Daytime Arctic Cloud Detection based on

Multi-angle Satellite Data with Case Studies

Tao Shi, Bin Yu, Eugene E. Clothiaux, and Amy J. Braverman

Abstract

Global climate models predict that the strongest dependences of surface air temperatures on increasing atmospheric carbon dioxide levels will occur in the northern high latitudes of Alaska, Greenland, and eastern Siberia (Giorgi and Bi, 2005). A systematic study of these dependences requires accurate Arctic-wide measurements, especially of cloud coverage as gains or losses in cloud coverage can, in turn, affect surface air temperatures. Cloud detection in the Arctic is extremely important for regional climate studies across the Arctic, but it is also quite challenging because of the similar remote sensing characteristics of clouds, ice- and snow-covered surfaces.

This paper proposes two new operational arctic cloud detection algorithms using Multi-angle Imaging SpectroRadiometer (MISR) imagery. The key idea embedded in these algorithms is to search for ice- and snow-covered surface pixels in the MISR imagery instead of cloudy pixels directly, as in the MISR operational algorithms. Through extensive exploratory data analysis, three physically useful features have been identified that differentiate well surface pixels from cloudy pixels for MISR images with different sunlight and weather conditions. They are the correlation of MISR images of the same scene from different MISR viewing directions, the standard deviation of pixel values across a scene, and a Normalized Differential Angular Index ($NDAI$) that characterizes the changes is a scene with changes in the MISR view direction. The first algorithm based on these three features, Enhanced Linear Correlation Matching (ELCM), thresholds the three features with two fixed cut-off values for the correlation and standard deviation features and one data-adaptive value for the $NDAI$ feature. Probability labels are obtained by using ELCM labels as training data for Fisher's Quadratic Discriminant Analysis (QDA), leading to the second (ELCM-QDA) algorithm. Both algorithms are automated, simple, and fast for operational processing of MISR data. The best available validation data, expert labels, are used to test the two algorithms.

Based on five million test pixels ELCM results are significantly better both in terms of accuracy (92%) and coverage (100%) when compared with two MISR operational algorithms,

one with an accuracy of of 80% and coverage of 27% and the other with an accuracy of 83% and a coverage of 70%. Our algorithms provide the best performances to date among all available operational algorithms using these MISR test data. The ELCM-QDA probability prediction is also consistent with the expert labels and more informative in our study. Expert labelled pixels mostly have ELCM-QDA probabilities very close to 1 or 0, and the pixels with probability predictions close to 0.5 all fall in the areas of images where the expert dose not have high confidence.

KEY WORDS: Classification; Clustering; Multi-angle Imaging SpectroRadiometer (MISR); Feature selection; Quadratic Discriminate Analysis (QDA); Remote sensing;

1 INTRODUCTION

Sensitivity of Earth's climate to increasing amounts of atmospheric carbon dioxide is a topic of general and scientific interest, and an important public policy issue as well. Today's global climate models (GCMs, or AOGCMs for coupled atmosphere-ocean global climate models) generally predict that global surface air temperatures will increase by 1.5 K to 3.5 K with a doubling of atmospheric carbon dioxide levels throughout the 21st century (Giorgi and Bi, 2005). These same models also predict that the strongest dependences of surface air temperatures on increasing atmospheric carbon dioxide levels will occur in the Arctic (Giorgi and Bi, 2005). As the Arctic warms, changes in the properties and distribution ice- and snow-covered surfaces, atmospheric water vapor and clouds can potentially lead to further warming and hence strong sensitivity to increasing amounts of atmospheric carbon dioxide. A systematic study of these dependences requires accurate Arctic-wide measurements (e.g., Francis et al. 2005), especially of cloud coverage as clouds play an important role in modulating the sensitivity of the Arctic to increasing surface air temperatures

3

(e.g., Kato et al. 2006).

Ascertaining the properties of clouds in the Arctic is a challenging problem because liquid- and ice-water cloud particles often have similar scattering properties to the particles that compose ice- and snow-covered surfaces. As a result, the amount of visible and infrared electromagnetic radiation emanating from clouds and snow- and ice-covered surfaces is often similar, which leads to problems in the detection of clouds over these surface types. Without accurate characterization of clouds over the Arctic we will not be able to assess their impact on the flow of solar and terrestrial electromagnetic radiation through the Arctic atmosphere and we will not be able to ascertain whether they are changing in ways that enhance or ameliorate future warming in the Arctic.

With the launch of the Multi-angle Imaging SpectroRadiometer (MISR) onboard the National Aeronautics and Space Administration (NASA) Terra satellite in 1999 novel electromagnetic radiation measurements made at nine angles became available for scientific study. The MISR sensor consists of nine cameras (Figure 1), with each camera viewing Earth scenes at a different angle in four spectral bands (blue, green, red, and near-infrared). The view zenith angles of the nine cameras are 70.5° (Df), 60° (Cf), 45.6° (Bf), and 26.1° (Af) in the forward direction, 0.0° (An) in the nadir direction and 26.1° (Aa), 45.6° (Ba), 60° (Ca) and 70.5° (Da) in the aft direction. The "f" in the letter designation of the cameras represents the "forward" direction and the "a" represents the "aft" direction. The Da camera collects data from a location seven minutes after the Df camera. The nominal resolution of the MISR radiances is 275 m by 275 m at the Earth's surface, but the blue, green and near-infrared radiances are aggregated onboard to a 1.1 km by 1.1 km resolution to reduce data transmission from the Terra satellite.

The MISR cameras cover a swath at the Earth's surface that is approximately 360 km wide and extends across the daylight side of the Earth from the Arctic down to Antarctica in about

forty-five minutes. There are 233 geographically distinct, but overlapping, MISR swaths, which are also called **paths**. MISR collects data from all paths on a repeat cycle of 16 days; that is, it covers the exact same path every 16 days. In the MISR data products each path is subdivided into 180 **blocks**, with the block numbers increasing from the north to south pole. Each complete trip of MISR around the earth is given its own **orbit** number.

The MISR angular radiances are completely novel and from the start exhibited strong indications of possible separability of clouds from snow- and ice-covered surfaces (Diner et al. 1999a). However, the MISR operational cloud detection algorithms (see details in Section 2.1) were designed before MISR was launched and they were not particularly targeted at detecting clouds over bright surfaces in polar regions. As a result, MISR operational algorithms do not work well over polar regions. Moreover, one orbit of data needs to be processed before the next orbit comes, so an operational cloud detection algorithm is necessary. The massive MISR data size (MISR collects 3.3 megabits per second on average and 9.0 megabits per second at peak) poses a hurdle for operational processing. To balance processing speed and the possibility of having both clear and cloudy pixels in one processing unit, we choose to consider three blocks of MISR images as one "**data unit**". As we show in the paper, this choice is validated empirically by our test results against expert labels.

Cloud detection in the Arctic is not only scientifically important, but also statistically challenging. Any classification framework is not readily applicable because it is impossible to get expert labels operationally, considering the size of the data and resource time constraints. A pure clustering framework is not suitable either since each data processing unit could be totally cloud covered or fully cloud free (e.g., Figure 4). When a data unit is fully cloud covered, the clusters correspond to different types of cloudy pixels, and not whether a pixel is cloudy or clear. Therefore, the challenge is how to combine classification and clustering and in a computationally efficient manner.

Our goal is to build operational cloud detection algorithms that can efficiently process the massive MISR data one data unit at a time. We use extensive exploratory data analysis to construct a labelling scheme, or detector, for each data unit by combining clustering information on the current data unit with the labelling scheme learned from the data unit collected on the prior (16 day earlier) orbit. That is, we sequentially build detectors to process operationally the massive MISR data. Our novel idea is to search for ice- and snow-covered surface image pixels instead of cloudy pixels, which is used in the MISR operational algorithms. We will show that this simple reversal of the detection target is fruitful. Modelling the surface has an advantage over modelling clouds in this clustering and classification framework because the surface does not change much between consecutive 16-day visits, while clouds are always different from one visit to the next. Specifically, our algorithms are based on three physically useful features – the correlation ($CORR$) of MISR images of the same scene from different MISR viewing directions, the standard deviation SD_{An} of pixel values across a scene, and a Normalized Differential Angular Index ($NDAI$) that characterizes the changes in a scene with changes in the MISR view direction – for characterizing the scattering properties of ice- and snow-covered surfaces. They are arrived at with ample exploratory data analysis by both statisticians and atmospheric scientists.

We propose an Enhanced Linear Correlation Matching (ELCM) algorithm based on thresholding the three feature with values either fixed or learned from the current or previous data unit. We envision that the ELCM algorithm is applied to MISR data units sequentially. The $CORR$ and SD_{AN} feature cut-off values, or thresholds, are set to fixed values during the operational processing because they are stable over data collected from different times and locations. The $NDAI$ threshold is initially set based on off-line analysis of the first set of data units. To classify subsequent data units the $NDAI$ threshold is either kept at its previous value or updated using a data-adaptive

6

algorithm applied to the current data unit. Labels resulting from the ELCM algorithm are then used to train the Fisher's Quadratic Discriminant Analysis (QDA) in order to produce probability labels, leading to our second (ELCM-QDA) detector. The QDA step produces the probability of "cloudiness" as a more informative probability prediction, although it does not necessarily improve overall prediction accuracy.

To assess the accuracies of the ELCM and ELCM-QDA algorithms, they are applied to extensive MISR data and tested against expert labels provided by one of the authors (Eugene Clothiaux). The expert labels are provided for 10 orbits of MISR data collected over the Arctic during the daylight season of 2002, and they represent the best available validation data to date in adequate quantities (5.086 million 1.1 km pixels).

The rest of the paper is organized as follows. Section 2 contains a brief review of MISR algorithms and introduces the MISR data investigated in this study. Section 3 first describes our proposed methodology, starting with our rationale for the study and computation of the features. Then (Section 3.2) we propose our Enhanced Linear Correlation Matching (ELCM) algorithm that combines the clustering information in the current data unit with the detection rule learned from the data unit on the previous visit. In Section 3.3 we demonstrate that probability predictions can be obtained over the party cloudy scenes if we use ELCM labels to train Fisher's Quadratic Discriminative Analysis (QDA). Section 4 contains the testing results. ELCM and ELCM-QDA results are compared against the expert labels along with the MISR operational algorithms. Strengths and weaknesses of our algorithms are explored. Finally, the expert labels are used to illustrate that our features are better than the raw radiance measurements in the sense of "separability" and "stability". Section 5 concludes the paper with a summary and a discussion of the potential scientific impacts of our algorithms.

7

2 MISR OPERATIONAL ALGORITHMS AND DATA DESCRIPTION

Viewing the atmosphere and surface from multiple angles, MISR has stereo capabilities that can be used to retrieve the elevation of objects, such as clouds, on or above the surface of the Earth. In addition to stereo information, MISR non-nadir camera angles provide the radiation scattering patterns of different objects (e.g. clouds). These two sources of novel information motivated the MISR Level 2 Top-of-atmosphere Cloud (L2TC) algorithm (Diner et al. 1999b), which produces two cloud masks: the Stereo Derived Cloud Mask (SDCM) and the Angular Signature Cloud Mask (ASCM). In Section 2.1, the MISR operational algorithms and their shortcomings in polar regions are briefly reviewed. In Section 2.2, we describe the data used in this study.

2.1 MISR Operational Algorithms

The L2TC algorithm uses MISR's stereo capability to detect clouds by comparing retrieved object heights to the underlying known terrain height. The rationale of the L2TC cloud height retrieval algorithm is based on the registration of MISR measurements to a known Earth-based reference ellipsoid. As shown in Figure 2, a cloud is registered to different reference ellipsoid locations for the different MISR cameras. The L2TC algorithm matches the same object in the different angle-dependent images, allowing object heights and horizontal velocities to be retrieved through simple trigonometric relationships. The SDCM is derived by comparing retrieved object heights with the known terrain heights and objects more than approximately 650 m above the terrain height are classified as clouds.

The ASCM is based on the Band- Differenced Angular Signature (BDAS), which is the difference between two solar spectral reflectances as a function of view angle (Di Girolamo and Davies 1994;

Diner et al. 1999b). Based on the relative contribution of Rayleigh scattering to total reflection, the longer solar ray paths are through Earth's atmosphere before reaching MISR, the larger the BDAS. Since radiation reflected from clouds has much shorter paths than radiation reflected from the Earth's surface, small BDAS indicates clouds. Over ice- and snow-covered surfaces in the Arctic, the difference between Bidirectional Radiance Factors (BRFs) of blue radiation form the Df camera and near infrared radiation from the Cf camera is thresholded to distinguish clouds from the surface.

Both algorithms have difficulties in polar regions, especially for low clouds that are often present in daytime Arctic regions. The SDCM can not detect low clouds because the distances between low clouds and the surface are often smaller than MISR height retrieval accuracy. The ASCM is good at detecting high and thin clouds, but also has difficulties detecting low clouds over terrain because the Rayleigh scattering contribution to BDAS increases rapidly with a decrease in cloud-top height (Di Girolamo and Davies 1994). Our approach, to be proposed in Section 3, attempts to overcome these shortcomings of the SDCM and ASCM by searching for surface pixels instead of cloudy ones.

2.2   Data

The data used in this study are collected from 10 MISR orbits of path 26 over the Arctic, northern Greenland and Baffin Bay. As we described earlier, the repeat time between two consecutive orbits over the same path is 16 days, so the 10 orbits span approximately 144 days from April 28 through September 19, 2002 (a daylight season in the Arctic). Path 26 was chosen for the study because of the richness of its surface features, which include permanent sea ice in the Arctic Ocean, snow-covered and snow-free coastal mountains in Greenland, permanent glacial snow and ice, and sea ice that melted across Baffin Bay over the 144 days. Six data units (MISR blocks 11–28) from

each orbit are included in this study. (Three of the sixty data units were actually excluded from this study because they lacked completely any ice- and snow-covered features.) Therefore, the data investigated contain 57 data units with 7,114,248 valid 1.1 km resolution pixels with 36 radiation measurements for each pixel. We use the 275 m red band radiation measurements to build some of our features so the actual data size is much larger. Our study concentrates on repeated visits over time so the expert can learn about the surface features in a data unit. This surface information helps the expert provide accurate labels for validation purposes.

Figure 3 shows images of three data units collected from MISR blocks 20–22 over three consecutive orbits (i.e. 13257, 13490, and 13723) of path 26. At first look it is hard to tell if different patches of images are clouds or packs of ice and snow. The images show that the ice- and snow-covered surfaces are as bright as, or even brighter than, clouds, which contradicts the traditional assumption, which is applicable at lower latitudes, that clouds are brighter than surfaces. As a result, traditional cloud detection algorithms based on this assumption can not be applied in polar regions. Furthermore, the differences between different patches of images are not a necessary result of cloud versus ice and snow because an entire data unit might be fully cloud covered or fully clear. An example of such cases is provided in Figure 4: fully clear in blocks 20–22 of orbit 12325 (Fig. 4a) and fully cloudy in blocks 14–17 of orbit 13490 (Fig. 4b).

To evaluate the performance of our proposed methods and existing MISR operational algorithms, one of the authors (Eugene Clothiaux) hand labelled the data. Because high quality ground-based measurements are rare in the polar regions (at most one pixel in each orbit) and satellite-based cloud radars and lidars do not yet exist, expert labelling of clear and cloudy scenes is currently the best method for producing validation data for assessing automated polar cloud detection algorithms. After the expert determines the labels (i.e. clear or cloudy) for a patch

of image pixels, tools developed by the MISR science team at Jet Propulsion Laboratory, called "misrdump" and "misrlearn" (Dominic Mazzoni, personal communication), are subsequently used to label the pixels in MISR nadir camera images as clear or cloudy. For example, labels for data shown in Figure 3 are plotted in Figure 5, with white for cloudy and gray for clear. Expert labels are only given to the pixels for which the expert is highly confident based on his knowledge of the measurements, with more ambiguous regions left unlabelled (black in Figure 5). As a result of this conservative labelling scheme, the expert labels covers about 71.5% (5,086,002) of the total valid pixels. These labels will be used in this paper to evaluate the performance of different cloud detection algorithms. Since expert labels are not available operationally, an operational cloud detection algorithm can not depend on such expert labels.

3 METHODS

We propose two new cloud detection algorithms for the Arctic that are automated, simple and fast for operational processing of the massive MISR data. Both algorithms are based on three physically useful features for identifying surface pixels, thereby obtaining cloudy pixels by exclusion. Each pixel is treated as independent of the others because the computational cost of most spatial models is too high to be implemented in operational data processing. Spatial smoothness of surface and cloudy pixels is taken into account implicitly in our algorithms, for two of our three features are based on local patches of pixels at the 275-m resolution. The three features were derived through extensive exploratory data analysis. For each data unit, we develop a clustering algorithm, Enhanced Linear Correlation Matching (ELCM), by thresholding the three features. For the first data unit, we find the best thresholds for the three features using the expert labels off-line. For subsequent data units from the same block range ELCM results for the data unit are combined with

the detectors learned from data collected on the previous visit. The success of this combination method relies heavily on all three features. The ELCM class labels are employed as inputs to train Fisher's Quadratic Discriminate Analysis (QDA) for soft, or probability, labels. We term this method ELCM-QDA.

3.1 Constructing Three Physical Features

MISR presents a massive data analysis problem. For each pixel there are 36 dimensions (4 wavelengths at 9 angles). For each data unit there are $196,608 (= 384 \, \text{rows} \times 512 \, \text{columns})$ pixels at the 1.1 km resolution and $3,145,728 (= 1536 \, \text{rows} \times 2048 \, \text{columns})$ pixels at the 275 m resolution. The massive sample size and relatively high dimension pose a serious challenge for even the most simple exploratory data analysis, such as scatter plots and summary statistics (e.g. means, variances, and correlations), over a particular region. In order to efficiently visualize all 57 data units that we use for testing our algorithms, we wrote a Graphic User Interface (GUI) with more than one thousand lines of MATLAB code. With the help of the GUI we investigated the distributions of a large collection of features over all data units. The features investigated include linear combinations of angular radiances, correlations among different angles and wavelengths, nonlinear transformations of radiances, spatial patterns of clouds, and smoothness of reflecting surfaces. Combining exploratory data analysis with specific domain knowledge, such as the fact that ice and snow surfaces scatter radiation more isotropically than clouds which have more forward scatter, we found three physically useful features that separate well surface pixels from cloudy ones.

Only MISR red band data are used for constructing the features for two reasons. First, based on exploratory data analysis and the scattering properties of liquid- and ice-water particles, all four bands have similar reflectance signatures over ice, snow and clouds (Fig. 6; e.g. Clothiaux et al. 2005). Second, only the red band radiances have 275 m spatial resolution, the highest spatial

resolution of MISR. During MISR operational processing, MISR red-band radiation measurements are registered to the reference ellipsoid in two different ways. One is the reference ellipsoid registration illustrated in Figure 2, whereby the radiation measurements are registered directly to the reference ellipsoid ignoring terrain. In the second terrain-projection registration process (Figure 6) the radiation measurements are projected (over land) to the terrain surface and then given the reference ellipsoid coordinates of the terrain surface. In our study we always use the radiation measurements that are projected to the underlying surface (i.e. surface registered data), which are the reference-ellipsoid projected measurements over the oceans and terrain-projected measurements over land surfaces.

Let $I_{m,\ell}^k$ denote the (m, ℓ) pixel of the 275 m resolution red radiation measurements for the kth MISR camera angle, where k runs from 1 to 9 and represents MISR cameras Df to Da.

First feature: CORR

The first feature, named CORR, is revised from Shi et al. (2002) and is the linear correlation of radiation measurements at different view angles. For clear scenes the radiation measured by different MISR cameras is scattered from the same surface (left panel of Figure 6) and the radiation in the different directions increases and decreases together with the magnitude of the source, namely the sun. Consequently, they should be strongly correlated. If a cloud is well above the underlying ocean or land surface, the different MISR camera radiation measurements registered to the same location are from different parts of the cloud (right panel of Figure 6). As a result, the correlation of the measurements from any two MISR cameras will be low. That is, the correlation of radiation measurements in any two MISR camera directions will be high for clear regions and low when clouds well above the surface obscure either one or both of the camera views.

With the above ideas in mind we define the first feature $CORR$ at 1.1 km resolution using the

275 m data. A 1.1 km pixel (i, j) $(i = 1, \ldots, 384, j = 1, \ldots, 512)$ in a data unit corresponds to a 4×4 group of pixels at 275 m resolution. At the 275-m resolution we include two additional rows and columns of pixels from all four sides of the 4×4 square to form a 2.2 km by 2.2 km square which contains 64 275-m resolution pixels. Using these 64 275-m resolution pixels, we denote the linear correlation coefficient between the k_1 and k_2 MISR view directions ($k_1 \neq k_2$ and $1 \leq k_1, k_2 \leq 9$) by

$$LC_{ij}^{k_1, k_2} = \frac{\sum\limits_{m=4i-5}^{4i+2} \sum\limits_{\ell=4j-5}^{4j+2} (I_{m,\ell}^{k_1} - \overline{I_{i,j}^{k_1}})(I_{m,\ell}^{k_2} - \overline{I_{i,j}^{k_2}})}{\sqrt{\sigma_{i,j}^{k_1} \sigma_{i,j}^{k_2}}}, \tag{1}$$

where $I_{m,\ell}^{k}$ is the 275-m resolution radiation measurements at location (m, ℓ), and $\overline{I_{i,j}^{k}}$ and $\sigma_{i,j}^{k}$ are the mean and standard deviation of the 64 radiation measurements for the kth view angle associated with location (i, j) at 1.1-km resolution.

The means are arithmetic averages and the standard deviations are given by

$$\sigma_{i,j}^{k} = \sqrt{\left(\frac{1}{64 - 1}\right) \sum\limits_{m=4i-5}^{4i+2} \sum\limits_{\ell=4j-5}^{4j+2} (I_{m,\ell}^{k} - \overline{I_{i,j}^{k}})^2}. \tag{2}$$

We expect low correlation for radiation measurements from clouds. However, high correlation between two cameras for high altitude cloud measurements could occur under rare circumstances, such as the ratio of along-track cloud speed to satellite speed equaling the ratio of cloud height to satellite height. Therefore, we define the first feature as the average of the correlation coefficients obtained from the MISR Af ($k = 4$)/An ($k = 5$) and Bf ($k = 3$)/An ($k = 5$) camera pairs,

$$CORR = (LC^{4,5} + LC^{3,5})/2,$$

because, in practice, this average is always relatively low for high altitude clouds. High values of $CORR$ suggest either clear (cloud-free) conditions or the presence of low altitude cloud that is registered to the same location on the underlying surface. Extremely smooth terrain surfaces, like

frozen rivers and packed snow, can also produce very low correlations because the variations in the radiation measurements over these smooth surfaces are mainly due to instrument noise. Therefore, if we declare clear for high $CORR$ groups of pixels and cloudy for low $CORR$ groups, we can produce two types of errors: smooth terrain surfaces in the absence of clouds can be classified as cloudy and low altitude clouds can be classified as clear, or cloud-free. To avoid these errors, we need two more features.

Second Feature: SD

To identify smooth surfaces, the standard deviation [Eq. 2] within groups of MISR An ($k = 5$) camera radiation measurements is useful:

$$SD_{\text{An}} = \sigma_{i,j}^5.$$

This feature has small values for radiation emanating from smooth surfaces over which the correlations between different MISR view directions are dominated by instrument noise. Because smooth surfaces in the Arctic are sometime quite limited in space, for example, frozen rivers or valley glacial ice flows, $\sigma_{i,j}^5$ should be computed for small groups of pixels.

Third Feature: NDAI

The third, and last, feature is the Normalized Differential Angular Index ($NDAI$) developed by Nolin et al. (2002) to separate low clouds from surfaces. Surface-leaving radiation at visible wavelengths is more isotropic from ice- and snow-covered surfaces than it is from low altitude clouds (Stephens et al., 1981). This observation motivated Nolin et al. (2002) to use the ratio

$$NDAI_{ij} = \frac{\overline{I_{ij}^1} - \overline{I_{i,j}^5}}{\overline{I_{ij}^1} + \overline{I_{i,j}^5}}, \tag{3}$$

where the average radiation measurements are over 16 275-m resolution measurements in a 1.1 km by 1.1 km group of pixels, to separate clouds from ice- and snow-covered surfaces. In our current

implementation MISR Df-camera measurements are compared with those from the An camera. Note that, in principle, $NDAI$ could also be used to detect high-altitude homogeneous clouds with large spatial coverage (right panel of Figure 6). But clouds with small spatial extents and multi-layer clouds with shadows sometimes display low values of $NDAI$, hence $NDAI$ itself is not sufficient to detect all clouds in the Arctic. In conjunction with the linear correlation ($CORR$) and surface smoothness (SD_{An} features, these types of clouds are easily separable from the surface. We will demonstrate in the results that these three features are sufficient to identify and distinguish most surface and cloud types.

To illustrate the information content in the three features for separating clouds from ice- and snow-covered surfaces consider the locations of expert labelled clear and cloudy pixels in the three-dimensional space spanned by $CORR$, SD_{An} and $NDAI$ (Figure 7). As the figure clearly indicates, expert labelled clear and cloudy points occupy disjoint regions in the space spanned by the features. Moreover, the reversed "L" shape patterns of the clear pixels confirm our observation that $CORR$ decreases as SD_{An} gets small. In addition to the clear separability of clear and cloudy pixels in the feature space, the features possess another desired property for building an efficient cloud detection algorithm: the distributions of features from clear surfaces are about the same over data collected from different geo-locations or at different times. The importance of this stability property of the features in our combined clustering and classification algorithms will be illustrated in Section 3.2.

3.2 Enhanced Linear Correlation Matching (ELCM) Algorithm

For operational processing of MISR data expert labels are not available to train a detector on the current data unit. As a result, detectors have to be trained either on the current data unit without expert labels or from data units collected previously. Moreover, the results of clustering algorithms

do not always reflect a separation between clear and cloudy conditions, especially when data units are fully cloud covered or totally clear (c.f. Figure 4). Here, we propose a method that combines clustering information on the current data unit with the detector learned from previous data units. To borrow a detector developed from previous data units and apply it to current data units, we require that the distributions of previously collected clear and cloudy pixels in the space spanned by the predictors be similar to those for the current data unit. That is, data collected at different times must have approximately the same patterns in the feature space. This stability and the separability described previously are two essential properties for any operational cloud detection algorithms and they are, as we will show, satisfied by our proposed ELCM and ELCM-QDA algorithms.

Unfortunately, the MISR radiation measurements themselves do not have these two required properties. The main differences among radiation measurements collected at different times and locations result from changing solar illumination directions and the magnitude of the solar radiation. For example, over Greenland the radiation measurements collected in June are systematically higher than those collected in September. We might normalize the measurements by the magnitude of the solar radiation, but this would not account for changes in the measurements that result from changing illumination directions of uneven terrain. Thus the distributions of radiation measurements of clear and cloudy pixels in June do not provide a good description of the corresponding pixels in September and a detector trained on measurements from June performs poorly if applied to September data. While the MISR An-, Cf- and Df-camera view directions best separate clear and cloudy pixels in a space spanned by three MISR radiation measurements, their overall performance is nonetheless poor (Figure 8).

Clear and cloudy pixels are not readily separable in a space of radiation measurements and our results indicate that they are not the best features for the cloud detection problem. The three

17

features that we constructed earlier do possess the separability required for building operational cloud detection algorithms and they are stable in time. High values of $CORR$ over ice- and snow-covered surfaces depend only on accurate data registration to the underlying surface and some variability in the terrain surface and both of these properties are not extremely sensitive to changes in magnitude and direction of the solar illumination. Low values of SD_{An} over smooth ice- and snow-covered surfaces depend only on instrument noise, which is relatively stable in time. The $NDAI$ feature, which characterizes the amount of forward to upward scatter, does not depend on the magnitude of the solar radiation because it is normalized. Unfortunately, the distribution of $NDAI$ over a data unit is affected by the angle between the MISR camera view direction and the solar illumination direction. The seasonal change of this angle leads to slightly different distributions with time. We address this issue by proposing to use a clustering method, called Enhanced Matching (EM), fitted to current $NDAI$ values together with distributions of $NDAI$ values learned from previous data units.

Given the nice properties of the space of three features, a natural and computationally efficient way to label pixels operationally is to threshold the features at fixed values for $CORR$ and SD_{An} and adaptively for $NDAI$. We call this approach Enhanced Linear Correlation Matching (ELCM) because two additional features were added to the linear correlation matching classification (LCMC) of Shi et al. (2002). The ELCM algorithm is as follows:

A 1.1 km by 1.1 km pixel is labelled clear when

$\cdot$ $SD_{\text{An}} < threshold_{\text{SD}}$ **OR**

$\cdot$ $CORR > threshold_{\text{CORR}}$ **AND** $NDAI < threshold_{\text{NDAI}}$.

When the above tests fail, the pixel is labelled as cloudy.

The rationale of the ELCM algorithm follows directly from our choice of three features. Ice- and

snow-covered surfaces under clear skies are either very smooth, with extremely small SD_{An}, or they have both high correlations between different view directions and relatively low forward scattering. Clouds are rarely, if ever, extremely smooth, producing low correlations between views and relatively strong forward scattering. After investigating the distributions of expert labelled clear and cloudy pixels from different orbits and comparing results obtained from different thresholds to the expert labels, we concluded that thresholds for $CORR$ and SD_{AN} are stable and robust across all data units and we set them to the empirically determined, and fixed, values of

$$threshold_{\text{CORR}} = 0.75 \quad \text{and} \quad threshold_{\text{SD}} = 2.0.$$

Appropriate thresholds for $NDAI$ varied from data unit to data unit, so this threshold must be learned adaptively.

The choice of threshold$_{NDAI}$

The clustering algorithm extracted from the current data unit and the threshold learned from the previous visit are used to set $threshold_{\text{NDAI}}$. For each data unit in our first orbit we perform a grid search of $threshold_{NDAI}$ from $[0, 1]$ in steps of 1×10^{-5} to identify the value that leads to the smallest classification error relative to the expert labels. This step ensures always having a threshold available from a previous data unit. The clustering algorithm is a one-dimensional mixture model of two Gaussian distributions fitted to the $NDAI$ distribution of the current data unit using the Expectation/Maximization (EM) algorithm (Dempster et al. 1977). The EM algorithm is stopped when the increment of likelihood is less than 1×10^{-7}. The dip between the two Gaussian distributions of the mixture model is found by a numerical grid search between the means of the two fitted Gaussian means. The dip, if it exists, is used as the threshold; the dip usually falls in the empirically determined range from 0.08 to 0.40. Otherwise, the threshold used in the previous (16 day earlier) visit is used. When fitting $NDAI$ distributions, the upper and lower 2.5% tails are

19

trimmed off to avoid outliers and extreme values in both tails, thereby improving robustness of the fitting scheme. For the same robustness considerations, we use the dip in the fitted distribution, instead of the Bayes rule cut-off, because the dip is more robust when the empirical distribution has a heavier tail than a mixture model of two Gaussian distributions.

An example of an $NDAI$ distribution and the mixture model fitted to it is illustrated in Figure 5. These results are based on MISR blocks 20–22 of orbit 13490. For this data unit the dip in the fitted distribution is quite clear, occurring at a value of 0.215 and falling in the expected range from 0.08 to 0.40, and the threshold from the previous visit is not needed. Using $threshold_{\text{CORR}} = 0.75$, $threshold_{\text{SD}} = 2.0$, and $threshold_{\text{NDAI}} = 0.215$ to classify the pixels in the data unit, we obtain a 96.16% agreement rate against the expert labels. Classification accuracies for MISR blocks 20-22 of orbits 13257 and 13723 are 87.53% and 94.36% (Figure 10). Our findings here are comparable to what we find for all 57 data units in the study, as we show in the results.

3.3   Probability Prediction by Training QDA on ELCM

The ELCM algorithm is designed to provide either clear or cloudy detection labels for each pixel with valid radiation measurements. Since the thresholds produced by the algorithm are not perfect, neither are its labels. Moreover, it is not always possible to classify accurately a pixel as clear or cloudy cloudy if it is partly cloudy, as always happens near cloudy boundaries to varying degrees. Therefore, reporting a probability of cloudiness is desirable and more informative than providing only a binary clear versus cloud labeling. The operational SDCM and ASCM products from MISR actually report results in four categories: "high confidence cloud", "low confidence cloud", "low confidence clear", and "high confidence clear". We use Fisher's Quadratic Discriminant Analysis (QDA) trained from ELCM labels to provide an estimate of probability, or confidence, of cloudiness.

If ELCM results show that an entire data unit is almost totally clear or cloudy with 98% of the labels in the same class, we report just the labels. Otherwise, the ELCM labels are taken as input data to train QDA on the three features. For the 57 data units in this study 32 are partly cloudy and 25 are either totally clear or cloudy.

In a two class classification problem, QDA models each class density as a multivariate Gaussian distribution:

$$f_s(x) = \frac{1}{(2\pi)^{p/2}|\Sigma_s|^{1/2}} \; e^{-\frac{1}{2}(x-\mu_s)^T \Sigma_s^{-1}(x-\mu_s)},$$

where $s = 1, 2$ denotes the class label. Let π_s be the prior probability of class s. The a posterior probability for x belonging to class s is then given by

$$P(x \in \text{Class } s | X = x) = \frac{f_s(x)\pi_s}{f_1(x)\pi_1 + f_2(x)\pi_2}.$$

The parameters π_s, μ_s, and Σ_s are estimated by the empirical (i.e. training data) class proportions, means and covariance matrices, and subsequently substituted into the above two equations. The estimated posterior probability $P(x \in \text{Class } s | x)$ serves as the predicted probability of cloudiness for a pixel.

Compared to other methods that produce probability estimates, such as logistic regression, QDA is easy to compute and more robust to errors in the training labels. First, computation of QDA only involves estimating the vector of means and the covariance matrix, which is much more efficient than iterative weighted refitting of logistic regression. Second, but more importantly, QDA is more robust to mislabels in the ELCM results that are training data for QDA. QDA models the joint distribution of $P(\text{label}, X)$ and logistic regression models only the conditional distribution of $P(\text{label}|X)$. In our setup, means and covariances of X are estimated from labels produced by the ELCM thresholding algorithm. These labels contain errors due to the threshold selection. However,

mistakes in labels do not heavily affect the estimates of means and covariances if the mistakes are only a very small proportion of the total training sample. In our application, the thresholds are indeed in the low density area of the feature distribution. As a consequences, thresholds that are slightly off their optimal values do not heavily affect the estimates of means and covariances in QDA. That is, estimates of $P(\text{label}|X)$ are resistant to errors due to slightly non-optimal thresholds.

The logistic regression models $P(\text{label}|X)$, which is fitted from training data labels. The logistic regression keeps faith to labels, even when they are produced from slightly non-optimal thresholds, and hence this model is more strongly affected by mislabels. Simply put, QDA works better for this problem than logistic regression. This phenomenon is also behind the recent surge of interest in semi-supervised learning (Zhu et al. 2003; Zhou et al. 2005), which aims to take into account the distribution of the predictor, or feature, distribution. A robust version of QDA based on Minimum Covariance Determinant estimators (c.f. Rousseeuw 1985; Croux and Haesbroeck 1999) is also tested, but it does not improve the estimates or the labelling results. Thus we use ELCM-QDA to provide probability labels.

ELCM-QDA results from the three data units discussed across Figures 3, 6 and 10 are illustrated in Figure 11. One indication that the ELCM-QDA algorithm is performing reasonably is the frequent occurrence of probabilities of 0.5 at cloud boundaries. We expect this result because boundary pixels are at the transition between clear and cloudy regions. This observation is consistent with our motivation to use probability labels. As we would also expect, the binary expert labels usually fall in-line with the correct high confidence clear and cloudy probability labels. That is, if only the high probability labels are reported, they have very high agreement with the expert labels.

4 RESULTS

To assess the performance of the ELCM and ELCM-QDA algorithms, as well as the MISR operational SDCM and ASCM algorithms, we now compare their results with the expert labels. We present case studies on data units with the lowest agreement rates between the algorithms and expert labels to provide insights into the reasons why the algorithms fail at certain geo-locations. The probability predictions of ELCM-QDA are also evaluated. In addition to comparing the results from all available operational algorithms with expert labels we also treat the expert labels as training data to show the good separability and stability properties of the three MISR features developed in this study.

4.1 Overall Comparisons of the ELCM, ASCM and SDCM Algorithms

The agreement rates of the ELCM and MISR ASCM and SDCM operational algorithms with expert labels are computed for the 5,086,002 pixels (i.e. 71.5% of the total data set) with expert labels. The results (Table 1) show that the ELCM algorithm agreement rate of 91.80% is better than both the MISR ASCM and SDCM algorithms with agreement rates of 83.23% and 80.00%. The coverage of the ELCM algorithm is 100%, while that for the ASCM and SDCM algorithms is 70.12% and 26.64%. A histogram of ELCM algorithm and expert label agreement rates for the 57 individual data units is illustrated in Figure 12. Across all data units the ELCM algorithm has agreement rates with expert labels at or well above 90% for the majority of cases. Several data units have agreement rates as low as 70% and we analyzed these data units in more detail to quantify any systematic biases in the algorithm or the expert labels.

4.2 Case Study of the Lowest Accuracy Data Unit

Figure 13 shows the data unit (orbit 13956, blocks 17–19) for which the ELCM algorithm and expert label agreement rate is only 71%, the lowest agreement rate for any data unit. Disagreements between the two mainly occurs over regions labelled clear by the expert. Upon further investigation of this data unit, we concluded that the expert was not in error in the labelling of these pixels and that the ELCM algorithm was in error. The major reason for errors in the ELCM algorithm results is that $CORR$ is low even though there is no cloud. The low correlation in this case is a result of poor registration of the MISR Af- and Bf-camera radiation measurement to the reference ellipsoid. As illustrated in Figure 6, this could occur for two reasons. First, if a MISR camera pointing direction is not well-characterized and has slight errors, images from this camera will not be accurately registered to the reference ellipsoid. The correlation of this camera's pixel values with those from any other camera will be low. Inaccurate camera models affect both oceanic and land-covered areas. Second, if terrain heights are not accurate in the MISR geo-location information database, terrain-projected radiation measurements will have an additional potential source of error in their registration. For polar regions both types of errors occurred early in the MISR mission and at this time we expect more problems of the second type because of known limitations in the MISR geo-location information database.

After checking ELCM results for this block range of all ten orbits, we concluded that the reduced $CORR$ feature for regions with sharp elevation changes (e.g. the coast line of Greenland) resulted from errors in terrain heights within the MISR database. An unfortunate result is that the ELCM algorithm has systematic errors in regions of rough topography that result from poor registration of MISR images in these regions. Currently, we do not have a solution to this problem because it requires updating the MISR geo-location database which is not planned for at this time.

24

4.3 Probability Prediction by ELCM-QDA

The ELCM-QDA algorithm does not improve overall agreement rates with expert labels relative to the ELCM algorithm (Figure 12). However, the ELCM-QDA algorithm goes beyond ELCM's binary labels of cloud versus clear by providing probability labels (Figure 14). Expert labelled clear and cloudy pixels almost always have probability labels close to 0 (clear) or 1 (cloudy), while pixels unlabelled by the expert are more likely to have intermediate values. That is, the expert was conservative when giving labels, leaving some obviously clear and cloudy pixels unlabelled while avoiding ambiguous pixels. The self-consistency that we find between probability and expert labels provides support for the efficacy of our feature selection and the effectiveness of the ELCM-QDA algorithm.

The MISR operational SDCM and ASCM algorithms report cloud detection results in four categories: high confidence clear, low confidence clear, low confidence cloudy, and high confidence cloudy. For the results in Table 1 only the high confidence MISR results are compared to the expert labels. If we compare both high and low confidence MISR SDCM and ASCM results to expert labels, SDCM and ASCM coverages are increased to 40.24% and 79.48%, but their accuracy rates drop to 72.9% and 76.23%. The ELCM algorithm is extremely fast in its computation and ideas from it might help improve MISR L2TC operational cloud detection in the Arctic.

4.4 Case Study of Separability and Stability of Features

Randomly sampled expert labels are used as training data to demonstrate the advantages of the three features, relative to the radiation measurements, for cloud detection in the Arctic. We perform two sets of experiments for this purpose: one for testing separability and the other stability. First we train QDA and logistic regression (LR, whose results are courtesy of Guilherme Rocha) using half

of the expert labels drawn at random from one of two data units. We then test the two detectors on the remaining half of expert labels for the data unit. Two groups of predictors are tested: the three features and the red radiation measurements from the nine MISR cameras. Results of this case study with the two data units, which are indicative of the results from all data units, make clear the better predictive powers of features as a result of their better separability than the radiation measurements (Table 2).

In the second experiment stability of the predictors is tested by applying QDA and LR trained with expert labels from a previous orbit (i.e. 13257) to a subsequent one (i.e. 13490). For this case study we obtained accuracy rates of 92.6% for QDA and 92.4% for LR using features and 90.3% for QDA and 90.4% for LR using the radiation measurements. The detectors trained on features outperform the ones using the measurements because the 2.3% and 2.0% gains of features over radiation measurements as predictors are significant for the 99,761 pixels in the data unit from orbit 13490. The three features turn out to be more stable than the radiation measurements.

The results also support our method of adaptively fitting $NDAI$ using current data. With the features as predictors the QDA trained on the previous visit to the data unit of orbit 13257 classifies the (current) data unit from orbit 13490 worse, with 92.6% accuracy, than the one trained on the current data unit itself, with 95.9% accuracy. As discussed earlier, degradation of detector performance is mainly due to changes in the $NDAI$ distribution from the first visit to the second, as illustrated in Figure 15. For this case the distributions of $NDAI$ over clouds differ markedly between the two visits, hence adaptively fitting $NDAI$ for the current data unit improves performance.

5 Summary and Potential Scientific Impacts

This paper presents a statistical study that is but the first of many that must be pursued to characterize accurately the properties of clouds over the Arctic, and eventually Antarctica. We have demonstrated that the three physical features $CORR$, SD_{An} and $NDAI$ based on MISR angular radiation measurements contain sufficient information to separate clouds from ice- and snow-covered surfaces. The new ELCM algorithm based on the three features is much more accurate and provides better spatial coverage than the existing MISR operational algorithms for cloud detection in the Arctic. The ELCM algorithm combines classification and clustering frameworks to make it suitable for operational MISR data processing. Computations within the ELCM algorithm are sufficiently fast to handle the massive MISR data. Results from the ElCM algorithm can be used to improve MISR L2TC cloud retrievals or train QDA to provide probability labels for partly cloudy scenes.

Improved daytime cloud masks from MISR, combined with those from the Moderate Resolution Imaging Spectroradiometer (MODIS) also on the Terra satellite, will enable a range of improved cloud and radiation studies in the Arctic. Of particular importance will be improving our understanding of the flow of visible and infrared radiation through the Arctic atmosphere and beginning to tease apart the response of clouds to, and feedbacks upon, changes in Arctic climate. Improved cloud properties for the Arctic will also translate to more accurate global climate model simulations of climate through improved model cloud physics. These studies will eventually enable us to study how changing cloud properties may either enhance or ameliorate any initial changes in the Arctic brought about by increasing concentrations of atmospheric carbon dioxide.

In addition to monitoring and modelling cloud properties in the Arctic, polar surface studies using satellite measurements require reliable cloud detection. One must know in surface studies that radiation making it to the satellite sensor is originating from the surface and not clouds above

it. Monitoring from space of glaciers and sea ice using visible and infrared wavelengths requires, as a first step, identification of clouds in the imagery. Surface studies are an important component of global monitoring of water resources and climate change.

At this time collaboration between statisticians and atmospheric scientists is important for best utilization of the massive data sets currently being collected for weather and climate studies. We have learned that working together provides the most efficient means for combining domain specific knowledge with the statistical knowledge necessary to explore these large data sets in robust ways.

References

[1] Clothiaux, E.E., Barker, H.W. and Korolev,A.V. (2005) *Observing Clouds and Their Optical Properties, In 3D Radiative Transfer in Cloudy Atmospheres*, Ed. A. Marshak and A.B. Davis, Springer, Berlin **Chapter 2**, 93-150.

[2] Croux, C., and Haesboeck, G.(1999). "Influence Function and Efficiency of the Minimum Covariance Determinant Scatter Matrix Estimator", *Journal of Multivariate Analysis*, **71**, 161-190.

[3] Dempster, A., Laird, N. and Rubin, D. (1977). "Maximum likelihood from incomplete data via the EM algorithm" (with discussion) *J. R. Statist. Soc B.* **39**: 1-38.

[4] Di Girolamo, L. and Davies, R. (1994). "A Band-differenced Angular Signature Technique for Cirrus Cloud Detection", *IEEE Trans. Geosci. Remote Sens.* **32**, 890-896.

[5] Diner, D.J., Asner, G.P., Davies, R., Knyazikhin, Y., Muller, J. Nolin, A.W., Pinty, B., Schaaf, C.B., and Stroeve, J. (1999a). "New directions in earth observing scientific applications of mulitangle remote sensing". *Bulletin of American Meteorological Society.* **80**, 2209–2228.

[6] Diner, D.J., Davies, R., Di Girolamo, L., Horvath, A. Moroney, C., Muller, J.-P., Paradise, S.R., Wenkert, D, and Zong, J. (1999b). "MISR level 2 cloud detection and classification algorithm theoretical basis". Jet Propulsion Lab., Pasadena, CA, JPL Tech. Doc. D-11399, Rev. D, 1999.

[7] Diner, D.J., R. Braswell, R. Davies, N. Gobron, Y. Jin, R.A. Kahn, Y. Knyazikhin, N. Loeb, J-P. Muller, A.W. Nolin, B. Pinty, C.B. Schaaf, and J. Stroeve (2004, submitted). "New Directions in Earth Observing: Scientific Applications of Multiangle Remote Sensing II. Illustrative Results from Terra". *Bulletin of American. Meteorological Society.*

[8] Fisher, A. R. (1925) *Statistical Methods for Statistical Workers*, Hafner Publishing Company.

[9] Giorgi, F., and X. Bi (2005) "Updated regional precipitation and temperature changes for the 21st century from ensembles of recent AOGCM simulations". *Geophys. Res. Lett.*, 32, L21715, doi:10.1029/2005GL024288.

[10] Mardia, K. V., Kent, J.T., and Bibby, S. M. (1979) *Multivariate Analysis.* Academic Press.

[11] Nolin, W. A., Fetterer, M.F., and Scambos, A.T. (2002) "Surface Roughness Characterization of Sea Ice and Ice Sheets: Case Studies With MISR Data". *IEEE Trans. Geosci. Remote 1Sensing.* **40**. **July** ,1605–1615.

[12] Rousseeuw, J.P. (1985), "Multivariate estimation with high breakdown point". *Mathematical Statistics and Applications edited by Grossmann, W. et al.*, B. Reidel, Dordrecht, 283–297.

[13] Shi, T., Yu, B., and Braverman, A. (2002). "MISR cloud detection over Ice/Snow using Linear Correlation Matching". *Technical Report 630, Department of Statistics, University of California.*

[14] Stephens, G.L., Campbell, G.G. and Vonder Haar, H.T. (1981). "Earth radiation budgets". *J.Geophys. Res.*, **86**, 97399760.

[15] Zhou, D., Schölkopf, B., and Hofmann, T. (2005). "Semi-supervised learning on directed graphs". *Advances in neural information processing systems.* **17** L. K. Saul, Y. Weiss and L. Bottou (Eds.), Cambridge, MA: MIT Press.

[16] Zhu, X., Ghahramani, Z., and Lafferty, J. (2003). "Semi-supervised learning using Gaussian fields and harmonic functions". *ICML 2003*

Table 1: Agreement rates relative to expert labels and coverages of the ELCM and MISR operational SDCM and ASCM algorithms.

	ELCM	SDCM	ASCM
Agreement with experts	91.80%	80.00%	83.23%
Coverage	100%	26.64%	70.12%

Table 2: Agreement rates relative to randomly chosen expert labels of results from the operational ELCM algorithm and the off-line QDA and LR methods applied to blocks 20–22 of orbit 13257 and 13490. "F" represents methods trained on the three features and "R" methods trained on radiation measurements.

Orbit	ELCM	QDA(F)	QDA(R)	LR(F)	LR(R)
13257	87.5%	91.8%	86.8%	93.2%	85.1%
13490	96.2%	95.9%	95.6%	96.1%	95.2%
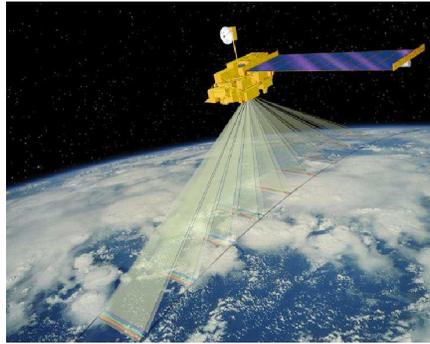
Figure 1: Cartoon illustration of the Terra satellite with the view directions of the nine MISR cameras.
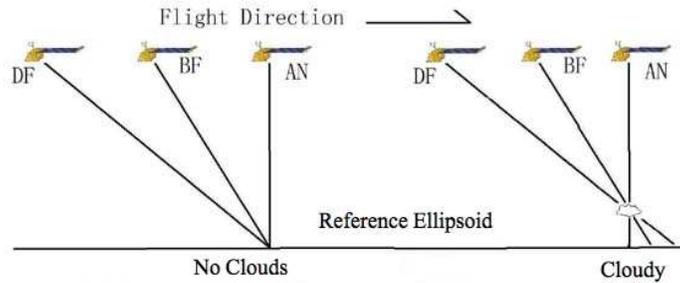
Figure 2: Registration of surface features and clouds to the reference ellipsoid. Note that only three of the nine MISR cameras are illustrated and that surface objects are registered to the same location while clouds are registered to different locations.
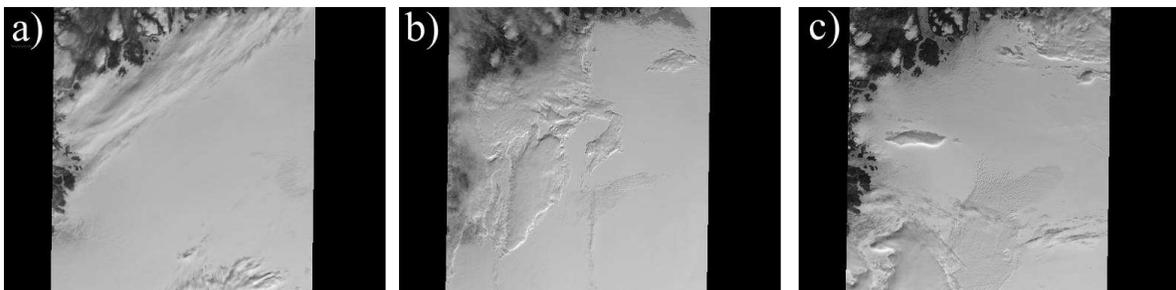
Figure 3: Data collected by the MISR An-camera for three consecutive orbits (i.e. 13257, 13490, and 13723) over blocks 20–22 of path 26.
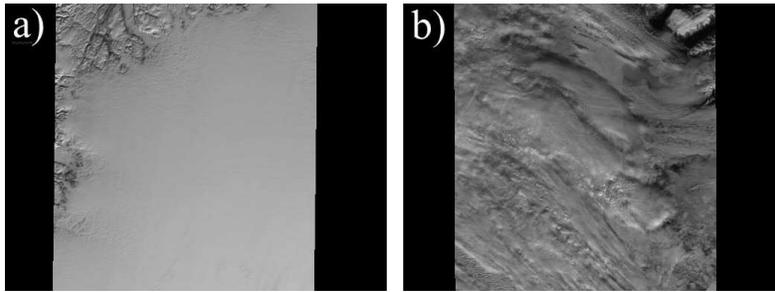
Figure 4: MISR An-camera images of a) a fully clear data unit (blocks 21-23, orbit 12325) and b) a fully cloudy one (blocks 15-17, orbit 13490).
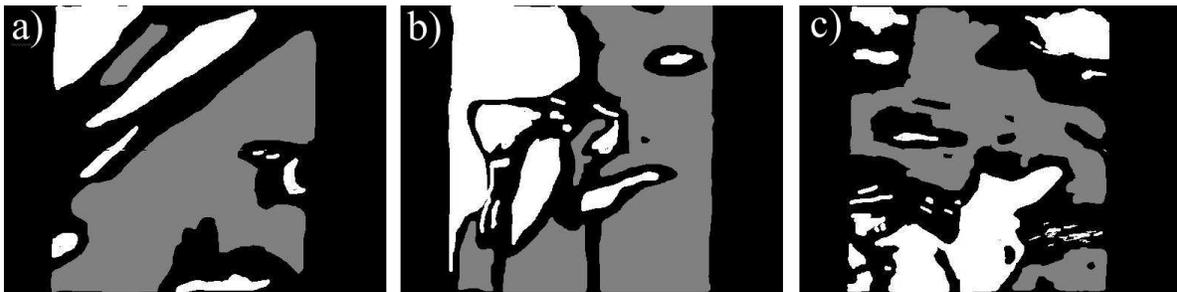
Figure 5: Expert labels for blocks 20-22 of MISR orbits a) 13257, b) 13490, and c) 13723. White represents high confidence cloudy, gray high confidence clear, and black unlabelled pixels.
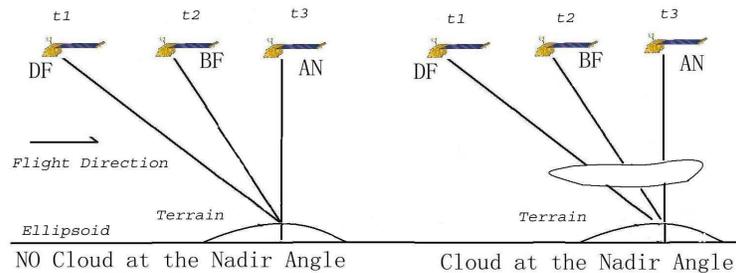
Figure 6: a) Registration of MISR radiation measurements to the terrain when no clouds are present. b) Registration of MISR radiation measurements to the terrain when a cloud is present; note that measurements from different parts of the cloud are now mapped to the same location on the terrain.
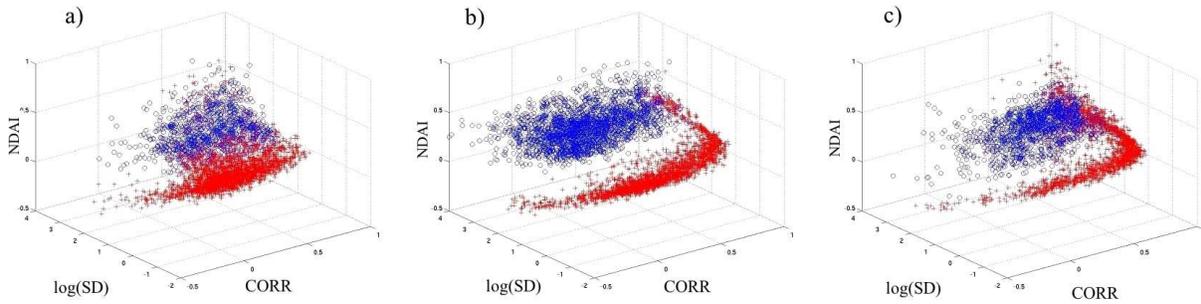
Figure 7: Plots of expert labelled clear (red) and cloudy (blue) pixels in the three-dimensional space spanned by $CORR$, $log(SD_{\mathrm{An}})$ and $NDAI$. Panels a), b) and c) correspond to panels a), b) and c) in Figures 3 and 5.
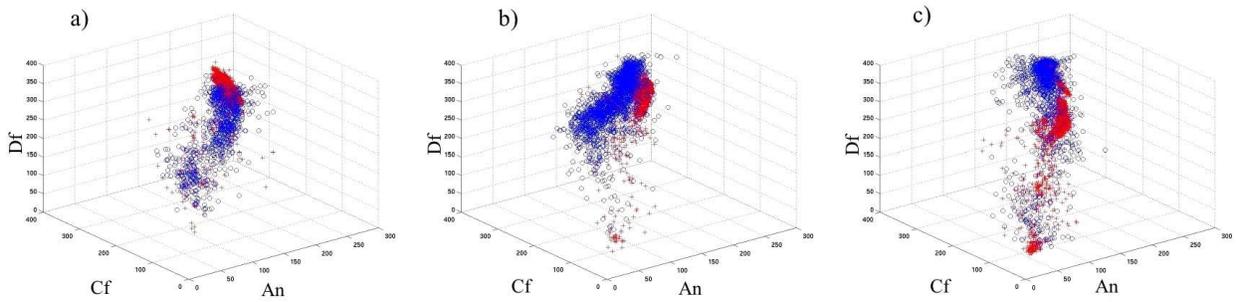
Figure 8: Plots of expert labelled clear (red) and cloudy (blue) pixels in the three-dimensional space spanned by the MISR An-, Cf- and Df-camera radiation measurements. Panels a), b) and c) correspond to panels a), b) and c) in Figure 3.
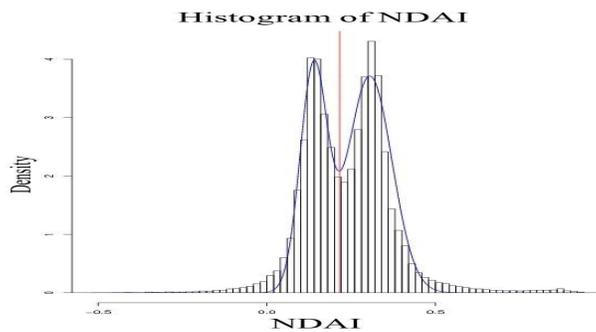

Figure 9: Histogram (vertical bars) of $NDAI$ from blocks 20–22 of orbit 13490 together with the fit (blue line) of the two one-dimensional Gaussian functions mixture model to it. The vertical red line indicates the $NDAI$ threshold derived from this histogram.
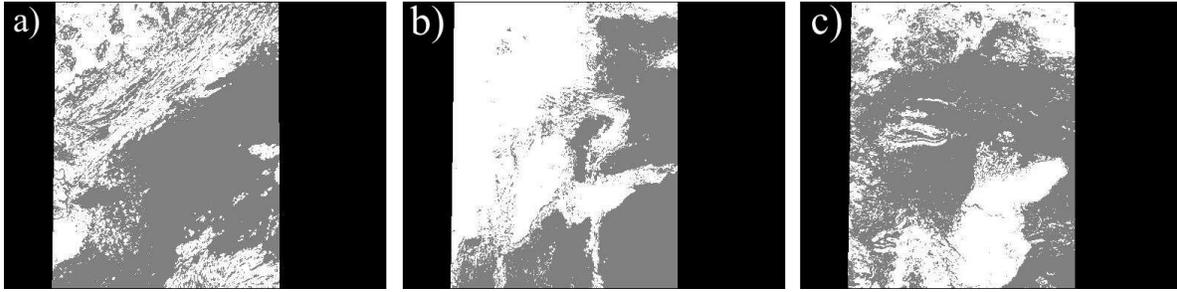
Figure 10: ELCM algorithm results for blocks 20–22 of MISR orbits a) 13257, b) 13490, and c) 13723; these data units are identical to those presented in Figures 3 and 5. White represents pixels classified as cloudy with gray for clear.
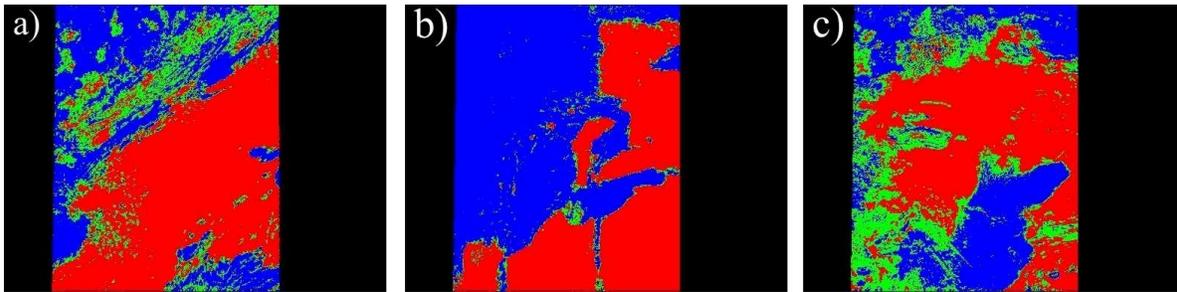

Figure 11: ELCM-QDA results for blocks 20–22 of MISR orbits a) 13257, b) 13490, and c) 13723, which are the same data units illustrated in Figures 3, 5 and 10. Red is for pixels for which $P(\text{cloudy}|x) < 0.2$, green for $0.2 \leq P(\text{cloudy}|x) \leq 0.8$, and blue for $P(\text{cloudy}|x) > 0.8$.
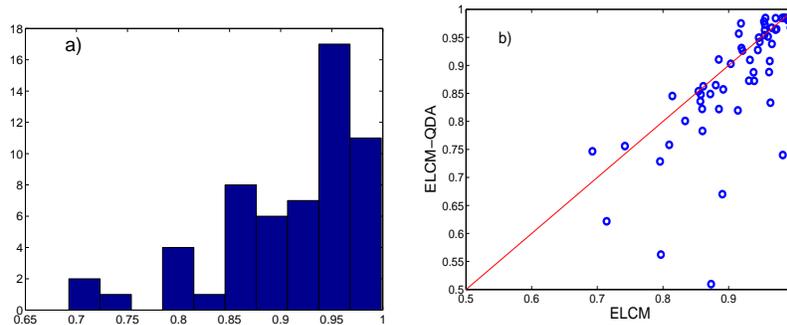

Figure 12: a) Histogram of ELCM algorithm and expert label agreement rates for the 57 individual data units in the study. b) Agreement rate of the ELCM-QDA algorithm with expert labels versus the agreement rate of the ELCM algorithm with expert labels for the 32 partly cloud data units in the study.
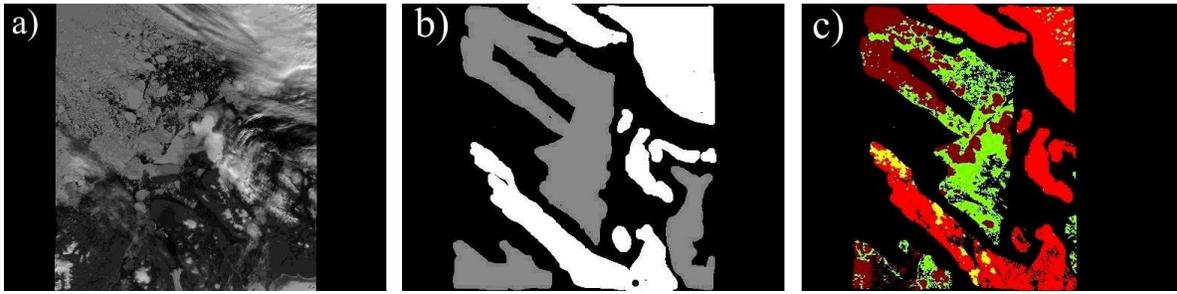
Figure 13: a) MISR An (nadir) camera red band image of blocks 17–19 of orbit 13956; b) Expert labels for the image in a); c) ELCM algorithm and expert label agreement map with red representing agreement on cloudy pixels, brown agreement on clear pixels, green for expert labelled clear but with ELCM classifying the pixel as cloudy, and yellow for expert labelled cloudy but with a clear label by ELCM.
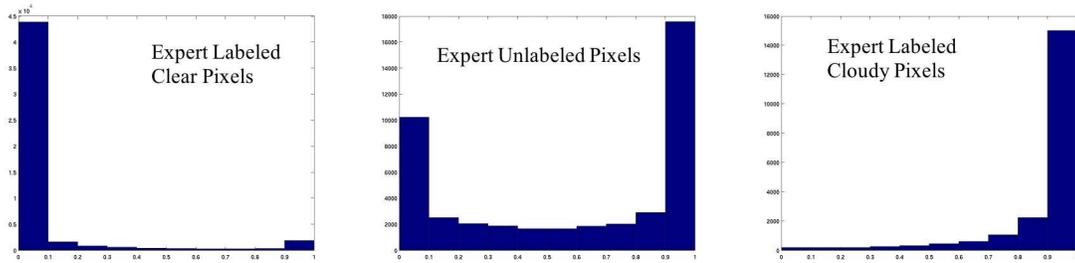

Figure 14: Histograms of predicted probability of cloudiness from the ELCM-QDA algorithm (Orbit 13723 block 20-22) for a) expert labelled clear pixels, b) unlabelled pixels, and c) expert labelled cloudy pixels.
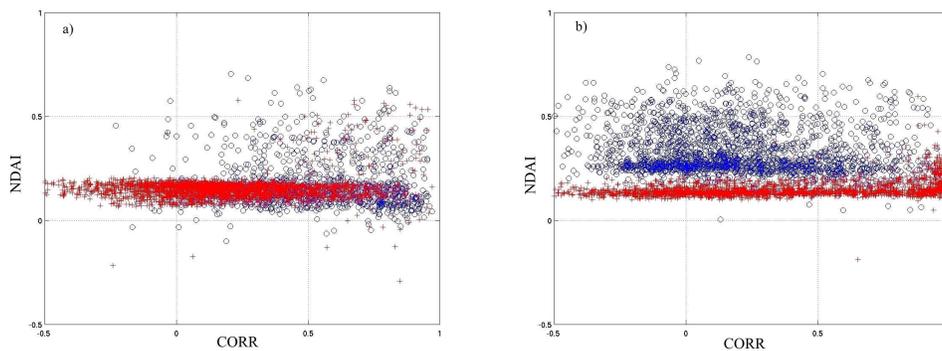

Figure 15: Scatter plots of $NDAI$ vs $CORR$ for MISR blocks 20–22 of orbits a) 13257 and b) 13490. Blue circles represent expert labelled cloudy pixels and red clear pixels.