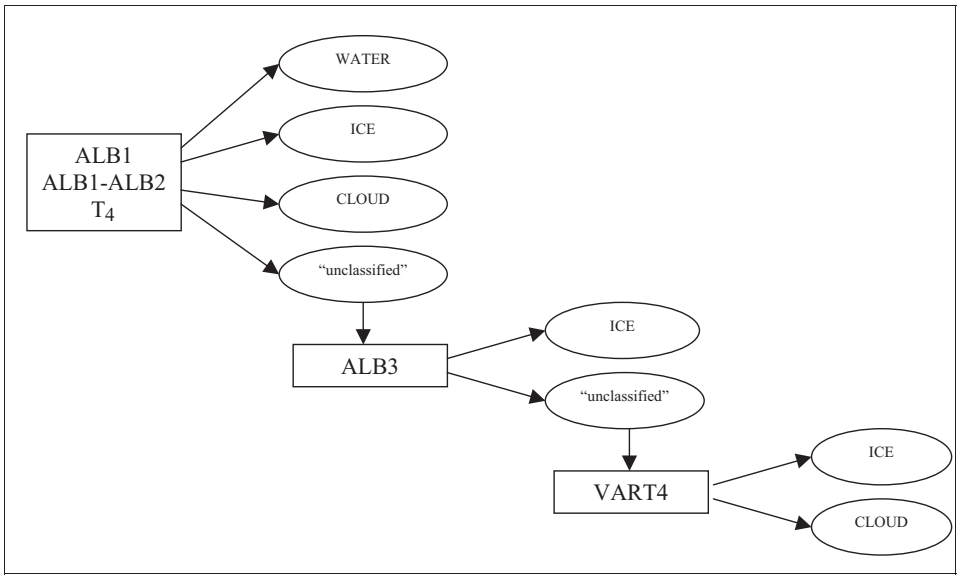

Figure 2. Flow chart of the cloud detection algorithm.

In the first step, for each pixel the following parameters are examined: normalized channel 1 albedo (ALB1), difference between channel 1 and channel 2 normalized albedo (ALB1–ALB2) and channel 4 brightness temperature (T_4). All the albedos are normalized by the cosine of the solar zenith angle in order to make the albedos independent from latitude and from the time of the satellite overpass.

Channel 1 albedo is the planetary albedo in the visible spectral range and it is due to the solar radiation reflected from the Earth's surface or from the cloud top layer; it has high values for clouds, ice and snow cover and low values for open water, land and vegetation. The difference between channel 1 and channel 2 albedo (or alternatively the ratio between channel 1 and channel 2 albedo) can be used to distinguish between ice and clouds. In case of snow or ice covered surfaces, the albedo decreases with increasing wavelength, while the particle effective mean radius responsible for cloud reflectance makes the cloud albedo almost constant in the visible spectral range (Key and Barry 1989). Consequently, ALB1–ALB2 tends to have great values for ice or snow, and low values for clouds. The channel 4 brightness temperature is related to the top of the atmosphere temperature and it has very different values for open water and clouds or ice. Typically, open water pixels are warmer than cloud or ice pixels.

In this first step the algorithm separates OPEN WATER from the other classes (CLOUD, ICE) and defines an additional class of 'unclassified' pixels (figure 2). Due to the overlapping of the spectral characteristics of ice and cloud surfaces, most of the pixels fall in the unclassified class. The threshold values chosen for the different classes are reported in table 1.

In the second step we used channel 3 albedo (ALB3) for a first discrimination of the unclassified pixels. In fact, several studies (Kidder and Wu 1984, Key and Barry 1989, Allen *et al.* 1990) have tended to demonstrate that channel 3 radiance could be a very useful parameter to separate low clouds from ice cover. Kidder and Wu (1984) showed that there is a great contrast in the channel 3 radiance between low clouds and snow/ice cover and, because they have a similar emission temperature, the contrast depends on the solar reflection contribution.

The channel 3 brightness temperature was converted to albedo by subtracting the thermal contribution estimated by the use of the channel 4 brightness temperature, with the following equation (Key and Barry 1989, Allen *et al.* 1990):

$$\mathrm{ALB3} = \frac{L_3 - B_3(T_4)}{F^0_{\mathrm{SW}} \cos Z - B_3(T_4)} \tag{2}$$

where L_3 is the measured radiance in channel 3, F^0_{SW} is the incident solar radiance in the electromagnetic range of channel 3 weighted by the cosine of the solar zenith

Table 1. Threshold values for the parameter used by the cloud detection algorithm in the first step pixel classification.

OPEN WATER	$\mathrm{ALB1} \leqslant 19.5$	$-7.2 \leqslant \mathrm{ALB1} - \mathrm{ALB2} \leqslant 5.9$	$267 \leqslant T_4 \leqslant 283$
CLOUD	$8.3 \leqslant \mathrm{ALB1} \leqslant 24$ $\mathrm{ALB1} \geqslant 91.7$	$\mathrm{ALB1} - \mathrm{ALB2} \leqslant -16.6$	$224 \leqslant T_4 \leqslant 273$
ICE	$24 < \mathrm{ALB1} < 91.7$	$10.83 < \mathrm{ALB1} - \mathrm{ALB2} \leqslant 15.56$	$247 \leqslant T_4 \leqslant 273$
ICE CLOUD (unclassified)	$24 < \mathrm{ALB1} < 91.7$	$-16.6 < \mathrm{ALB1} - \mathrm{ALB2} < 10.8$	$247 < T_4 < 273$

angle and $B_3(T_4)$ is the contribution to the measured radiance from thermal emission of the viewed surface. $B_3(T_4)$ is estimated by Planck equation using as thermal emission temperature the T_4 deduced by AVHRR channel 4 measurements assuming the transmittance of the intervening atmosphere to be equal to 1.0. In this second step pixels with ALB3 $\leqslant$ 1.4 are classified as ICE; the remaining pixels are subjected to further analysis.

The last step (figure 2) consists in the analysis of one texture characteristic of the image to separate the last unclassified pixels in ICE and CLOUD. Usually clouds have a very rough shape, and temperature could vary a lot within a small area determining an increase of the variance, while ice has a smoother temperature distribution. The brightness temperature variance of channel 4 (VART4) is therefore evaluated in a box of $N \times N$ pixels, around each undecided pixel, to try to classify it.

The dimension of the cell is an important issue: if it is too large it can blur the boundaries between classes, if it is too small it may not permit an adequate description of the textural features to distinguish between classes (Key 1990). Following the statistical consideration made by previous studies (Ebert 1987, Lubin and Morrow 1998) we selected a box made of 32×32 pixels (about $32\,km^2$). For values of VART4 $\leqslant 0.2^{\circ}C$ the pixel is classified as ICE, in the other case as CLOUD.

4. Evaluation of the algorithm performances

4.1. *Visual control*

The cloud detection algorithm was applied to all of the 154 daytime images and an image mask of CLOUD, ICE and OPEN WATER was produced for each satellite pass.

The produced masks are then subjected to a visual control made by an expert in satellite images. The control was made by displaying the channel 1 surface albedo together with the image mask. The areas of misclassified pixels were digitalized by the expert on the image and the pixels were counted.

In general the analysis of the ALB1 image allows the expert to discriminate quite easily between the image areas covered clouds or ice and the open water regions. Only in the presence of high cloud amount over ice, the similarity between the reflectance characteristics of the two surfaces makes visual control very difficult. Visual control can be considered as the 'truth' because the human observer does not analyse the single pixel, but he/she makes a scene classification taking into account several global characteristics: his/her own cloud pattern knowledge, the total cloud or ice distribution in the image and the knowledge of the surface conditions in the area resulting from the analysis of the previous and following images.

Figures 3–4 show some examples of the masks produced by the cloud detection algorithm together with the channel 1 albedo image. For each image we used the expert's definition of misclassified pixels to define the percentage error as the ratio between the number of erroneous pixels and the total number of pixels.

Figure 3 is an example of classified image corresponding to the satellite pass of 16 March 1998 at 10:02 UTC. In the albedo image, the Svalbard Islands covered by ice can be easily recognized and the algorithm correctly classifies them as ICE.

The algorithm correctly identified as OPEN WATER the large portion of the ocean south and west of the islands clearly visible in the albedo image. In the west and south-west part of figure 3(*a*) two spiral shaped atmospheric systems of low non-homogeneous clouds are present over the ocean as well as a thin and compact cloud system cutting the image from north west to south-east. In the corresponding flag

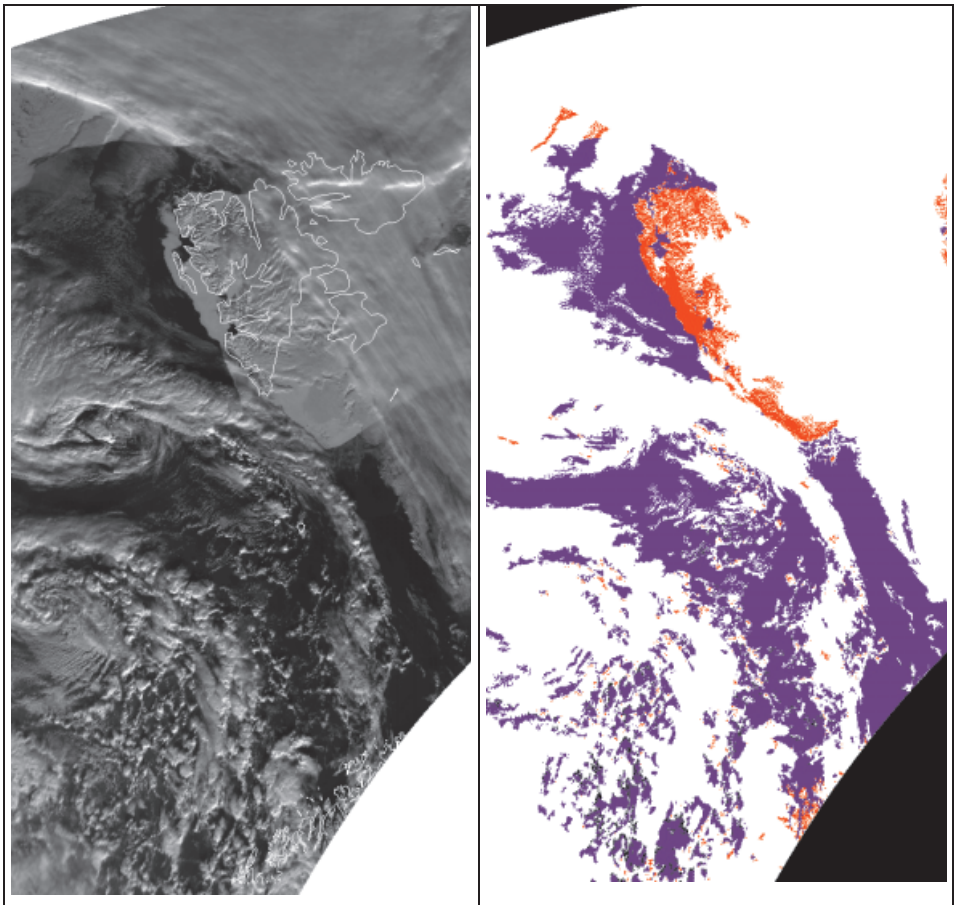

Figure 3a | Figure 3b

Figure 3. Results of the automatic classification algorithm for the 16 March 1998, 10.02 UTC image: (*a*) channel 1 albedo in percent (for the colour scale see figure 1), (*b*) pixel classification mask: cloudy pixel in white, open water pixel in blue, ice in red. The areas not covered by the AVHRR passage are in black (lat. 68–83 N, long. 0–30 E).

mask these situations are correctly classified. In the north-western part of the image, near the ice edge, the algorithm was also able to detect a very thin layer of clouds even if the ice beneath it is quite visible. The expert recognized as misclassified only a few isolated CLOUD pixels present over Svalbard. Here the roughness of the land surface covered by ice/snow increases the value of VART4 and the algorithm erroneously defines these pixels as clouds. The percentage error for this image is only 0.1% and it is due to the latter misclassification and to some open water pixels, near the cloud boundary in the southern part of the image, which remained as unclassified. This error is due to the cloud borders that contaminate the pixels modifying the surface radiances, thus excluding them from the preceding classes. Consequently, in the first step of the process the algorithm is not allowed to define these pixels as OPEN WATER.

Figure 4 shows the image corresponding to the satellite pass of 30 March 1998

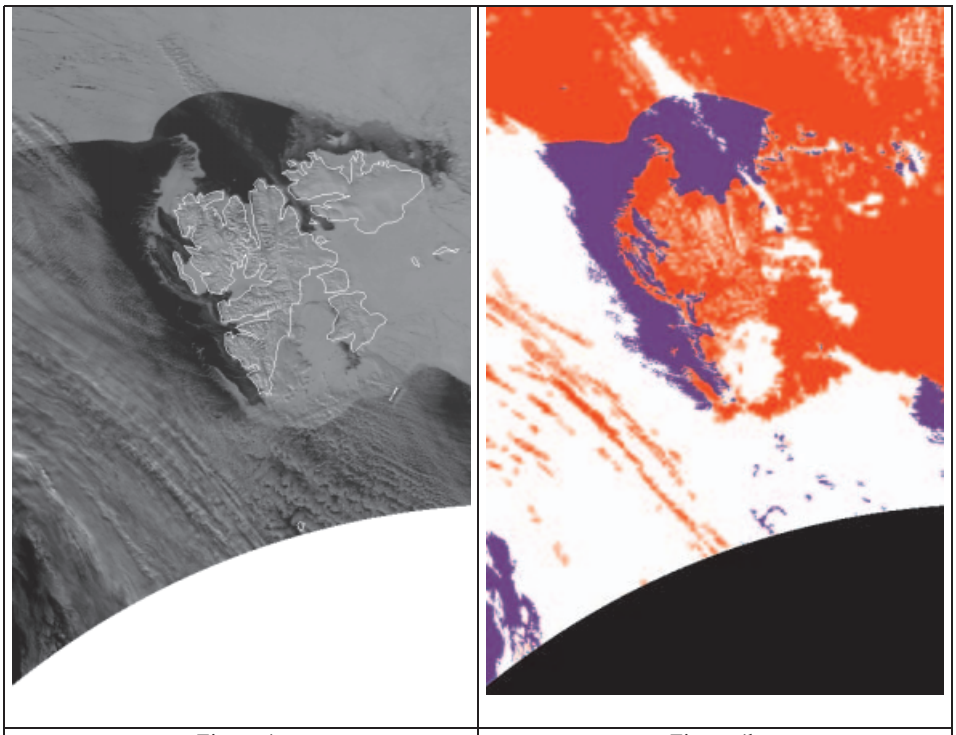

Figure 4a | Figure 4b

Figure 4. Results of the automatic classification algorithm for the 30 March 1998, 12.35 UTC image: (*a*) channel 1 albedo in percent (for the colour scale see figure 1), (*b*) pixel classification mask: cloudy pixel in white, open water pixel in blue, ice in red; areas not covered by the AVHRR passage are in black (lat. 72–82 N, long. 0–30 E).

at 12.35 UTC. In this case the ice pack covering the northern part of the image and the area between the Svalbards and the western limit of the image is free from clouds, as well as the Svalbard Islands themselves. The algorithm correctly recognizes the ice cover on both sea and islands and it is also able to identify a thin layer of low clouds over the ice pack that crosses the image from north-west to south-east. A very large layer of stratocumulus covers the majority of the open water and a small part of the ice south of Svalbard. The corresponding image mask correctly classifies the majority of these clouds. In the central part of figure 4(*a*) an opening polynya (area of open water surrounded by sea ice) is also visible; this polynya is known as Whaler's Polynya and it appears every year in the same period (Smith *et al.* 1990). The algorithm mainly classifies correctly most of these pixels as ice or open water, being the polynya surface not defined by our classification category. However, a thin strip of pixels in the central part of the polynya area is misclassified as clouds. The error for this image is 4.2% and it is due to the misclassification of some cloud streets west of the Islands (figure 4(*a*)) defined by the algorithm as ICE (figure 4(*b*)).

The digitalization of the misclassified areas in each of the 154 images allows us to quantify the error for each image and to identify the main error categories. The percentage error, resulting from this expert visual control, as function of the mean cloud cover is shown in figure 5. It is evident that in most of the images the error is

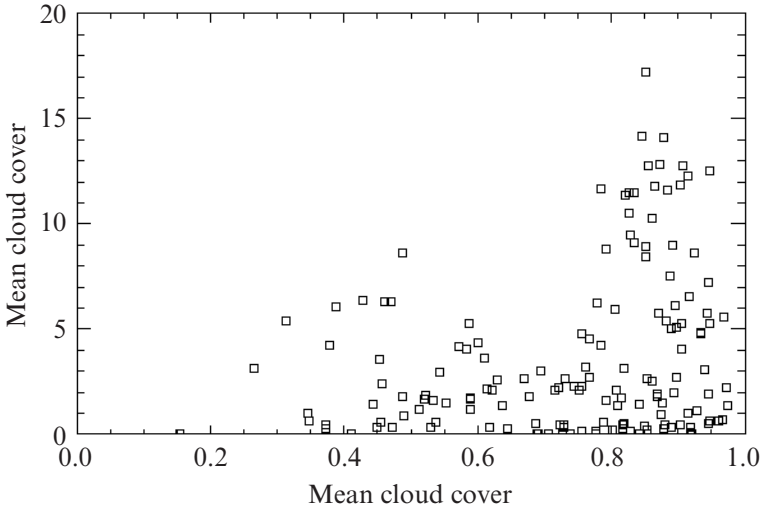

Figure 5. Percentage error resulting from the expert visual control as function of the image mean cloud cover.

less then 5% and it is uniformly distributed for all the cloud cover values. Only 11% of the images (17 images) result to have an error greater than 10%; the maximum error value is of 17.2% and it is evident that the error increases for images in which cloud cover is grater than 80%. The synthesis of the visual control reported in table 2 shows that the algorithm correctly classifies 13 images: 10 of these images are those used in the definition of the threshold values and they obviously have zero error. Thirty-six images have percentage error lower than 1%; in this case the error is due to some isolated pixels that are misclassified. For 60 images the error is between 1% and 5%. Table 2 confirms the good algorithm performance resulting from the qualitative inspection of figure 5.

Often ice is classified as clouds in the early morning images when the high value of the solar zenith angle causes high values of the normalized albedo, with an incorrect classification of the pixels as CLOUDS in the algorithm's first step. The same error occurs for the scenes relative to the end of the study period, when the ice pack is melting and the consequent increase in temperature variance tends to locate the pixels in the CLOUD class. On the contrary, cloud pixels are classified as ice, in presence of compact, homogeneous and low cloud cover. Another frequent

Table 2. Summary of the results of the visual expert evaluation.

Error percentage (%)	Number of images
$\% = 0$	10* + 3
$0 < \% \leqslant 1$	36
$1 < \% \leqslant 5$	60
$5 < \% \leqslant 10$	28
$\% > 10$	17

*10 is the number of the images used to define the limit values.

error concerns thin layers of cloud, or maybe haze, over ice: the algorithm 'sees' the ice under the clouds and then classifies the area as ICE. This kind of error can be considered as not serious in case of a further use of this cloud detection method as an input for a surface radiative budget. In fact, the presence of a thin layer of clouds or a layer of compact, homogeneous and thick clouds has a quite different influence on the incoming solar radiation.

4.2. *Comparison with meteorological total cloud cover observations at Ny-Ålesund station*

The image masks produced by the algorithm can be used to quantify the cloud cover fraction as it is usually visually measured by a meteorological observer. The cloud cover fraction is a meteorological variable provided by each meteorological station. The presence of the meteorological station in Ny-Ålesund allows for an independent control for the cloud detection algorithm by comparing satellite estimates of cloud cover around the station with direct observations (Goodman and Henderson-Sellers 1988).

To make the comparison suitable for the validation of the cloud detection method, we need to define the extension of the sky area seen by the meteorological observer. This area depends on the station position (altitude of the station, distance from obstacles, etc.) and visibility conditions. Defining a cell of $N \times N$ pixels cloud cover was obtained by the ratio of the number of pixels classified as CLOUD to the total number of pixels in the cell. The cloud cover estimate was reported in oktas, consistently with the meteorological observations. The dimension of the cell can be chosen taking into account the maximum radius of visibility for the surface observer that in general ranges from 30 to 55 km depending on cloud cover, station location and altitude (Malberg 1973). In the case of the Arctic region in which often fog or low clouds are present, the visibility of the observer can be smaller than 30 km. In order to check the cell size we increased N from 17 to 37 pixels and compared the AVHRR-derived total cloud cover with the meteorological observation in Ny-Ålesund.

Meteorological observations of total cloud cover were carried out at 6, 12 and 18 UTC every day by the Norsk Polarinstitutt at the Koldewey station in Ny-Ålesund and were transferred directly to the Global Telecommunication System (GTS) where they contribute to weather forecasting. The total cloud cover for meteorological observations is reported following the international code. Code number 9, corresponding to obscured sky, describes a meteorological situation (usually fog or snow precipitation) that does not allow a realistic description of the weather conditions. For our comparison we have only one case of class 9 and we have considered this case as completely overcast sky (class 8) (Rossow *et al.* 1993).

For the comparison we have used the satellite passes in a range of ± 1 h around the observation. The resulting total number of match-up points is 42.

In figure 6 the histogram of the error difference between satellite cloud amounts and observations as function of the cell dimensions is reported. It is evident that the difference is minimum for N equal 21 (with 21 cases of perfect correspondences and with a lower dispersion). This value indicates that in general the meteorological observer has a visibility around 20 km.

In figure 7 the observed cloud cover is plotted together with the satellite estimate for the 21×21 cell showing that 50% of the observations were correctly estimated. Most of the cases of perfect agreement are for total cloud cover (that is also a very

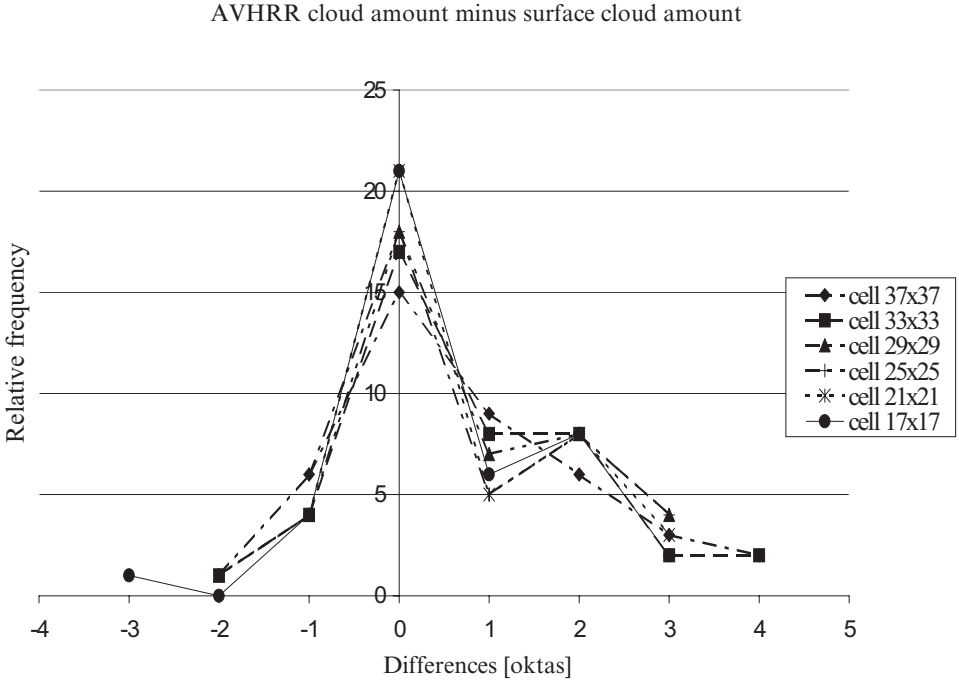

Figure 6. Difference between the AVHRR cloudiness estimates in a $N \times N$ pixel cell and cloud amount meteorological observations at Ny-Ålesund station as function of the cell N.

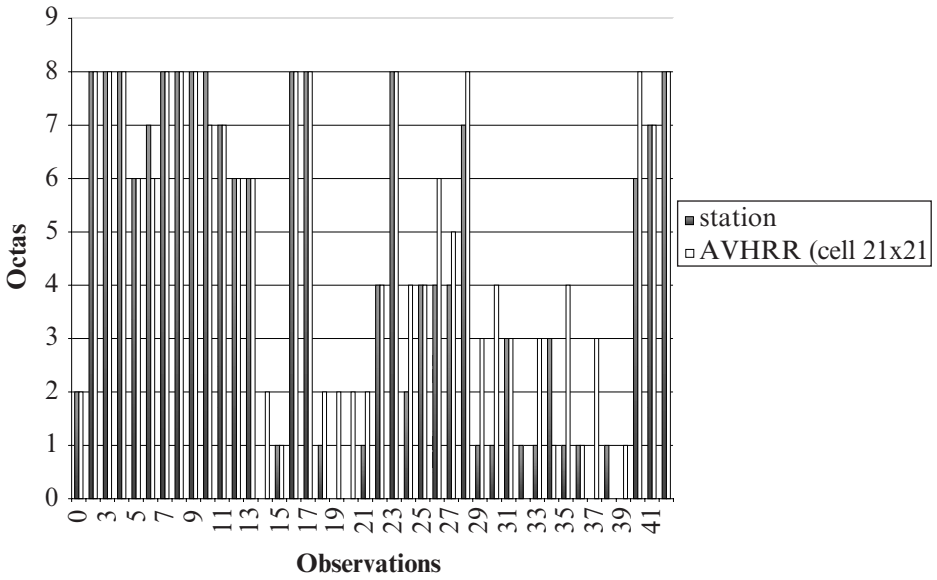

Figure 7. Comparison between the cloud cover meteorological observation (in grey) at the Ny-Ålesund station and the AVHRR cloud cover estimated (in white) in a 21×21 pixel cell centred on the station.

frequent situation in our match-up file, i.e. 10 observations). Moreover, if we consider an error of ± 1 oktas, which is the typical error for cloud cover meteorological observations (Seze *et al.* 1986, Henderson-Sellers *et al.* 1987, Rossow *et al.* 1993), the agreement percentage increases to 71%.

Figure 7 shows that in general, the AVHRR cloud cover tends to overestimate with respect to meteorological observations (see $N = 21$ case in figure 6). It is quite difficult to define the reasons of this difference: it is evident that both methods have specific problems (Malberg 1973, Henderson-Sellers *et al.* 1987). First of all the cloud cover meteorological observation is subjective and its error is function of the expertise of the observer, especially in case of presence of a partially covered sky. In case of sky cover by low clouds also an expert observer could not easily recognize the openings between clouds and in presence of fog or, in the polar case, of the Arctic haze, the surface evaluation of cloud amount could be very difficult. In general also the orographic situation of the station can induce a systematic error in the definition of the cloudiness. On the other hand, satellite cloud cover estimates are much more objective but their goodness depends on the error in the cloud detection introduced by the pixel classification algorithm. In general thin cirrus clouds are difficult to be recognized in a satellite scene by any cloud detection algorithm. Moreover the satellite horizontal resolution (1 km) does not permit to resolve sub-pixel cloud variations (Coakley and Bretherton 1982, Rossow *et al.* 1993).

An additional problem is introduced in our inter-comparison by the inexact time correspondence between ground and satellite observations. We have chosen a time interval of ± 1 h around the surface observation in order to maximize the number of match-up data and minimize the changes in the synoptic weather conditions.

Moreover, the Ny-Ålesund station is located on land covered by ice, which is a problematic case for our surface classification algorithm (see preceding section). The careful analysis of the data seems to indicate that the worst results of the comparison with the meteorological observations are obtained for those images with a high percentage error in the expert's evaluation. In particular the three cases in which a difference as high as three oktas was observed (satellite passes of 7 April 1998 at 11.06 UTC, 11 April 1998 at 05.15 UTC) the percentage errors resulting from the visual control were 5.7%, 8.6% and 6.3%, respectively. A careful analysis of their classification mask reveals that misclassified pixels are present in the 21×21 pixel box centred on the Ny-Ålesund station. This can give a justification for the large discrepancy that occurred in these cases.

4.3. *Sea-ice maps: comparison with SSM/I sea-ice concentration maps*

The quality of sea-ice maps, produced by the AVHRR surface classification algorithm for cloud-free regions, can be controlled using SSM/I sea-ice concentration maps at $0.25 \times 0.25^{\circ}$ spatial resolution obtained from NSIDC. For this end AVHRR sea-ice masks have been re-binned each 25×25 pixels in order to achieve a spatial resolution similar to NSIDC SSM/I maps and to be able to estimate for each 25×25 pixel box a percent of sea-ice pixels with respect to the total number of cloud free pixels in the box. If the percent of ice-flagged pixels was lower than 10% the cell is classified as open water. Of course the presence of clouds does not permit to produce daily AVHRR sea-ice concentration maps. For this reason we limited our analysis to the comparison of SSM/I and AVHRR 10 day average maps.

Figures 8 and 9 show the ice distribution deduced from AVHRR and SSM/I, respectively for the three 10-day periods of the ARTIST experiment. In general the

spatial distribution of sea-ice obtained from AVHRR data is very similar to the SSM/I concentration maps even if some details of the transition area between Arctic pack and open sea appears better described in the SSM/I maps.

The observation of the evolution of the sea-ice limit moving from the first 10-day period (figure 8(*a*)) to the second (figure 8(*b*)) shows that some ice melted in the northern part of the Svalbard Islands. In the third 10-day window (figure 8(*c*)) this tendency is interrupted and the area north of the Svalbard Islands appears to be completely covered by ice again. The same behaviour is observed in the SSM/I time series where ice concentration in the same area decreases from the first to the second period and increases again in the third period. One stimulating though speculative hypothesis is that the inversion of the seasonal tendency could be due to some contribution of ice moving south after the breakup of the edge of the polar pack or ice reforming during a wind event. Looking at the southern ice edge in the AVHRR images, it can be observed that it is shifting northwards going from the first to the last 10-day period, according to the seasonal evolution. In figure 8(*c*) a relatively large area of sea-ice surrounded by open water is present at the western edge of the image. The comparison with the corresponding SSM/I sea-ice map (figure 9(*c*)) confirms the idea that this is not an artefact of the AVHRR cloud detection algorithm, even though the concentration is quite different.

This can be considered as an indirect validation of the cloud detection algorithm. Also AVHRR data can then be used to map the seasonal evolution of the limits of

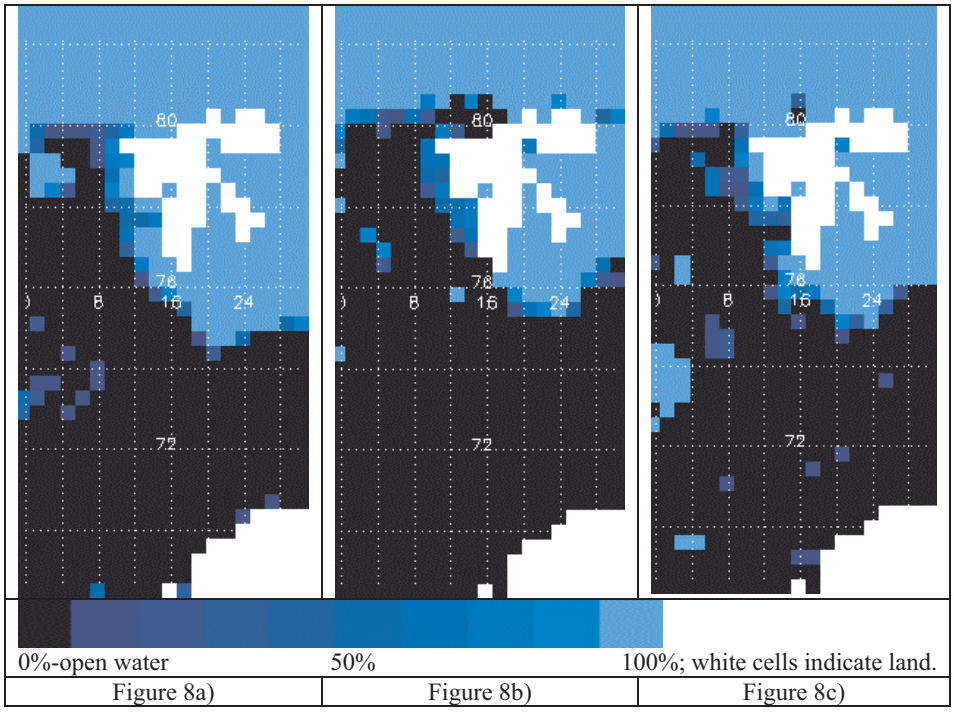

Figure 8. Percent of ice pixels (relatively to total cloud free pixels) selected by the AVHRR classification algorithm: (*a*) map for the period 16 March 1998 to 25 March 1998; (*b*) map for the period 26 March 1998 to 4 April 1998; (*c*) map for the period 5 April 1998 to 15 April 1998 (lat. 68–83 N, long. 0–30 E).

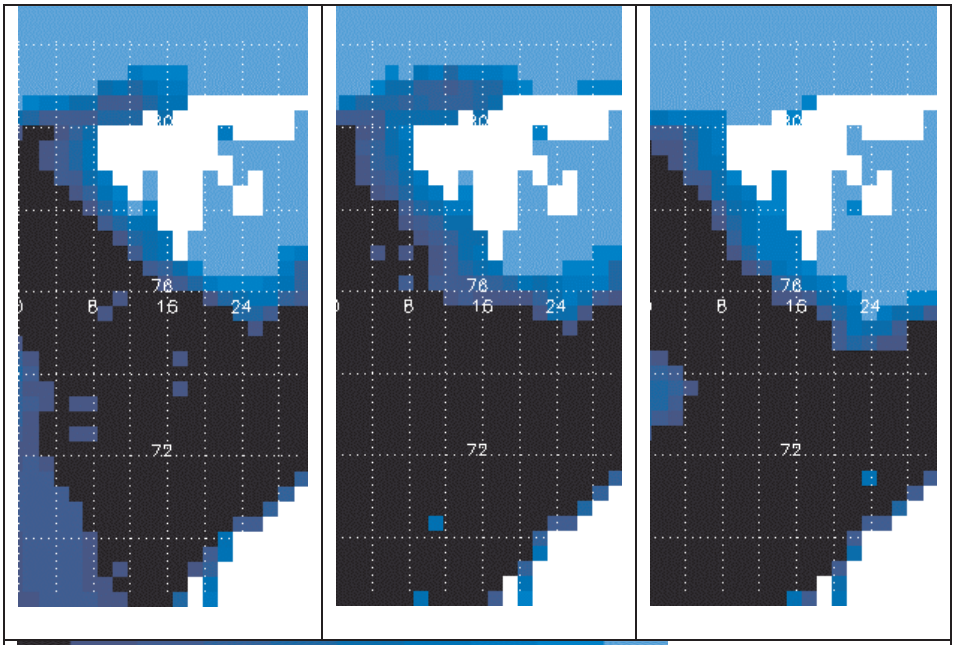

Figure 9. Ten-day mean sea-ice concentration from SSM/I data: (*a*) map for the period 16 March 1998 to 25 March 1998; (*b*) map for the period 26 March 1998 to 4 April 1998; (*c*) map for the period 5 April 1998 to 15 April 1998 (lat. 68–83 N, long. 0–30 E).

the sea-ice boundary line with high spatial resolution to complement other information deduced from SSM/I. The last major difference is in the south-west corner of the first 10-day average images, where SSM/I displays an unexpected sea-ice concentration of 10–20%.

In order to make the comparison more objective, contingency tables for each couple of 10-day maps and for the global comparison are provided (table 3) with the corresponding statistical coefficients (table 4). Being the AVHRR estimate of the mean ice concentration percent strongly biased by cloud cover, the contingency table was computed dividing the pixels in only two classes: ice and open water. Due to the very different spatial resolution of AVHRR and SSM/I data and to some error in the SSM/I land mask, the land regions in two datasets do not exactly overlap. Thus the land regions are defined by those pixels classified as land in one or both of the two datasets.

Usually, the contingency table is used to describe the accuracy of the forecast with respect to the observations. In our case we take as 'forecast' the AVHRR classification in ice or open water, and as 'observations' the SSM/I classification of the same two classes. The total number of comparisons is 2085.

A perfectly match between the two datasets will exhibit null values for the diagonal, where the two numbers 107 and 0 are seen in table 3 for the final total comparison. For an imperfect comparison, further parameters can be used to assess

Table 3. Contingency tables for the three ice distribution maps and for final global comparison.

Figure 10(*a*) and 11(*a*): first decade		SSM/I ICE	SSM/I OPEN WATER	Marginal total for AVHRR
AVHRR	ICE	315	0	315
	OPEN WATER	41	339	380
	Marginal total for SSM/I	356	339	Total number 695

Figure 10(*b*) and 11(*b*): second decade		SSM/I ICE	SSM/I OPEN WATER	Marginal total for AVHRR
AVHRR	ICE	247	0	247
	OPEN WATER	49	399	448
	Marginal total for SSM/I	296	399	Total number 695

Figure 10(*c*) and 11(*c*): third decade		SSM/I ICE	SSM/I OPEN WATER	Marginal total for AVHRR
AVHRR	ICE	304	0	304
	OPEN WATER	17	374	391
	Marginal total For SSM/I	321	374	Total number 695

Final total comparison		SSM/I ICE	SSM/I OPEN WATER	Marginal total for AVHRR
AVHRR	ICE	866	0	866
	OPEN WATER	107	1112	1219
	Marginal total For SSM/I	973	1112	Total number 2085

the degree of accuracy (Wilks 1995). The 'hit rate' (H), or 'correct proportion' for the comparison is $H = 0.9486$. This parameter is the fraction of the total realization in the sample ($n = 2085$ in our case) for which the forecast correctly anticipated the subsequent event or non-event. Another coefficient that can be taken into account is the 'critical success index' (CSI), i.e. the number of correct 'ice' forecasts divided by the total number of occasions on which that event was forecast and/or observed; the worst possible index is 0, the best is 1. In our case CSI = 0.89. Next, the 'probability of detection' (POD) is the ratio of correct forecasts to the number of occasions in

which the event occurred. In our case POD is equal to CSI because the term in the first row and second column is zero. The last coefficient is the 'false-alarm rate' (FAR), i.e. the ratio of failed forecasts to total forecasts for, e.g. the 'ice' event. FAR expresses failures so smaller values of FAR are to be preferred. With our data FAR = 0, which is the best-obtained result. In table 4 are reported all the coefficients corresponding to each contingency table.

Overall, it seems that the statistical analysis has provided good results and thus represents a further validation for the cloud detection algorithm.

5. Conclusions

A cloud detection algorithm for the Arctic region was implemented using AVHRR data relative to the ARTIST field phase (from 16 March to 15 April 1998). The proposed algorithm is a simplified version of the Ebert (1987) method and was developed to work on AVHRR scenes, typical of the transition seasons in the Arctic region. The algorithm classifies each pixel of the AVHRR scene in three classes: ice over sea or land, open water and clouds. As in every threshold method, the algorithm is based on the definition of the limit values of some selected satellite derived parameters (e.g. albedos, temperature, etc.) for each defined class. To develop the algorithm we used 10 representative satellite passes and the threshold values have been defined after several trials, analysing the effect of the chosen limits on the pixel classification of the 10 selected images.

The algorithm was applied to the 154 available images and for each satellite pass a flag mask of cloud, ice and open water was produced. The results were then subjected to three different controls.

The pixel classification algorithm performance was evaluated by an expert in satellite imagery. The areas of misclassified pixels were digitalized by the expert on the image and the pixels were counted. This permitted to quantify the error resulting from the expert's visual control. The pixel classification results to be quite accurate: 70% of the images (109) has an error less than 5% and only 11% of the images (17) has errors greater than 10%. The maximum error value is 17.2%.

The analysis of the misclassified areas in each of the 154 images allowed us not only to quantify the error for each image but also to identify the main error categories resulting from the expert visual control (§4.1). The analysis of the error distribution as function of cloud cover averaged on the entire AVHRR scene (figure 5) shows that the images with error less 5% are uniformly distributed. The error increases only for cloud cover greater than 80%. This is not surprising because the expert visual control indicates as the most frequent error the misclassification of cloudy pixels as ice and vice versa.

Although we try to make the visual control as objective as possible, it is obviously subjective and depends on the capability of the image observer to distinguish the

Table 4. Coefficients relative to the contingency tables.

	First decade	Second decade	Third decade	Total
H	0.9410	0.9294	0.9755	0.9486
CSI (= POD)	0.8848	0.8344	0.9470	0.8900
FAR	0	0	0	0

different surfaces. In any case, this visual control gives a global judgement on the entire image and a general idea of the error types of the method. In order to give further evaluation of the algorithm performance cloud cover and ice cover images were produced from the flagged image and compared with independent observations.

A comparison with the cloud cover meteorological observations made from the land station at Ny-Ålesund can be considered as an objective method to define the quality of the results, even though the meteorological observations themselves may have some intrinsic problems (see §4.2 for a brief discussion). Satellite estimations of cloud cover in an $N \times N$ pixel cell around Koldewey (Ny-Ålesund) meteorological station were compared with meteorological observations and confirmed the good performance of the algorithm. In fact, using a 21×21 pixel cell, 50% of the observations was perfectly estimated by satellite. If we consider an error of ± 1 oktas, i.e. is the typical error of cloud amount observations, the agreement percentage increases to 71%. In only three cases we observed a difference as high as three oktas.

The last quality control is the comparison between the 10-day sea-ice cover maps (obtained from the ice, cloud and open water masks) and the sea-ice distribution maps obtained from SSM/I data. The visual qualitative comparison shows a very similar distribution between AVHRR and SSM/I sea-ice cover. The use of contingency tables and related coefficients confirms the good result of the comparison. This further indirect validation of the proposed pixel classification method seems to indicate the possibility to use the ice mask resulting from the application of the algorithm to compute cloud-free ice-cover maps by averaging 10 days' images. Then AVHRR data can be used to map the seasonal evolution of the limits of the sea-ice boundary line with high spatial resolution to complement other information deduced from SSM/I.

It is important to underline some conclusive and remarkable points regarding the presented algorithm. First, the auto-consistency of the algorithm requires no other additional data source for pixel classification. Secondly, this method requires a limited number of parameters and processing steps to classify pixels in the AVHRR scenes, thus a low computing time. Moreover, the small number of thresholds involved in this method permits an easy tuning of the proposed algorithm for the other condition. This can considerably reduce its main limitation, i.e. the choice of fixed thresholds that can diminish the obtained performances in other conditions (Goodman and Henderson-Sellers 1988). Consequently, this represents a very good methodology easily improved and applicable to other contexts. Looking at the ISCCP (International Satellite Cloud Cover Climatology) results it is well known that the largest errors in cloud cover and properties actually occur in the polar regions (Rossow *et al.* 1993). The uniqueness of the polar cloudiness and, more in general of the meteorological and climatic conditions, lead research toward a more specific study of these areas.

An obvious, but not so straightforward, improvement of this cloud detection algorithm could be its application to night-time images, stressing more the use of the infrared channels and the introduction of channel 5 in the analysis procedure. This work represents our next effort.

References

Allen, R. C., Durkee, P. A., and Wash, C. H., 1990, Snow/cloud discrimination with multispectral satellite measurements. *Journal of Applied Meteorology*, **29**, 994–1004.

Cavalieri, D. J., Gloersen, P., and Campbell, W. J., 1984, Determination of sea ice parameters with Nimbus-7 SMMR. *Journal of Geophysical Research*,, 5355–5369.

Coakley, J. A., and Bretherton, F. P., 1982, Cloud cover from high-resolution scanner data detecting and allowing for partially filled fields of view. *Journal of Geophysical Research*, **87**, 4917–4932.

Ebert, E., 1987, A pattern recognition technique for distinguishing surface and cloud types in the Polar regions. *Journal of Climate and Applied Meteorology*, **26**, 1412–1427.

Goodman, A. H., and Henderson-Sellers, A., 1988, Cloud detection and analysis: a review of recent progress. *Atmospheric Research*, **21**, 203–228.

Henderson-Sellers, A., Seze, G., Drake, F., and Desbois, M., 1987, Surface-observed and satellite-retrieved cloudiness compared for the 1983 ISCCP special study area in Europe. *Journal of Geophysical Research*, **92**, 4019–4033.

Key, J., and Barry, R. G., 1989, Cloud cover analysis with Arctic AVHRR data: 1. Cloud detection. *Journal of Geophysical Research*, **94**, 18521–18535.

Key, J., 1990, Cloud cover analysis with Arctic Advanced Very High Resolution Radiometer Data 2. Classification with spectral and textural measures. *Journal of Geophysical Research*, **95**, 7661–7675.

Kidder, S. Q., and Wu, H., 1984, Dramatic contrast between low clouds and snow cover in daytime 3.7 μm imagery. *Monthly Weather Review*, **112**, 2345–2346.

Kilpatrick, K. A., Podesta, G. P., and Evans, R., 2001, Overview of the NOAA/NASA AVHRR Pathfinder algorithm for Sea Surface Temperature and associated matchup database. *Journal of Geophysical Research*, **106**, 9179–9197.

Lubin, D., and Morrow, E., 1998, Evaluation of an AVHRR cloud detection and classification method over the Central Arctic Ocean. *Journal of Applied Meteorology*, **37**, 166–183.

Malberg, H., 1973, Comparison of mean cloud cover obtained by satellite photographs and ground based observations over Europe and the Atlantic. *Monthly Weather Review*, **101**, 893–897.

Rao, C. R. N., and Chen, J., 1995, Inter-satellite calibration linkages for the visible and near infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-7, -9, -11 spacecraft. *International Journal of Remote Sensing*, **16**, 1931–1942.

Robinson, I. S., 1985, *Satellite Oceanography* (Chichester: Ellis Horwood).

Rossow, W. B., Garder, L. C., and Lacis, A. A., 1989, Global, seasonal cloud variations from satellite radiance measurements: I. Sensitivity of analysis. *Journal of Climate*, **2**, 419–458.

Rossow, W. B., Walker, A. W., and Garder, L. C., 1993, Comparison of ISCCP and other cloud amounts. *Journal of Climate*, **6**, 2394–2418.

Seze, G., Drake, F., Desbois, M., and Henderson-sellers, A., 1986, Total and low cloud amounts over France and southern Britain in the summer of 1983: comparison of surface-observed and satellite-retrieved values. *International Journal of Remote Sensing*, **7**, 1031–1050.

Smith, S. D., Muench, R. D., and Pease, C. H., 1990, Polynyas and leads: an overview of physical processes and environment. *Journal of Geophysical Research*, **95**, 9461–9479.

Swift, C. T., and Cavalieri, D. J., 1985, Passive microwave remote sensing for sea ice research. *Earth Observation System*, **66**(49), 1210–1212.

Tsay, S. C., Stamnes, K., and Jayaweera, K., 1989, Radiative energy budget in the cloudy and hazy Arctic. *Journal of Atmospheric Sciences*, **46**, 1002–1018.

Welch, R. M., Sengupta, S. K., Goroch, A. K., Rabindra, P., Rangaraj, N., and Navar, M. S., 1992, Polar cloud and surface classification using AVHRR imagery: an intercomparison of methods. *Journal of Applied Meteorology*, **31**, 405–420.

Wilks, D. S., 1995, *Statistical Methods of the Atmospheric Sciences* (New York: Academic Press).