March 1999

THE ASTER POLAR CLOUD MASK ALGORITHM THEORETICAL BASIS DOCUMENT

By: Ronald M. Welch, Denise Berendes, Todd A. Berendes,
Kwo-Sen Kuo and Antonette M. Logar

Department of Atmospheric Science
University of Alabama in Huntsville
Global Hydrology and Climate Center
Huntsville, AL 35899

THE ASTER POLAR CLOUD MASK ALGORITHM
THEORETICAL BASIS DOCUMENT

Ronald M. Welch, Denise Berendes, Todd A. Berendes and Kwo-Sen Kuo

Department of Atmospheric Science
Global Hydrology and Climate Center
University of Alabama in Huntsville
Huntsville, AL 35899

and

Antonette M. Logar

Department of Mathematics and Computer Science
South Dakota School of Mines and Technology
501 East Saint Joseph Street
Rapid City, South Dakota 57701-3995

TABLE OF CONTENTS

4

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

With the growing awareness and debate over the potential changes associated with global climate change, the polar regions are receiving increased attention. Since greenhouse forcings are expected to be amplified in polar regions (Wetherald and Manabe, 1980; Schlesinger and Mitchell, 1987; Steffen and Lewis, 1988), these regions may act as early warning indicators of actual climate shifts. Global cloud distributions can be expected to be altered by increased greenhouse forcing. In the polar regions, cloud cover changes can be expected to have a significant effect on sea ice conditions (Shine and Crane, 1984) and on regional ice-albedo feedbacks (Barry *et al.*, 1984). In particular, polar stratus is very important to the polar heat balance and directly affects surface melting (Parkinson *et al.*, 1987). However, to monitor changes in polar surface conditions and polar cloudiness, more accurate procedures must be developed to distinguish between cloud and snow-covered surfaces.

Owing to the similarity of cloud and snow-ice spectral signatures in both visible and infra-red wavelengths, it is difficult to distinguish clouds from surface features in the polar regions. In the visible channels, thin ice, ice fragments, wet ice, and pancake ice have low albedos and can be misinterpreted as water, melt ponds, or as thin cloud/haze. Persistent surface inversions and low clouds in winter, and near isothermal structure and extensive stratiform clouds in summer, limit discrimination in the infrared channels (Steffen *et al.*, 1993).

2. BACKGROUND

2.1 Overview

The experience of many investigators has been considered in the classification methodology described in this document. Some of them include: Key and Barry (1989); Li and Leighton (1991); Ebert (1987, 1989, 1992); Key (1990); Allen *et al.* (1989); Saunders and Kriebel (1988); Raschke *et al.* (1992); McGuffie *et al.* (1988); Ormsby and Hall (1991); Sakellariou and Leighton (1988); Crane and Anderson (1984); Simpson and Humphrey (1990); King and Tsay (1993); Menzel and Strabala (1994); Welch *et al.* (1988a,b, 1990, 1992); Tovinkere *et al.* (1993); Rossow (1989); Stowe *et al.* (1991); Rossow *et al.* (1989a,b); Seze and Rossow (1991a,b); Kuo *et al.* (1988).

Typical cloud masking algorithms assume that clouds can be detected using visible and in-frared channel thresholds. Reflectance thresholds typically are set about 3% above the back-ground, and thermal thresholds typically are set about 3°C below the background. Other ap-proaches rely upon bispectral thresholding (Minnis and Harrison, 1984) and a variety of statistical methods (Saunders and Kribel, 1988). However, Rossow *et al.* (1989b), Stowe *et al.* (1989), and many others have reported difficulties associated with polar cloud cover retrievals. Indeed, Land-sat imagery shows that clouds often are darker than the background snow and ice (Welch *et al.*, 1990). In particular, cloud cover often is confused with melt ponds, thin ice, and pancake ice.

Much of the experience in cloud masking described in the literature is based on low resolu-tion non-polar imagery (e.g., AVHRR LAC and GAC). Relatively little work has been performed on high spatial resolution polar imagery. In particular, there is virtually no experience with high spatial resolution polar imagery for the full complement of ASTER channels (especially multispec-tral imagery in the thermal IR region of the spectrum from 8-12 μm).

There are three main factors that must be addressed in the development of an operational polar cloud masking algorithm: (1) proper identification and labeling of sample or test regions, (2) the choice of features, and (3) the choice of the classification methodology.

No classification scheme can be expected to produce accurate results if the labeling of samples or test regions is incorrect. A wide variety of information needs to be examined before labeling a region. A new Interactive Visual Image Classification System (IVICS) has been developed which provides a wide variety of analysis tools to the user. This system has greatly facilitated the selection of pure training samples and accurate labeling. This is particularly important in polar scene analysis where erroneous labeling is problematic. By virtue of using this system, very high classification accuracies ($>95\%$) have been attained for AVHRR data of polar regions (Berendes *et al.*, 1996).

A relatively new intuitive approach has been developed for feature ranking and selection. The distributional overlap between each pair of classes is examined. Those features that most frequently show the least overlap and are least correlated with other features are selected as the best features for the classifier. Several studies have shown that textural signatures can be used in regional AVHRR classification schemes to distinguish between cloud, snow-covered mountains, solid and broken sea ice and floes, and open water (e.g., Welch *et al.*, 1990, 1992; Ebert, 1987, 1989; Key, 1990; Rabindra *et al.*, 1992; Tovinkere *et al.*, 1993). We have indicated in past versions of this document that textural features would be an important element in the feature vector. However, after comprehensive testing, we have found that, in pixel (as opposed to regional) level classification of high spatial resolution Landsat TM data, textures provide little improvement in accuracy. Therefore, because they incur significant computational expense, we have concluded that only spectral signatures are necessary in the feature vector for our classification strategy.

The classifier of choice in this algorithm is a hybrid between some simple thresholding techniques, use of ancillary information, and back propagation neural networks. Through thresholding and use of ancillary information, the class ambiguity of a pixel feature vector is reduced. The ambiguity is then resolved through appropriate selection of one of a number of especially trained neural networks.

The ASTER operational duty cycle is estimated to be approximately 8%. We estimate that 90% of the time the ASTER is on will be during the daytime. During the remaining 10% of the cycle time, it will be on at night. Both a daytime and nighttime algorithm are being developed; however, more developmental emphasis is on the daytime algorithm. Polar regions are defined in this algorithm to consist of all high latitude regions poleward of 60°N and 60°S.

This Algorithm Theoretical Basis Document is organized as follows. Section 2 includes an overview of the problem (this section), the instrument characteristics, a summary of the developmental data sets, the labeling procedure, and a description of the class spectral characteristics. Section 3 contains the algorithm description, including preprocessing, classification, and spatial context consistency tests. Section 3 also presents results obtained from testing of the classifier on Landsat TM data. Finally, Section 4 considers the constraints, limitations, and assumptions used in the algorithm and describes the validation methodology and quality assurance plan.

2.2 ASTER instrument characteristics

ASTER will provide data in three spectral regions using three separate radiometer subsystems. These are the visible and near-infrared (VNIR) subsystem being provided by NEC, the short wavelength infrared (SWIR) subsystem provided by MELCO, and the thermal infrared (TIR) subsystem provided by FUJITSU. The instrument band passes, radiometric accuracies, and radiometric and spatial resolution are given in Table 1. The VNIR includes a separate, single-spectral-band (0.76-0.86 μm, channel 3B) radiometer inclined backward at an angle of 27.7° to the other sensors to provide a 15-m same-orbit stereoscopic imaging capability. A wide dynamic range and multiple gain settings will help ensure useful data for a variety of investigations.

The swath width for all three systems is 60 km. The ASTER instrument has a cross-track pointing capability of 8.55° for the SWIR and TIR subsystems, and 24° for the VNIR subsystem. This gives cross-track observing ranges on the ground of approximately 136 km and 343 km, respectively, ensuring that any point on the globe will be accessible at least once every 16 days for the SWIR and TIR, and once every five days for the VNIR. However, in most instances, all three radiometer systems will image the same 60-km ground swath.

TABLE 1. Spectral and spatial characteristics of the Advanced Spaceborne Thermal Emission Reflectance Radiometer (ASTER). Asterisk indicates the stereo band. Stereo B/H ratio 0.6.

ASTER					
Wavelength Region	Band Number	Spectral Range	Radiometric Accuracy	Radiometric Resolution	Spatial Resolution
	1	0.52-0.60	+/- 4%	≤ 0.5%	15m
VNIR	2	0.63-0.69	+/- 4%	≤ 0.5%	15m
	3	0.76-0.85*	+/- 4%	≤ 0.5%	15m
	4	1.60-1.70	+/- 4%	≤ 0.5%	30m
	5	2.145-2.185	+/- 4%	≤ 1.3%	30m
	6	2.185-2.225	+/- 4%	≤ 1.3%	30m
SWIR	7	2.235-2.285	+/- 4%	≤ 1.3%	30m
	8	2.295-2.365	+/- 4%	≤ 1.0%	30m
	9	2.360-2.430	+/- 4%	≤ 1.3%	30m
	10	8.125-8.475	1-3K	≤ 0.3K	90m
	11	8.475-8.825	1.3K	≤ 0.3K	90m
TIR	12	8.925-9.275	1-3K	≤ 0.3K	90m
	13	10.25-10.95	1-3K	≤ 0.3K	90m
	14	10.95-11.65	1-3K	≤ 0.3K	90m

2.3   Developmental dataset

Landsat TM data currently provides a good match to ASTER data, both spatially and spectrally, and serves as the primary developmental dataset for the daytime version of the algorithm. A total of 82 quarter scenes are currently being used in this development. Twenty-four of them were obtained over coastal Antarctica or over sea ice near Antarctica. The other 58 were obtained over northern latitude regions located in places such as Greenland, Iceland, and Alaska. Labeled samples of contiguous pixel regions have been extracted from all of these scenes and total more than 3700. They represent approximately 200,000 pixels. Results from testing on the labeled samples are presented in section 3.6. The classifier also has been applied to each of the 82 scenes in the inventory. Pertinent information for each of the 82 circumpolar Landsat TM scenes is shown in Table 2. A wide range of conditions are included in this data set, such as: thin cirrus, thin and thick stratocumulus, cumulus, fog, ice-covered land, snow-covered mountains, glaciers, snow-free land, broken sea ice, slush, and melt ponds.

Table 2.   Summary of pertinent information for the Landsat TM daytime developmental data set.

Label	Location	Sat ID	Quad #	Path	Row	Lat (deg)	Lon (deg)	Date Year-Mn-D
l311	Antarctic	4	1	030	115	-77.030	-153.50	1989-03-03
l312	Antarctic	4	2	102	107	-66.944	114.97	1989-03-28
l313	Antarctic	4	4	116	106	-65.579	94.72	1989-03-14
l314	Antarctic	4	2	078	108	-68.265	150.46	1989-02-15
l315	Antarctic	4	2	060	114	-75.875	163.95	1989-02-01
l316	Antarctic	4	2	124	108	-68.267	79.41	1989-02-18
l317	Antarctic	4	4	128	108	-68.280	73.31	1989-03-18
l318	Antarctic	4	1	090	107	-66.933	133.44	1989-02-19
l319	Antarctic	4	2	040	116	-78.110	-173.51	1989-02-21
l320	Antarctic	4	1	030	115	-77.027	-153.52	1989-02-15
l321	Antarctic	4	3	199	112	-73.459	-44.69	1989-01-30
l322	Antarctic	4	2	118	107	-66.940	90.24	1989-03-12
l323	Antarctic	4	1	050	117	-79.101	165.61	1989-01-10
l324	Antarctic	4	2	008	113	-74.690	-112.42	1989-02-05
l325	Antarctic	4	1	207	110	-70.896	-52.44	1989-02-07
l326	Antarctic	4	2	040	116	-78.116	-173.48	1989-03-09
l327	Antarctic	4	2	062	113	-74.696	164.16	1989-03-03
l328	Antarctic	4	2	210	107	-66.945	-51.87	1989-03-16
l329	Antarctic	5	2	070	109	-69.614	161.23	1989-01-14
l330	Antarctic	4	2	008	113	-74.694	-112.39	1989-02-21
l331	Antarctic	4	2	215	106	-65.583	-58.21	1989-03-19
l332	Antarctic	4	1	035	116	-78.113	-165.76	1989-03-06
l333	Antarctic	4	1	215	107	-66.935	-59.67	1989-03-03
l334	Antarctic	4	2	014	118	-80.001	-145.11	1989-02-15
p1_1	Greenland - S Coast	5	1	006	015	64.202	-50.39	1987-07-20
p1_2	Greenland - S Coast	5	2	006	015	64.202	-50.39	1987-07-20
p1_3	Greenland - S Coast	5	3	006	015	64.202	-50.39	1987-07-20
p1_4	Greenland - S Coast	5	4	006	015	64.202	-50.39	1987-07-20

Table 2. (Continued)

Label	Location	Sat ID	Quad #	Path	Row	Lat (deg)	Lon (deg)	Date Year-Mn-D
p2_1	Jones Sound	4	1	043	006	75.869	-89.17	1988-07-01
p2_2	Jones Sound	4	2	043	006	75.869	-89.17	1988-07-01
p2_3	Jones Sound	4	3	043	006	75.869	-89.17	1988-07-01
p2_4	Jones Sound	4	4	043	006	75.869	-89.17	1988-07-01
p3_1	Fairbanks, AK	5	1	069	015	64.238	-147.96	1984-09-22
p3_2	Fairbanks, AK	5	2	069	015	64.238	-147.96	1984-09-22
p3_3	Fairbanks, AK	5	3	069	015	64.238	-147.96	1984-09-22
p3_4	Fairbanks, AK	5	4	069	015	64.238	-147.96	1984-09-22
p4_1	Denali Nat Park, AK	5	1	070	016	62.873	-150.77	1986-06-15
p4_2	Denali Nat Park, AK	5	2	070	016	62.873	-150.77	1986-06-15
p4_3	Denali Nat Park, AK	5	3	070	016	62.873	-150.77	1986-06-15
p4_4	Denali Nat Park, AK	5	4	070	016	62.873	-150.77	1986-06-15
p5_1	Camden Bay, AK	5	1	071	011	69.615	-145.20	1985-06-03
p5_2	Camden Bay, AK	5	2	071	011	69.615	-145.20	1985-06-03
p5_3	Camden Bay, AK	5	3	071	011	69.615	-145.20	1985-06-03
p5_4	Camden Bay, AK	5	4	071	011	69.615	-145.20	1985-06-03
p6_1	Drew Point, AK	5	1	078	010	70.922	-154.16	1985-07-22
p6_2	Drew Point, AK	5	2	078	010	70.922	-154.16	1985-07-22
p6_3	Drew Point, AK	5	3	078	010	70.922	-154.16	1985-07-22
p6_4	Drew Point, AK	5	4	078	010	70.922	-154.16	1985-07-22
p7_1	Deviation Peak, AK	5	1	079	013	66.951	-160.79	1985-07-13
p7_2	Deviation Peak, AK	5	2	079	013	66.951	-160.79	1985-07-13
p7_3	Deviation Peak, AK	5	3	079	013	66.951	-160.79	1985-07-13
p7_4	Deviation Peak, AK	5	4	079	013	66.951	-160.79	1985-07-13
p8_1	Novaya Zemlya	5	1	178	008	73.501	55.84	1986-03-16
p8_2	Novaya Zemlya	5	2	178	008	73.501	55.84	1986-03-16
p8_3	Novaya Zemlya	5	3	178	008	73.501	55.84	1986-03-16
p8_4	Novaya Zemlya	5	4	178	008	73.501	55.84	1986-03-16

Table 2. (Continued)

Label	Location	Sat ID	Quad #	Path	Row	Lat (deg)	Lon (deg)	Date Year-Mn-D
p9_1	Iceland	5	1	217	015	64.236	-16.62	1985-08-05
p9_2	Iceland	5	2	217	015	64.2 36	-16.62	1985-08-05
p9_3	Iceland	5	3	217	015	64.236	-16.62	1985-08-05
p9_4	Iceland	5	4	217	015	64.236	-16.62	1985-08-05
p10_1	Iceland	5	1	219	015	64.239	-19.73	1984-09-17
p10_2	Iceland	5	2	219	015	64.239	-19.73	1984-09-17
p10_3	Iceland	5	3	219	015	64.239	-19.73	1984-09-17
p10_4	Iceland	5	4	219	015	64.239	-19.73	1984-09-1-
p11_1	Greenland - E Coast	5	1	227	012	68.305	-28.07	1986-08-14
p11_2	Greenland - E Coast	5	2	227	012	68.305	-28.07	1986-08-14
p11_3	Greenland - E Coast	5	3	227	012	68.305	-28.07	1986-08-14
p11_4	Greenland - E Coast	5	4	227	012	68.305	-28.07	1986-08-14
wrang_1	Wrangell Mtns, AK	5	1	065	017	61.497	-144.02	1986-09-16
wrang_2	Wrangell Mtns, AK	5	2	065	017	61.497	-144.02	1986-09-16
wrang_3	Wrangell Mtns, AK	5	3	065	017	61.497	-144.02	1986-09-16
wrang_4	Wrangell Mtns, AK	5	4	065	017	61.497	-144.02	1986-09-16
forbin_2	Forbindel Glacier, Grnld	5	2	230	012	68.302	-32.69	1986-09-04
forbin_4	Forbindel Glacier, Grnld	5	4	230	012	68.302	-32.69	1986-09-04
mala_1	Malaspina Glacier, AK	5	1	062	018	60.103	-140.27	1985-08-07
mala_2	Malaspina Glacier, AK	5	2	062	018	60.103	-140.27	1985-08-07
mala_3	Malaspina Glacier, AK	5	3	062	018	60.103	-140.27	1985-08-07
mala_4	Malaspina Glacier, AK	5	4	062	018	60.103	-140.27	1985-08-07
chug_1	Chugach Mtns, AK	5	1	068	017	61.485	-148.58	1985-08-01
chug_2	Chugach Mtns, AK	5	2	068	017	61.485	-148.58	1985-08-01
chug_3	Chugach Mtns, AK	5	3	068	017	61.485	-148.58	1985-08-01
chug_4	Chugach Mtns, AK	5	4	068	017	61.485	-148.58	1985-08-01

The Landsat TM sensor has 7 channels with spectral bandwidths as follows: Channel 1 (0.45-0.52 _ m), Channel 2 (0.52-0.60 _ m), Channel 3 (0.63-0.69 _ m), Channel 4 (0.76-0.90 _ m), Channel 5 (1.55-1.75 _ m), Channel 6 (10.4-12.5 _ m), Channel 7 (2.08-2.35 _ m). These bands are located over visible, near-infrared, and thermal infrared window regions. The spatial resolution of the Landsat TM data is 28.5 m at nadir, except for the thermal IR channel, Band 6, which has a spatial resolution of 120 m. Calibration of pixel brightness to reflectance (Bands 1-5 and 7) and brightness temperature (Band 6) is performed according to Markham and Barker (1986).

Since late in 1997, we have been using the MODIS Airborne Simulator (MAS) data for algorithm development. This instrument has 50 channels. The advantages of the MAS data set over LANDSAT TM are that 1) the MAS channels are more comparable to those of ASTER in both wavelength and spatial resolution, and
2) MAS has three thermal infrared (TIR) channels (channels 42, 44 and 45) that correspond well to ASTER TIR channels (11,13 and 14). MAS data used in this development effort have been taken from the ARMCAS and ALASKA data sets. One problem in using these data is the fact that the classifier appears to find the data sets to be slightly different in calibration. This will be addressed later.

2.4   Pixel labeling

A critical aspect of this algorithm development is the extraction of an accurately-labeled set of pixel subregions. Accurate labeling is the key to accurate classification. Therefore, it is important to provide an analyst performing the sample extraction and labeling with as much information as possible. Figure 1 shows an example of the Interactive Visual Image Classification System (IV-ICS) which displays three-band color overlays. A series of pull-down menus are available to the analyst, which allow a wide range of channel displays and image processing functions. By default, all bands are histogram equalized for contrast enhancement. However, any combination of band differences and band ratios can be designed and displayed on command.
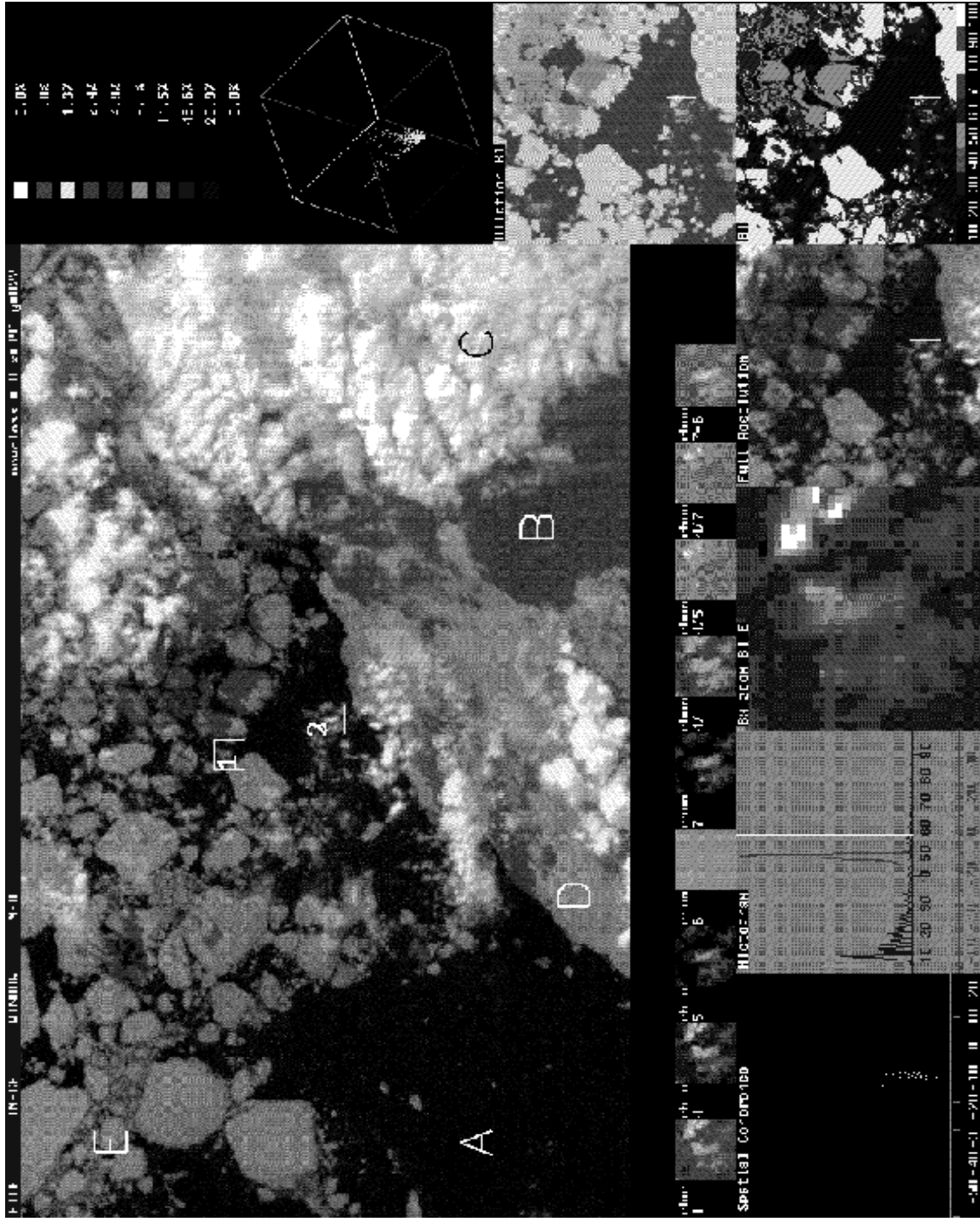
Figure 1. *Example display of Interactive Visual Image Classification System (IVICS) used in pixel labeling.*

The large central image in Figure 1 is a full spatial resolution subsample of the original image. The region labeled "A" is water, "B" is shadow on ice, "C" is stratocumulus cloud, "D" is ice-covered land, and "E" is broken sea ice. Directly under the central image are 10 small regions which display Landsat TM channels 1, 4, 5, 6, 7, 4/1, 4/5, 4/7, and 7-6. The analyst immediately can examine the region outlined in the box in each of these channels. These 10 small regions also are used as icons for mouse control, as explained below. Starting at the lower left corner of the monitor and moving to the right, a series of special purpose displays are provided. First a spatial coherence (2 x 2 pixel plot of mean versus standarddeviation) (Coakley and Bretherton, 1982) is provided. Next are histograms of all three channels where channels 4 and 5 are displayed in terms of albedo, and channel 6 is displayed in terms of temperature (°C). In the lower center is an 8x zoom of channel 1, followed by a three-color enlarged display of the selected regions (shown in black and white here). At the bottom right, a color density sliced version of channel 1 is displayed to the operator; the percentage of pixels in each color range is given at the top right. In addition, to the far right, a morphological dilation (Serra, 1982) is shown for channel 1. The analyst has various options for each of these special displays which are activated by clicking the mouse on any of the 10 icons. Finally, three-dimensional cluster analysis is shown in the center right cube; this cube displays 3-D clusters, computed from the selected region (i.e., box), for the selected red, green, and blue bands. The cluster cube rotates continuously, but can be halted with the mouse button. A trained analyst can see immediately from the three-dimensional cluster display if the boxed region contains pure classes. In Figure 1, the boxed region labeled as "2" in the main display contains both cloud and water, which is clearly differentiated in the 3-D cluster display.

IVICS-2000 is a new visualization and classification tool that has been developed for satellite data. It has been developed using OSF Motif and OpenGL, both industry standards, for each maintenance and later upgrades. It also allows for cross-platform deployment. IVICS-2000 replaces the SIVIS and earlier version of IVICS described above.

SIVIS used the Silicon Graphics Inc., (SGI) GL graphics library as its user interface builder. It has now been replaced by OpenGL, so the maintenance and upgrade of SIVIS became a major difficulty. IVICS-2000 provides most of the original functionality of SIVIS and IVICS with additional visualization options. One such option is the capability to examine classification results produced by the classifier. This allows the user to investigate classification ambiguity derived from incorrect sample labeling. IVICS-2000 greatly enhances our capability to refine the classifier as new results are obtained. IVICS-2000 is currently deployed but with constant upgrades taking place.

2.5   Classes

2.5.1   Daytime   mask

The primary motivation for this algorithm development is the production of a polar cloud mask (i.e., distinguishing cloud from clear); however, a necessary byproduct for classifier training is the distinction among several cloud and clear classes. The following clear and cloud classes were arbitrarily selected for labeling and, subsequently, for classification:

Clear	Cloud
water	thin cloud over ice/snow
slush/wet ice/thin ice	thin cloud over water
ice/snow	thin cloud over land
shadowed ice/snow	thick cloud

land
shadowed land

The major distinction in the clear classes is between land and some phase of water (liquid, frozen, freezing/melting). The major distinction in the cloud classes is between thin and thick clouds. If the cloud is thin, often the underlying surface (ice/snow, water, land) can also be determined. In this context, thin cloud implies one in which the optical properties are affected by the albedo of the underlying surface.

2.5.2   Nighttime   mask

With the adoption of MAS data, the development of the nighttime cloud mask became possible. Daytime MAS images were used for the nighttime algorithm development, since labeling of samples was done by visual inspection. However, only data from thermal channels were used in training the nighttime algorithm. Visual inspection in the visible channels appears to be the only feasible way to assess the performance of the nighttime classifier.

2.6   Class   spectral   characteristics

2.6.1   Water

Water is very dark (i.e., of low reflectance) in all 3 ASTER VNIR bands (i.e., Bands 1-3) and all 6 SWIR bands (i.e., Bands 4-9). In non-polar scenes, the reflectance of water is typically less than 0.06; however, in polar scenes water can appear to have a higher reflectance due to the presence of subpixel resolution ice and slush, resulting in reflectance that can be as high as 0.1. Fortunately, specular reflection is not a problem in ASTER imagery due to the small observation angles. In the ASTER TIR window bands (i.e., Bands 10-14), water is greater than 273 degrees K in non-polar regions. However, for the same reason that the VNIR and SWIR Band reflectance can be atypically bright, the sea water temperature can be as low as 271K and appear to be as low as 268K due to the presence of subpixel scale slush. Water in streams, lakes, and coastal areas, especially near river mouths, is more problematic. The water can be shallow and transparent to the bottom, or contaminated with silt and vegetative matter. This type of water can have the same spectral characteristics as wet or shadowed land.

2.6.2   Land

Land spans a relatively broad range of reflectance in the ASTER VNIR and SWIR bands and can be ambiguous in any of those bands with water, ice, or cloud. Fortunately, most land surfaces (including bare soil and vegetation) are brighter in ASTER Band 3 than Band 1, and the difference increases linearly with increasing reflectance in Band 3. By using ASTER Band 3 and Band 1 together, most land surfaces can be resolved from water, ice, and cloud regions (e.g., Li and Leighton, 1991, used AVHRR Bands 1 and 2 in the same manner). Unfortunately, some land surface types manifest unusually small differences in ASTER Band 3 - Band 1 reflectance, and they are difficult to resolve from cloud regions. Through the use of land cover and ecosystem maps, the classifier potentially can be adapted to these regions. It is also possible that these land regions will be resolvable in the ASTER SWIR Bands 5-9 due to some unique spectral features or ASTER TIR Bands 10 through 14 due to some unique emissivity characteristics. In all the daytime imagery that is currently being used in the development, land surfaces are always above 273K. While this condition is not universally present, especially in winter, this information is valuable for some seasons and geographical locations, and is useful for resolving the aforementioned unusual land surfaces and ice. Wet land surfaces are a classification problem as they are ambiguous spectrally, often with shadowed land surfaces (either caused by topography or by clouds) and sometimes with water. Land surfaces in mountainous regions can be ambiguous with thin cloud over land. Their reflectance is erroneously retrieved due to topography. Topography dramatically im-

pacts the observation geometry and can introduce bidirectional reflectance and lighting effects that are impossible to model without high resolution elevation models. Navigation of the imagery provides the potential for customizing classification in these regions.

2.6.3   *Freezing   or   frozen   water   (slush-ice-snow)*

Ice and snow span a large range of reflectance in the ASTER VNIR bands depending on the temperature and age of the ice or snow. Wet or melting ice, thin ice, and slush can all have reflectance as low as 0.05 in the ASTER VNIR bands and are ambiguous with water, thin cloud, and shadowed or wet land surfaces. Ice packs and snow can have moderate to very high reflectance in the ASTER VNIR bands and are frequently brighter than clouds. In general, it is impossible to resolve clouds from slush-ice-snow in the ASTER VNIR bands. A similar problem is also true for the ASTER TIR Bands 10-14. From a temperature standpoint, slush-ice-snow are everywhere ambiguous with clouds. In polar regions, one cannot assume that clouds are colder than the underlying surface. Frequently occurring inversions in the polar regions result in cloud top temperatures greater than the underlying slush-ice-snow surface. However, slush-ice-snow can be unambiguous with cloud in ASTER TIR Band 10 minus Band 13 and Band 13 minus Band 14. Emissivity differences from ASTER Band 10 to Band 13 are typically larger in clouds than in slush-ice-snow resulting in larger temperature differences for clouds. Atmospheric path differences from ASTER Band 13 to Band 14 are also larger for clouds, again resulting in larger temperature differences. These two temperature differences augment the resolution of clouds from the surface in the daytime algorithm, and form a major basis for resolving clouds from the surface in the nighttime algorithm. In the daytime algorithm, ASTER Band 4 is critical to resolving slush-ice-snow from

clouds. Relative to clouds, solar radiation in the 1.6 to 1.7 μm region (i.e., Band 4) is strongly absorbed by snow and ice; this is due to the larger ice crystal particle sizes and smaller single scattering albedos of snow/ice. Ice cloud absorbs more strongly than does water cloud, but less than ice and snow. In general, using ASTER Band 3 and Band 4 together, slush, ice, and snow are spectrally linear, from water to the brightest fresh snow surface, and other classes are spectrally orthogonal to them; however, the slope of the line is scene dependent.

2.6.4   *Clouds*

Of all the classes, clouds span a larger range of reflectance and temperatures than does any other class. This is due to the fact that clouds can be semi-transparent, and their spectral reflectance characteristics can be similar to that of the underlying surface. Thin clouds in daytime imagery can be resolved using some of the spectral characteristics described in previous paragraphs. Resolving thin clouds in nighttime scenes requires the use of Band differences. Emissivities of thin cloud vary significantly across the ASTER TIR bands. The difference in emissivity is largest for smaller particles, like those found in water and ice clouds. In ASTER Band 10 minus Band 13 vs. Band 13 minus Band 14, pixels from cloudy regions scatter widely away from zero, while pixels from clear regions are concentrated in a narrow region nearer to zero. Opaque clouds in daytime imagery can easily be resolved from surface features by virtue of their uniquely high reflectance in both ASTER Bands 3 and 4. However, opaque clouds can be more difficult to resolve in nighttime scenes because the ASTER TIR band differences (Band 10 - Band 13 and Band 13 - Band 14) are small and are more similar to surface regions. Most surface regions have temperatures greater than 220 K, depending on their location and season. Very cold clouds can be resolved based on temperature alone.

2.6.5   *Glaciers   and   moraines*

Moraines are a collection of gravel, rock, and dirt deposited by glaciers that can be wet or icy. They pose a particular classification problem in that they can be spectrally ambiguous with water, ice, land, or cloud. The ice in glaciers can also be difficult to classify correctly, as it is frequently contaminated with the same material found in moraines. Here navigational and land characterization information are especially important for classification.

3.   ALGORITHM DESCRIPTION

3.1   Overview

The methodology implemented in this algorithm can be characterized as hybrid, hierarchical, or multi-stage. The intent is to link several techniques in such a way that efficiency and speed are optimized while not compromising classification accuracy. A relatively simple technique based on both adaptive and fixed thresholds is used to reduce the pixel class ambiguity, followed by a set of especially trained back propagation neural networks which resolve the remaining ambiguity and classify the pixel. To the maximum extent possible, the classification strategy is derived in knowledge of physical phenomenology.

The algorithm has four stages or levels. In the first stage, the image is navigated, normalized, and sub/super sampled. In the second stage, a preclassification is performed to reduce the class ambiguity of a feature vector, with small computational expense, through the use of a few key features and ancillary information such as coastline data, land/water mask, land character mask, and ecosystem mask. In the third stage, additional features are computed and a set of back propagaton neural networks are used to resolve the class of the feature vector to one of ten possible classes. In the fourth stage, a simple spatial concistency test is performed on the classification mask and pixels that are not spatially consistent with their neighbors are reclassified.

Details concerning the ASTER Polar Cloud Mask also can be found in Logar et al, 1998. Details concerning the ASTER preflight and inflight calibration program and the validation procedure for Level 2 products can be found in Thome et al, 1998.

The algorithm consists of the following elements (Fig. 2):

ALGORITHMFLOW

↓

| STAGE 1: |
| Preprocessing - MedianFilter Sample data to 30 m Normalize Day/Night Navigation Earth Surface Types |

↓

```
┌─────────────────────────────────────┐
│              STAGE 2:               │
│  Preclassification                  │
│  (class ambiguity                   │
│  reduction)                         │
└─────────────────────────────────────┘
                 ↓
┌─────────────────────────────────────┐
│              STAGE 3:               │
│  Back propagation                   │
│  Neural Networks                    │
│  (one of ten classes)               │
└─────────────────────────────────────┘
                 ↓
┌─────────────────────────────────────┐
│              STAGE 4:               │
│  Spatial Consistency Test           │
└─────────────────────────────────────┘
```
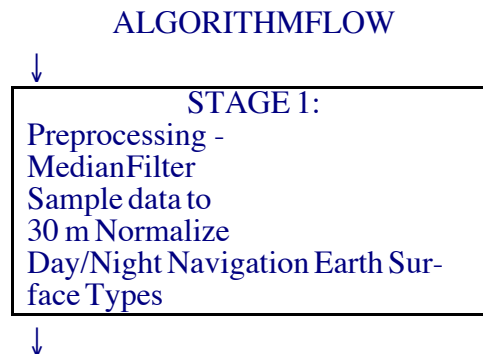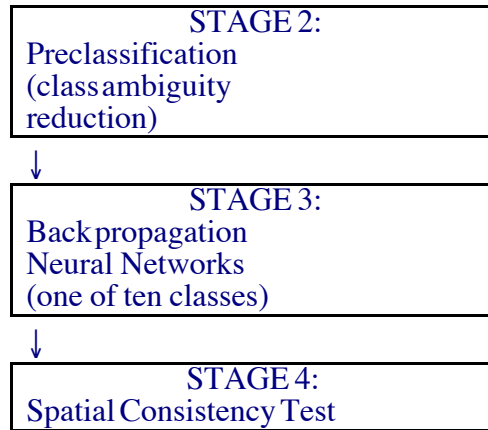
Figure 2: *Algorithm flow chart*

Each pixel is classified using a byte code. The current coding scheme is: (1) water, (2) slush/wet ice, (3) ice/snow, (4) thin cloud over ice/snow, (5) thin cloud over water, (6) thin cloud over land, (7) thick cloud, (8) land, (9) shadow on ice/snow, and (10) shadow on land. The slush/wet-ice class includes melt ponds, slush ice, and thin ice. This class has values of reflectivity intermediate between that of water and ice/snow, and with temperatures near 0°C. If the optical thickness of the clouds are sufficiently thin such that their reflective characteristics are impacted by the albedo of the underlying background, then the classifier attempts to identify the underlying background in classes 5, 6, and 7. In classes 9 and 10 the shadowing can be due either to cloud or topography. The distinctions among classes 4 though 7, between 8 and 10, and among 2, 3, and 9 are more difficult to make than among the more basic classes of water, ice/snow, cloud, and land.

3.2   Stage 1: Preprocessing

First, a 3x3 Median Filter is applied to the data to eliminate "salt and pepper" noise. The ASTER data are obtained at three different spatial resolutions (i.e., VNIR-15 m, SWIR-30 m, TIR-90 m, see Table 1). The classification is derived at 30 m spatial resolution in the daytime algorithm. Therefore, the VNIR data is subsampled, by half, to 30 m pixel spacing. The TIR data is supersampled, by 3, also to 30 m pixel spacing. The SWIR data remains unaltered. In the nighttime algorithm, this step is skipped and the classification is performed at the resolution of the TIR sensor (i.e., 90 m). In the daytime algorithm, the VNIR and SWIR band DNs are normalized for solar irradiance, solar zenith angle, observation angle, and calibration coefficients. The TIR band DNs are converted to temperature. Finally, the scene is navigated to a world data bank system, with coastlines, oceans, lakes, rivers, marshes, reefs, permanent ice regions, deserts, and salt beds noted. The data bank information is used in stage 2 to augment the class ambiguity reduction process.

3.3   Stage 2:   Preclassification (or class ambiguity reduction)

The second stage is designed to reduce the class ambiguity of a pixel to 2 to 4 classes based on adaptive thresholds and the aforementioned ancillary information. The adaptive thresholds are derived from the scene statistics. The class ambiguity of pixels with spectral feature vectors that are located close to class cluster centers can in many cases be reduced to two classes. Pixels with spectral feature vectors that are located between class cluster centers and are ambiguous retain their 10 class ambiguity into Stage 3. Based on results obtained from applying this stage to Landsat TM and MAS imagery, it is estimated that, on the average, the CPU time can be reduced by 50% through the use of this preclassification process.

18

Five features are utilized in this stage. They are B4, B5, B6, ATAN(B4,B5), and B2 • cos(10°) – B4 • sin(10°). (Note the band numbers indicate Landsat TM bands. Landsat TM bands 2, 4, 5, and 6 will be supplanted by ASTER bands 1, 3, 4, and 14 when applied to ASTER imagery; a similar assignment can be made with MAS channels, not shown). A sequence of tests are performed based on these 5 features. The tests utilizing ATAN(B4,B5) and B2 • cos(10°) – B4 • sin(10°) are determined adaptively for each scene. They are derived from the histogram of the feature for the scene being classified and are key to the algorithm's performance. Significant spectral variability is apparent in our 82-scene Landsat TM and MAS circumpolar datasets. Some of the variability is due to the natural variations in the atmospheric path and some is due to the natural variability of a specific class. For example, in the land class, the spectral characteristics of boreal forests are not the same as wetlands or bogs and are not the same as bare rock in treacherous mountain ranges. In the ice/snow class, the spectral characteristics of fresh snow are different from those of ice floes or glaciers. By using adaptive thresholds it is possible to compensate for some of this natural variability occurring in each class type.

The feature ATAN(B4,B5) is especially important in separating all types of frozen water surfaces from cloud over those surfaces. The distribution of all types of frozen water (slush, wet ice, pancake ice, snow, etc.) are distributed narrowly in the ATAN(B4,B5) feature while cloud is distributed much more broadly and uniformly. An example of this characteristic is shown in Figure 3a for one scene. The optimum threshold for separating ice from cloud is indicated in the figure. Another example is shown in Figure 3b for another scene. Again the optimum threshold is indicated but note that its location is significantly different from that in Figure 3a. If the threshold from Figure 3a were applied to the scene corresponding to Figure 3b, significant numbers of misclassification errors occur. The converse is also true when applying the threshold of Figure 3b to the scene corresponding to Figure 3a.

The same methodology is applied to B2 • cos(10°) – B4 • sin(10°); however, in this case the adaptive thresholding is important to distinguishing land from cloud over land. The importance of adaptive thresholding is demonstrated in Figure 4a and b in which histograms of B2 • cos(10°) – B4 • sin(10°) are shown for two scenes. Again, if the threshold for one scene is applied to the other, significant numbers of pixel misclassification occur. The 10° rotation is performed to adapt to the soil line in this feature space. This angle varies from scene to scene but 10° was chosen as a nominal value based on observations of our Landsat TM data set.
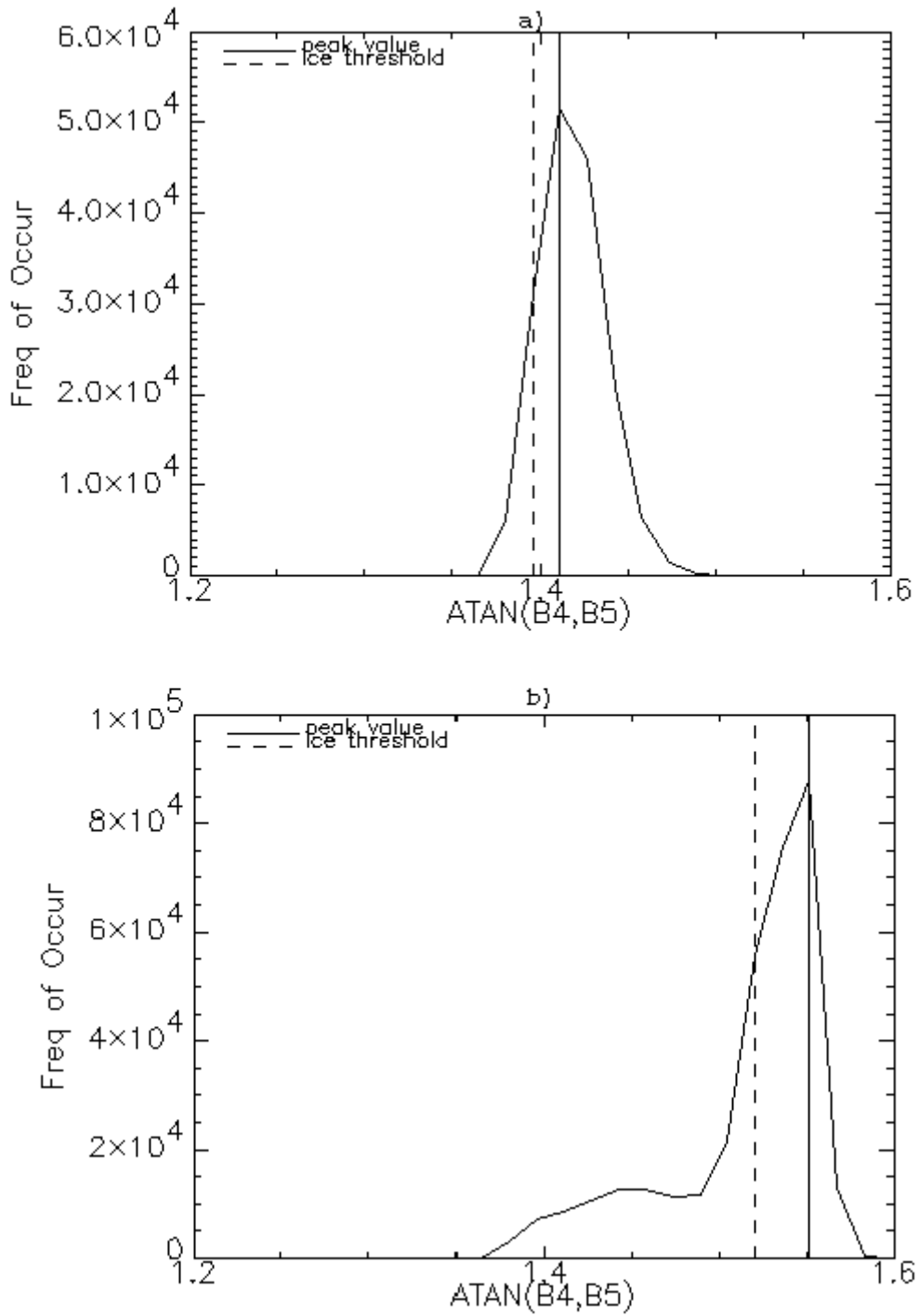
Figure 3. *Histogram of arctan (4,5) for one Landsat TM Antarctic scene with (a) a relatively low threshold, and (b) a relatively high threshold.*
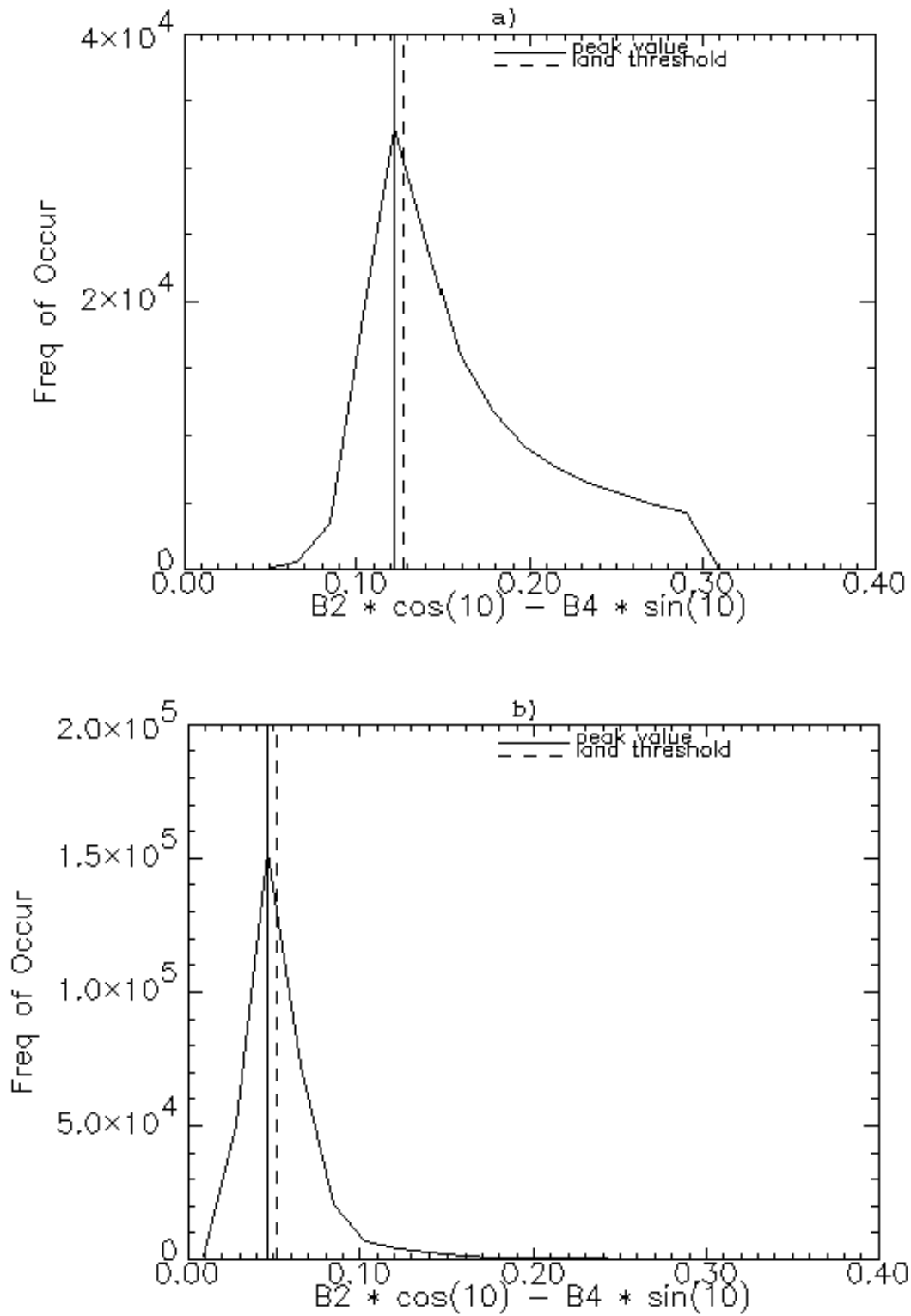
Figure 4. *Histogram of B2•cos (10°) - B4•sin (10°) for one Landsat TM northern hemisphere polar region scene with (a) a relatively high threshold, and (b) a relatively low threshold.*

21

The remainder of the tests, using the aforementioned 5 features, rely on fixed thresholds and were derived from analysis of the histograms of those features for approximately 200,000 labeled pixels. Most of the tests are conditional on previous tests such that a pixel is only classified if it has not been already. The following eight combinations of classes are possible as output from this stage and are provided as input to the next stage (i.e., the set of back propagation neural networks):

> land, shadow on land
> one of the four cloud classes
> water, slush/wet ice
> slush/wet ice, ice/snow, shadow on ice/snow
> slush/wet ice, ice/snow
> thin cloud over water or land, thick cloud
> thin cloud over water or land, land, shadow on land
> any class (no class ambiguity reduction)

3.4   Stage 3: Back propagation neural networks

3.4.1   Overview

This stage applies a more sophisticated and, consequently, a more computationally expensive technique than in the previous stage (i.e., Stage 2). Here more spectral features are computed. They include the basic 14 ASTER bands, but also include derivative spectral features such as normalized differences, arctangents, and ratios. The goal in this stage is to be able to classify at the rate of one million pixels (approximately one quarter of an ASTER scene) in less than 5 minutes using a modest workstation (e.g., SGI Indy). The key to this stage is a comprehensive set of representative samples for all classes. Approximately 3700 samples (i.e., about 200,000 pixels) have been extracted from 82 Landsat TM quarter scenes   (likewise, a large number of samples have been extracted from the ARMCAS and ALASKA MAS data sets). These samples provide the basis for training the neural networks. Currently, the networks are being trained on Landsat TM and MAS data.

3.4.2   Features

Potentially, a total of 140 features are computed for each pixel to be used by one of the back propagation neural networks. They are as follows (Note that the band numbers refer to Landsat TM.) The appropriate ASTER bands will be supplanted when computing these features for ASTER data):

a) The seven reflectance/temperature values for Bands 1-7;
b) The 21 unique band ratios for all paired combinations of seven bands (Note: Unique implies that, for example, Band 1 / Band 2 is computed but not Band 2 / Band 1);
c) The 21 unique band differences for all paired combinations of seven bands (e.g., Band 1 - Band 2);
d) The 21 unique arctangents for all paired combinations of seven bands (e.g., arctan (Band 1, Band 2));
e) The 21 2-D Euclidean distances for all paired combinations of seven bands (e.g., $\sqrt{(band1)^2 + (band2)^2}$);
f) The 3-D Euclidean distances for the following three-way combinations: [1,4,5], [1,4,6], [1,5,6], [4,5,6] (e.g., $\sqrt{(band1)^2 + (band4)^2 + (band5)^2}$);
g) The 21 unique normalized differences for all combinations of seven bands (e.g., (Band 1 - Band 2) / (Band 1 + Band 2));

h) The sets of hue, saturation, and intensity (HSI) for the following eight three-way combinations: [1,4,5], [1,4,6], [1,5,6], [4,5,6], [1,1-4,6], [1-4,4-5,6], [1-4,4-5,5-6], [4-5,5-6,1-5]. Often it is useful to cluster points within a 3D subspace that can be represented within the RGB color cube in which each of the three channels is assigned to one of the three colors. This model, however, does not directly relate to the notions of hue, saturation and intensity. HSI space can be related to RGB space by defining the Intensity axis along the diagonal from (0,0,0) to (1,1,1) in the RGB color cube. Saturation then is the distance from the Intensity axis and ranges in value from 0 to 1. Hue is the angle around the Intensity axis with respect to some reference and ranges in value from 0 to 2π. The RGB to HSI transformation equations are given in Ballard and Brown (1982). The HSI metrics provide the best separability among all 140 features for some class pairs.

3.4.3 Feature selection

Training a neural network with feature vectors containing 140 elements is computationally expensive, extremely time consuming, and generally impractical. Therefore, it is necessary to reduce the input vector size and make it as small as possible. This reduction of features for input into the networks is achieved through the use of two metrics of distributional separation. In the first metric, histograms for each feature and pair of classes are constructed. The histogram ranges are scaled to the minimum and maximum feature values of both classes and discretized into 256 bins. The histogram contains the frequency of occurrence for each class, normalized by the number of samples in the class. From each pair of histograms Overlap, O, is computed as follows:

$$O(f)_{ij} = \sum_{x=1}^{256} I_f(x) J_f(x)$$

where f is a specific feature, i and j are the 2 classes under comparison, I and J are the corresponding histograms for classes i and j, and x is the histogram bin number.

An additional metric, divergence (Richards 1993), is also computed for each feature and class pair. It is defined as:

$$DIV(F)_{ij} = \left| m_i - m_j \right| / (s_i + s_j)$$

where m_i and m_j are the means for classes i and j, and s_i and s_j are the standard deviations for those classes. The features with the lowest overlap (zero representing no overlap) and highest divergence are rated as the top separating features for a given class pair. For a selected set of classes, the features are then ranked by the number of times they are rated as a top feature in a pairwise comparison of all classes. The top ranked features then are selected as input to a given neural network.

3.4.4 Weight projection in neural network training

The neural networks use a method of weight projection which produce a decrease in training time of approximately 50% (Logar *et al.*, 1992, Logar *et al.*, 1993). During training, a sequence of weight values are maintained for each weight. A quadratic-least-squares curve is then calculated for each weight and a prediction is made for a future value based on the curve. With the predicted weights in hand, an error is calculated. If the error has decreased, the new weights are retained. Otherwise, the old weights are retained and training resumes.

3.4.5 Back propagation neural network

The multilayer perceptron network, trained using the back propagation algorithm, used in this development, is described in many sources such as Rumelhart *et al.* (1986), and, more recently, in an excellent review article by Paola and Schowengerdt (1995). The network is composed of three layers of perceptrons: an input layer, a single hidden layer, and an output layer. Each perceptron receives weighted inputs and computes a sum, then generates an output using a non-linear and continuous function. The sigmoid function:

$$f(x) = \frac{1}{1 + e^{-x}}$$

is used in this application.

During training the error is computed as the aggregate squared difference between the desired output and the output actually produced. The weights are then adjusted using a gradient descent to minimize the error. The network is trained until the error reaches a minimum. This procedure does not guarantee a global minimum but most often finds a local minimum which is a good solution.

Currently, several sets of back propagation neural networks have been trained — one each for the southern and northern hemisphere polar region Landsat TM data sets – and one each for the ALASKA and ARMCAS MAS data sets. To determine the best features to use for training each of these sets of networks, the feature selection technique described in section 3.4.3 is applied. Eighty randomly-generated data sets were derived from the total set of approximately 200,000 labeled pixel samples (40 data sets each from the southern and northern hemisphere polar region data). Of the 40 data sets for each latitude extreme, 10 sets each representing 10%, 33%, 50% and 66% random selections from the total available data set were selected. After deriving the overlap and divergence metrics for each class pair, for each feature, the features were ranked based on the number of times a feature is top rated as explained in section 3.4.3. All features that are top rated less than 20 times are eliminated leaving 39 and 25 selected features that were used in training the northern and southern hemisphere polar region classifiers, respectively. Experience indicates that using additional lower-ranked features does not improve performance

Once the input vector composition is determined, each network topology can be determined. The number of elements in the feature vector determines how many nodes are in the input layer, (for the current Landsat TM data set, 39 for northern latitude and 25 for southern latitude). The hidden layer contains

$$\frac{\# \text{ of input nodes}}{2}$$
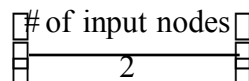
nodes. The number of output nodes in each network corresponds to the number of classes from Stage 2. The training rate is held constant at 0.1 and the momentum rate at 0.5 for all experiments. Experience has shown that these values produce good results for a wide variety of data sets. Training is terminated when the change in error is less than a predefined threshold, approximately 1% of the previous error, for 50 consecutive iterations, or when the error begins to increase.

3.5   Stage 4:   Spatial consistency test

One final test is performed on the entire classification mask from Stage 3. Spatial consistency is tested by examining the eight nearest neighbors of each pixel. If six or more of the nearest neighbors are of the same class, then the pixel is reclassified to the class of the most frequently occurring near neighbor.

3.6   Testing of the classifier using labeled samples derived from Landsat TM and MAS data - confusion matrices

The labeled samples extracted from the Landsat TM circumpolar developmental data set and the MAS data sets each were split into two sets corresponding to a fundamental geographic separation; that is, southern hemisphere polar region and northern hemisphere polar region data. The southern latitude data was obtained over coastal Antarctica and, consequently, only 7 of the 10 classes are present in these scenes (water, wet ice, ice/snow, thin cloud over ice/snow, thin cloud over water, thick cloud, and shadow on ice/snow). In this region of the world, bare land is rarely seen and is not present in any of the imagery; therefore, no samples for land, shadow on land, or thin cloud over land are available. The northern latitude data does contain samples for all 10 classes. Due to the presence of land classes in the northern latitude data, the classifier performs better when trained on each set individually. Note that another factor possibly contributing to a difference in the two data sets is that the southern latitude data was obtained from Landsat 4 while most of the northern latitude data was obtained from Landsat 5. The two regions also differ physiographically and climatologically. For example, much of the southern polar region is covered by a snow- and ice-laden continent surrounded by sea ice and open ocean, while the central area of the northern polar region is ice-covered ocean which is surrounded by land masses which are frequently ice- and snow-free during the summer months.

Note that there were inconsistencies found with the ALASKA and ARMCAS MAS data sets. There appears to be a calibration inconsistency between these data sets to which the classifier is sensitive. Additional testing is necessary to identify the basis for these inconsistencies. While not major in scope, these variations lead to decreases in classification accuracy of about 3%. That is, a classifier based upon the ALASKA data set produces accuracies about 3% lower than does the classifier based upon the ARMCAS data set, and vice versa.

Confusion matrices were generated which show a comparison of the classification results with the actual or known classes of the labeled samples. The elements of the confusion matrix are normalized to percent of the total number of test samples for the class tested. The results are presented in three confusion matrices for each of the northern and southern latitude data sets (Tables 3 and 4, respectively). The first matrix shows the accuracy of the classifier in its most important role - as a cloud mask. It is a 2 by 2 matrix in which all the classification results for the clear classes and the cloud classes have been accumulated together. For example, the value in the first row and column show the percent of samples from cloud classes classified as one of the cloud classes while the value in the second row and column show the percent of samples from clear classes classified as one of the clear classes. The off-diagonal elements, row one - column two and row two - column one, show the errors; that is, they show the percent of samples from cloud classes classified as one of the clear classes and the percent of samples from clear classes classified as one of the cloud classes, respectively. The second matrix shows the classification results for those samples from clear classes correctly classified as one of the clear classes while the third matrix shows the classification results for those samples from cloud classes correctly classified as one of the cloud classes. These two matrices depict the accuracy of the classifier in its secondary role in distinguishing between clear classes or cloud classes. The diagonal elements in these two matrices indicate the classification accuracies for each class while the off-diagonal elements indicate the percent inaccuracies or "confusion" of the classifier. The accuracy over all classes for each of these matrices is shown below the matrix (i.e., the percent of correctly classified samples). In the case of the second matrix for the cloud classes, the main distinction is thin or thick cloud and, if the sample is from a thin cloud, the distinction is the underlying surface (water, land, ice/snow). In the case of the third matrix for the clear classes, the main distinction is between land and some phase of water (liquid, frozen, melting) and shadowed or unshadowed.

Confusion matrices, 3 each for the northern and southern hemisphere polar region data, are shown in Tables 3 and 4, respectively.

Table 3.   Confusion matrices for the northern hemisphere polar region Landsat TM daytime samples.

Cloud vs. Clear				Cloud Classes			
	Cld	Clr		4	5	6	7
Cld	97.3	2.5	4	93.0	0.3	4.5	2.3
Clr	2.7	97.5	5	2.6	94.2	8.0	0.0
Unk	0.0	0.0	6	2.8	5.5	86.9	0.4
Total:	97.4		7	1.6	0.0	0.6	97.2
			Unk	0.0	0.0	0.0	0.0
			Total:	92.1			

Clear Classes						
	1	2	3	8	9	10
1	96.3	7.4	0.0	0.4	7.1	9.3
2	2.4	82.3	0.5	0.0	5.5	0.8
3	0.1	3.4	96.9	0.8	5.3	0.9
8	0.6	0.4	0.9	95.7	15.4	27.6
9	0.3	6.4	1.7	0.1	63.9	6.3
10	0.5	0.2	0.0	3.0	2.9	55.1
Unk	0.0	0.0	0.0	0.0	0.0	0.0
Total:	90.4					

Table 4.   Confusion matrices for the southern hemisphere polar region Landsat TM daytime samples.

Cloud vs. Clear	Cld	Clr
Cld	94.3	2.7
Clr	5.7	97.3
Unk	0.0	0.0
Total:	95.9	

Cloud Classes	4	5	7
4	83.8	9.2	12.8
5	8.1	83.7	21.1
7	0.4	0.9	65.1
Unk	0.0	0.0	0.0
Total	84.5		

Clear Classes	1	2	3	8
1	92.3	3.6	0.1	0.8
2	7.3	87.9	6.1	4.9
3	0.0	5.3	91.2	2.9
8	0.4	3.3	2.7	91.4
Unk	0.0	0.0	0.0	0.0
Total:	90.3			

These results indicate that the algorithm is approximately 95 percent accurate in distinguishing between clear and cloudy pixels. They also indicate that it is approximately 90 percent accurate in distinguishing among clear classes and 85-90% accurate in distinguishing among cloud classes.

3.7   Application of the ASTER Polar Cloud Mask algorithm to Landsat TM quarter scenes

The above results suggest that the algorithm should work well as a Landsat TM circumpolar scene classifier (for the classes selected). However, these results are based on tests from labeled samples. These labeled samples necessarily were selected as unambiguous and spectrally homogeneous representatives of specific classes. The imagery, overall, is spectrally more diverse than the samples are. A confusion matrix generated from classification results of labeled samples may indicate, for example, that only 3 percent of the thin cloud over ice/snow samples are misclassified as ice/snow; however, in a given scene, if more than 3 percent of the scene is comprised of pixels with feature vectors of thin cloud over ice/snow that manifest this type of misclassification, then visual examination of a color-coded classification mask would lead one to state that the classifier is less than 97 percent accurate. The confusion matrices indicate an upper estimate on the accuracy of the classifier, on the average. An important part of the process of evaluating the performance of a classifier, when applied at the pixel level to an entire scene, is visual examination of the scene classification results. Unfortunately this process is subjective and it is difficult to arrive at a quantitative estimate of the performance. Our subjective evaluation of the masks indicates that the classifier is performing the cloud masking function at greater than the 90% level. More details of the cloud mask can be found in Logar et al, 1998, *The ASTER Polar Cloud Mask*, IEEE Transactions on Geoscience and Remote Sensing, p1302-1312.

Researchers often question whether the more advanced techniques, such as Neural Networks, indeed possess advantages over the conventional methods in classifying cloud regions. In 1998 we compared three classification schemes,
   1) the conventional maximum likelihood,
   2) the innovative paired histogram, and
   3) the back-propagation neural network,
in search for the best classifier in terms of accuracy and computational efficiency (Berendes et al, 1998). The NN emerged from the comparison as the best choice, achieving 98.69% accuracy on daytime channels in distinguishing cloud versus non-cloud pixels in MAS data sets. An NN proved to be more computationally efficient as well, sometimes by a factor of two, than other methods. The nighttime NN classifier also achieved the highest accuracy, 86.95%.

4.   CONSTRAINTS, LIMITATIONS, AND ASSUMPTIONS

4.1   Validation

4.1.1   Introduction

4.1.1.1   Measurement and science objectives. The overriding objective of the ASTER Polar Cloud Mask product is to identify or classify all pixels in imagery obtained poleward of 60N and 60S as cloud or clear. Depending on the type of user, the product can be used to mask out all cloudy pixels for surface studies (e.g., ice process studies) or, conversely, all clear pixels for polar cloud studies. The product will be available on request for both daytime and nighttime imagery; however, a different scheme will be used in each case. The daytime algorithm will utilize visible, shortwave IR, and thermal IR for the feature space. In the nighttime algorithm, only the 5

ASTER thermal IR channels can be used. In the case of daytime imagery (estimated to be 90 percent of the total data obtained from ASTER), as a secondary objective, the underlying surface of thin cloud will be identified as either water, land, snow/ice, or unknown. Also pixels identified as clear will be further classified into one of the 6 subclasses of water, wet/thin ice, ice/snow, shadow on ice/snow, land, and shadow on land.

4.1.1.2 Mission. ASTER will fly on the EOS AM-1 platform which is planned for launch in September 1999. The ASTER instrument is described in the EOS Reference Handbook and a brief description of it is provided in 2.2 of this document.

4.1.1.3 Science data products. The ASTER Polar Cloud Mask constitutes one, level 2 data product which is a coded pixel map. It will be accompanied by metadata containing statistics for percent occurrence of each class and cloud/clear fraction, in addition to any "pass through" information derived from input streams for radiance/reflectance/temperature. Pass through information will include general quality assurance information such as the presence and location of bad pixels.

4.1.2   Validation   criterion

4.1.2.1 Overall approach. Currently Landsat TM and MAS data are being used as a surrogates for ASTER in the testing and validation of the daytime algorithm. The instruments have similar spatial resolution characteristics (i.e., Landsat TM bands 1-5, and 7 at 28.5m and band 6 at 120m; ASTER Bands 1-3 at 15m, bands 4-9 at 30m, and bands 10-14 at 90m). The spectral resolution capabilities of some of the bands are very similar, but some are significantly different. For example, bands 2, 3, 4, and 5 of Landsat TM are a good match for ASTER bands 1, 2, 3, and 4. However, band 7 in Landsat TM is a broad shortwave IR band covering the 2.1 to 2.4 µm region, while ASTER has 5 discrete 50 nm (approximately) wide bands centered between 2.1 and 2.4 µm. In addition, Landsat TM has only one thermal IR band covering the 8-12 µm range while ASTER has 5 contiguous, higher spectral resolution bands within the same region.

As mentioned previously, MAS data is being used for both daytime and nighttime algorithm development. MAS offers a very wide range of channels, many of which overlap the ASTER channels. This is particularly valuable for the thermal channels, especially in the 8.5, 11 and 12 µm regions.

In validating the operation of the daytime and nighttime algorithms on Landsat TM imagery, three approaches are being used. The first and third are quantitative approaches and the second is more subjective. All three have limitations but, in general, they complement each other. They are described following.

In the first method, the algorithm is applied to a labeled set of samples. To date approximately 3700 contiguous pixel regions (made up of several hundred thousand pixel samples) have been extracted and labeled by a human expert trained in identifying features in polar imagery. For every pixel, in every contiguous sample, the classification result from the algorithm is compared to the labeling and a "confusion" matrix is generated indicating the percentage of classification for each combination. If the algorithm performed perfectly, the confusion matrix would be diagonal in which each diagonal element was 100 percent. This validation method somewhat overestimates the accuracy of the algorithm because the human expert tends to select spectrally homogeneous and unambiguous samples which are generally classified at a higher accuracy rate than for an entire scene. However, it serves to provide an indication of the upper limit of performance and, empirically speaking, indicates within 5 percent the overall scene classification accuracy, on the average.

This method serves as the basis for validating the accuracy of the algorithm over the life of the product.

The second approach is somewhat subjective and involves visual comparison of the classification result with the imagery by a human expert. The expert has access to a tool that allows him to augment his analysis through the use of three-band overlays, and various image processing techniques such as contrast stretching and histogram equalization. When available, more than one expert is used and, generally, their subjective estimates of the accuracy are within 5 percent of each other.

The third approach is an attempt to quantify the overall scene classification accuracy. This is derived from a tool that a human expert uses to label randomly selected regions within sets of Landsat TM imagery. It is very much like the process that an expert uses to extract labeled samples except the computer randomly selects the region to be labeled as opposed to the expert selecting the region. The random selection of samples by the computer provides for a more objective estimate of classification accuracy when these samples are compared against the results obtained from the classifier.

The results obtained from these methods will also be compared with independent observations from ground-, air-, and other satellite-based observations. Comparisons with these other types of observations will be conducted over an extended period of time for a variety of circumpolar regions.

4.1.2.2 **Sampling requirements and trade-offs.** Although current testing and validation of the algorithm is based on approximately 100 Landsat TM quad scenes, due to the limited areal coverage of each quad scene (approximately 100 km by 100 km), the entire circumpolar region is represented at a fraction of a percent. Antarctica is only represented in 24 scenes for 3 months in 1989 and only over coastal areas. Likewise, MAS scenes only are available for a limited time and a limited area. These sample sets poorly represent the polar regions both in space and time. Currently it is not possible to define the accuracy of the classification, for example, by latitude, ecosystem, season, etc.; however, during post-launch, as ASTER obtains data over various polar regions, representative samples will be extracted and included in the training set for the classifier. Analysis of the distributional nature of important features for classification will be conducted periodically. If warranted either the classifier will be retrained or if correlative analysis manifests a unique condition by latitude, ecosystem, etc., an additional version of the classifier will be installed to accommodate it. (Note an additional version of the classifier implies a new file containing a set of weights for the neural network classifier.)

4.1.2.3 Measures of success. If the classification accuracy for cloud/clear is greater than 95 percent, when applying the algorithm to all available labeled pixel samples (several hundred thousand to date), the validation will be considered successful. During the prelaunch phase, samples are extracted primarily from Landsat TM and MAS. During the post-launch phase, labeled samples will be extracted primarily from ASTER and secondarily from Landsat 7 ETM+.

Once confidence has been gained with the cloud masking algorithm, the results will be compared to independent observations. There are problems with comparing satellite-derived results with ground-based lidars, radars, and surface-based observations. For example, the relatively small field of view of the lidars and radars, the aspect angle, and time of observation as compared to the satellite sensor are important. Similar problems must be considered for surface observer estimates of cloud cover location. Cloud sides are seen by surface observers and satellites at different observational angles and from opposite sides of the cloud (top vs. bottom). There are two approaches for comparing the independent observations with satellite cloud masks. First is to compare the imager pixels closest to the ground site with temporally averaged (e.g., 10 minutes)

lidar or radar data. The second approach is to perform a temporal average of the ground-based data over a period of hours and to compare it with a spatially averaged cloud mask.

4.1.3   Pre-launch algorithm test/development activities.

4.1.3.1   Field experiments and studies. The most efficacious approach for algorithm development and testing has been through the use of the existing NASA government archives of Landsat TM and MAS data. In terms of both spectral and spatial resolution, and circumpolar region coverage, Landsat TM and MAS data serve as the best surrogates for ASTER data. The best quantitative method for validating the algorithm results is through analysis of the cloud mask and testing of the algorithm on labeled samples

4.1.3.2   Operational surface networks. It is anticipated that the following products will be used in pre-launch activities:

1.   National Weather Service observations (especially in polar regions such as in Fairbanks and Anchorage, AK)
2.   Ceilometer network (limited to wintertime conditions at continental U.S. sites)
3.   DOE ARM site data from Oklahoma (during wintertime conditions)

4.1.3.3   Existing satellite data. To the degree affordable, Landsat TM data has been purchased to support the testing and validation of the algorithm. In addition, more Landsat TM data has been obtained at no charge through data sharing agreements. The large ALASKA and ARMCAS MAS data archive is being utilized.

4.1.4   Post-launch activities

4.1.4.1   Planned field activities and studies. The validation effort for this algorithm will take advantage of any data obtained from field studies conducted post-launch in which polar-like conditions are present. No special field activities are planned for validation of this product. When coincident cloud masking data is available from Landsat 7 (ETM+), comparisons will be made. In the case of reasonable agreement, supporting validation is provided. If they do not, additional analysis will be required. The validation effort using surface observations will take advantage of the enhanced surface based measurement capabilities located at the DOE ARM sites in Alaska and Oklahoma. The use of the ARM site data from Oklahoma will be limited to wintertime conditions, especially when snow and clouds are present during the time of overpass. An opportunity for validating the algorithm in the detection of thin cirrus will also occur during overpasses of Salt Lake City, UT using the long range lidar located there (Sassen).

4.1.4.2 ASTER STAR Sites for CERES and MODIS

Note: Daytime acquisitions include all 14 VNIR, SWIR and TIR channels
         Nighttime acquisitions include all 5 TIR channels

In Priority Order:

I.   *ARM sites*: 5 sites, **1200 total scenes** over 5 years, (58% during daytime)
      Purpose is to help validate retrievals of cloud properties, radiative fluxes and aerosol
      properties. The sampling strategy is dependent upon how rapidly the surface properties

and clouds change. The Oklahoma ARM site is the most variable for land and cloud both in terms of season and day-to-day. ARM sites have the best characterization of the atmosphere anywhere in the world, and these are the highest priority CERES and MODIS acquisitions. The tropical ocean and polar ARM sites experience slower and fewer changes in cloud type and surface conditions with season; therefore, they require fewer acquisitions.

Site	Lat/Long	Requirement
ARM, Oklahoma	40.03/-105.27	4 daytime/mo * 60 mo = 240
		2 night/mo * 60 mo = 120
ARM, Manus Island	-2.06/147.43	2 daytime/mo * 60 mo = 120
		2 night/mo * 60 mo = 120
ARM, Nauru Island	-0.53/166.92	2 daytime/mo * 60 mo = 120
		2 night/mo * 60 mo = 120
ARM, Kiritimati Is. (Christmas Is)	1.87/-157.33	2 daytime/mo * 60 mo = 120
		2 night/mo * 60 mo = 120
ARM, Barrow*	71.27/-156.83	1 daytime/mo * 60 mo = 60
		1 night/mo * 60 mo = 60

*Note that the polar scenes will be 2 daytime per month (polar day) plus 2 nighttime per month (polar night).

I. ***BSRN   sites*** (Baseline Surface Radiation Network): Purpose is to obtain verification of surface radiation and cloud properties in climatic regimes not covered by the ARM sites. These sites have the world's most accurate surface radiation measurements. Each site requires 1 daytime and 1 nighttime acquisition per month for each of the 60 months of the mission, or 120 acquisitions per site. 14 sites * 120 acquisitions/site = **1680 total scenes**.

BSRN Site	Lat/Long
Alice Springs, Australia	-23.70/133.87
Aswan, Egypt	24.00/33.00
Bermuda	32.37/-64.70
Bondville, Illinois	40.05/-88.37
Boulder, Colorado	40.03/-105.30
Carpentras, France	44.05/5.05
Florianopolis, Brazil	-27.58/-48.52
Fort Peck, Montana	48.00/-105.00
Georg von Neumayer, Antarctica*	-70.65/-8.25
Goodwin Creek, Mississippi	34.00/-90.00
Ny Alesund, Spitsbergen, Norway*	78.93/11.93
Ping Chuan, China	28.00/102.00
Syowa base, Japan, Antarctica*	-69.00/39.35
ToravereObs/Estonia	58.33/26.73

II.   **_SURFRAD and AERONET sites_**: These sites are primarily to characterize different aerosol types around the world. <u>Daytime sampling only</u>. 1 daytime acquistion per month for each of the 60 months, or 60 acquisitions per site.
20 sites * 60 acquistions/site = **1200 total scenes.**

Site	Lat/Long	
Sable Island, Nova Scotia	43.93/-60.01	Perturbed Marine
K'puszta, Keszcemet, Hungary	46.97/19.55	Perturbed Continental
Cheeka Peak, Washington	48.30/-124.62	Clean Marine
Tucson, Arizona	32.23/-110.95	Desert
GSFC, Maryland	39.03/-76.88	Urban
Bondoukoui	11.85/-3.75	West Africa
Ouagadougou	12.20/-1.40	West Africa
Catalina Island, California	34.00/-119.00	Urban
Cuiaba	-15.50/-56.00	Biomass Burning
Dry Tortugas	24.60/-82.80	Florida
Bermuda	32.37/-64.70	Atlantic
Capo Verde	16.73/-22.94	W. Coast Africa
Ascension Island	-7.98/-14.41	Atlantic
Barbados	13.00/-60.00	Caribbean
La Reunion	-20.00/55.50	S. Indian Ocean
Mongu	-15.50/23.00	S. Central Africa
Male	5.00/74.00	S. India, ocean
Tromelin	-16.00/54.50	Biomass Burning
Okinawa	25.50/128.00	S. China Sea
Crete	35.00/25.00	Mediterranean

III.   **_Polar sites_**: These sites are in addition to the ARM and BSRN sites listed above. They are to cover major types of snow/ice/cloud conditions in the polar regions. Note that the polar scenes will be 2 daytime per month (during polar day) plus 2 nighttime per month (during polar night). Average 1 acquisition/month * 60 months (=60 acquisitions/site) * 9 sites = **540 total scenes.**

Polar Site	Lat/Long
Alert Airport, Canada	82.31/-62.17
Baker Lake Airport, Canada	64.18/-96.05
Upernavik, Greenland	72.47/-56.10
Norilsk, Russia	69.20/88.06
Franz Josef Land, Russia	80.00/55.00
McMurdo, Antarctica	-77.51/166.40
Palmer, Antarctica	-64.46/-64.05
Casey, Antarctica	-66.01/111.05
Brockton, Antarctica	-78.48/-174.40

V.   **_Ocean sites_**: These sites are chosen to characterize the Scherr Global Cloud Climate regions not covered by the existing land and island site data listed above. Specific ocean locations are listed rather than strips, since clouds will not be independent in adjacent 60 km ASTER scenes along an orbit track. 1 daytime and 1 nighttime acquisition per month is requested for each site * 60 months
 (= 60 acquisitions/site) * 16 sites = **960 total scenes**.

Lat/Long
53.00/-150.00
43.00/-153.00
37.00/-123.00
30.00/-119.00
20.00/-133.00
15.00/-103.00
 7.00/-90.00
-7.00/-137.00
-10.00/65.00
-20.00/-75.00
-28.00/37.00
-30.00/-75.00
-33.00/-142.00
-40.00/150.00
-45.00/-145.00
-55.00/-147.00

IV. ***Special MODIS sites:*** These sites have been selected by the MODIS team. The Tibetan Plateau is expected to be an especially difficult region for cloud masking. Two others sites are in and out of the monsoon region, one in the middle of India and one that is often cloud covered, in Borneo. 1 daytime and 1 nighttime acquisition is requested for each of the three sites/mo * 60 months = 120 scenes/site. **Total 360 scenes** over 60 months.

Site	Lat/Long
TibetanPlateau	33.00/90.00
India	17.00/77.50
Borneo	00.00/115.00

Summary of Scenes:

I.	1200
II.	1680
III.	1200
IV.	540
V.	960
VI.	360
Total	5940

4.1.4.3 Validation of the International Satellite Cloud Climatology Project (ISCCP) Cloud Droplet Size Products in stratocumulus cloud fields

TargetAreas: Four scenes/year are requested for each of the following six target areas for five years, for a total of 120 scenes. The approximate date for the scene acquisitions should be the first half of the month of Jan, April, July and Oct of each year.

Sites:

Equator, 75W	Near nadir of GOES-8
Equator, 135W	Nadir of GOES-9
10S, 80W	West coast of South America

15S, 77W	West coast of South America
33N, 120W	West coast of North America
45N, 125W	West coast of North America

Research Outline and Objectives: Cloud microphysics is an important factor in climate change studies. Currently, we are developing retrieval schemes to be used by ISCCP, headed by Dr. Bill Rossow at NASA GISS, to generate cloud particle size products. To date the satellite data used in this process has been AVHRR radiances. While the results have veen verified through ground observations, there are two potential problems that need further investigation. The first is the effect of sub-pixel cloud cover and the second is the viewing geometry effect on cloud droplet size retrievals. We propose to use the combination of ASTER and GOES data to evaluate these two effects.

Reasons why ASTER data is necessary: To order to evaluate the effect of fractional cloud cover on particle size retrievals using AVHRR data at 1 km resolution at nadir, one requires imagery at much finer resolution. In evaluation of the viewing geometry effect, we must eliminate the possible effect of fractional cloud cover at sub-pixel scales. With its spatial resolution of 15m in the VNIR and spectral coverage from visible to infrared, ASTER data is ideal for this investigation.

4.1.4.4   Rainforests of the Cordillera De Tilaran of Costa Rica

Target Areas:

10 35' N 84 45'W
11 05' N 84 45'W
10 05' N 84 45'W
10 35' N 84 15'W
10 35' N 85 15'W

5 sites x 5 times/year (Jan, Feb, Mar, Apr, May) x 5 years = 125 total acquisitions

Research Outline and Objectives: To relate aspects of ground-level cloudiness to vegetation structure and species distribution of the rainforests of Costa Rica. This is of critical importance to hydrological management, biodiversity conservation, and land-use planning in the montane tropics during an era of dramatic environmental changes. This project will correlate floristic and structural features of the tropical montane ecosystems in the Cordillera de Tilaran or northwestern Costa Rica with features of the trade wind cloud field. This will provide quantitative data on the impact of variations of trade wind cumulus on the local biogeography of vegetation structure and composition of a tropical mountain range at a spatial scale relevant to evolutionary and metapopulation processes. Topographic dissection of tropical mountain ranges, particularly those exposed to directionally reliable winds, and the consequent edaphic and climatic heterogeneity, result in steep gradients in vegetation structure and ecosystem function, habitat patchiness, and species ranges that are highly fragmented. This environmental heterogeneity makes for explosive speciation with the result that cloud-forested tropical mountains are major centers of biodiversity. Yet, at the same time, the isolated nature of habitat units and their limited size dictate steep environmental gradients which makes these regions sensitive to environmental changes – particularly those involving the cloud fields which literally define the distribution of cloud forest. The data on the relationships between the cloud field and montane forest structure and local species distributions will provide quantitative assessments of the ecological impacts of changes in cloud fields due to climatic oscillations like ENSO.

Reasons why ASTER data are necessary: High spatial resolution satellite imagery is necessary for this project, both to determine cloud base height and for forest mapping. Cloud base level will be

determined from cloud shadows and eill be compared to those determined daily at specific ground monitoring sites in the upper Rio Guacimal/San Luis valley, the Monteverde/Cerro Plano plateau, and the upper Penas Blancas valley. Vegetation sampling will be conducted at 50 sites in a stratified partially randomized array between 1000 and 1860 m elevation (20 sites on the Pacific slope and 30 sites on the Caribbean side). From previous work in the area, anticipated spatial accuracy for the ground validation work is to within 200m. This will be improved using GPS technology. Geographic data for each site will include elevation, azimuth and slope. Detailed height to canopy strata, defined as regions of foliage bearing leaves vertically separated by at least 3m from layers above and below, tree height, crown breadth and trunk size will be measured. Certain forest types are readily distinguishable from satellite data, as are clumps of some species of canopy trees. Ground data will provide exact locations of homogeneous vegetation as well as those of known mixed composition. A neural network will be trained to inventory forest vegetation and changes in that vegetation at the sub-pixel scale. Ground personnel equipped with GPS will verify the location of road intersections, outcrops, peaks and stream confluences for use as tie points. It is essential to have the highest possible spatial and spectral resolution for these studies.

4.1.4.5   Satellite   Archeology

Area and Purpose: Eastern Utah for an archeological mapping project of ancient Anasazi roads which are currently being destroyed in the wake of urban expansion and development. Previous research has shown that prehistoric roads in Chaco Canyon and elsewhere in northwest New Mexico link important prehistoric Pueblo (Anasazi) "great houses," dwelling units, and ritual centers. These roads are characteristically up to 10 meters wide and very straight, passing over hills and valleys rather than following landscape contours. Recently, archaeologists have discovered segments of similar roads in Southeast Utah. The SE Utah roads also link Anasazi great houses and other structures typical of the late 11th through 13th centuries, suggesting that the Anasazi of that era were far more integrated socially and politically than heretofore realized. However, road-connected sites also contain components from the 10th and 11th century, raising the possibility that the roads significantly predate the great houses.

One of the major questions of SW archaeological research is how the Anasazi were able to live and prosper for centuries in the austere climate of the Colorado Plateau. Several hypotheses have been advanced to address this issue, including the use of roads as distribution networks. Understanding the age and function of the roadway system may provide insight into how the Anasazi sustained themselves for hundreds of years as they traded and transported goods across their landscape. Other hypothesized road functions range from communications links in alliance networks to architectural expressions of myth and ritual on a mythico-sacred landscape. Understanding the age and function of the roadway system may provide insights into how Anasazi communities interacted with one another and their surrounding landscape, enabling them to sustain themselves successfully for centuries.

Period: ASTER data, twice each year, for four subsequent years. It would be best if this data were acquired for each year in the early spring (April) and late summer (August-September). Previous data acquisition missions in Chaco Canyon have demonstrated the utility of these seasonal data sets for prehistoric road detection.

Area: Our study area is a rectangular area described by the following coordinates:
109 25' W to 109 45' W and from 37 15'N to 37 45'N. The center point is:
 37 30'N/ 109 35'W.

4.1.4.6   New EOS-targeted coordinated field campaigns. With the aforementioned comparison of ground-, air-, and satellite-based observations with the cloud mask, there will still be a deficiency of validation information. It would be beneficial to plan field campaigns in the circumpolar regions, or join already planned field campaigns there to fill in data gaps and augment the validation process.

4.1.4.7 Needs for other satellite data. Landsat enhanced thematic mapper plus (ETM+) data is needed to augment the transition of the algorithm from Landsat TM to ASTER data. It is also needed for comparison of cloud masking products in the validation process.

4.1.4.8 Measurement needs (in situ) at calibration/validation sites: land, buoys, etc. Currently, the algorithm is designed to take advantage of geographic databases, such as coastlines, ecosystems, land character, elevation, etc.; however, the coarse spatial resolution currently available from these datasets makes their use limited. Improved higher spatial resolution surface characterizations are needed.

4.1.4.9 Needs for instrument development (simulator). Not applicable.

4.1.4.10 Geometric registration site. Not applicable.

4.1.4.11 Intercomparisons (multi-instrument). The cloud mask obtained from ASTER will be intercompared with that obtained from Landsat 7 ETM+. Other product developers (e.g., MODIS SNOMAP and ICEMAP, CERES cloud mask) utilizing lower spatial resolution data will be comparing their results with this ASTER derived product.

4.1.5 Implementation of validation results in data production

The confusion matrices derived from testing of the algorithm on labeled pixel samples will be included in a product file as an annotation element. There are currently 4 cloud classes and 6 clear classes. The matrices will be much like those shown in Tables 3 and 4. This will provide the user with an indication of the relative accuracy of each class, as well as cloud/clear classification accuracy.

4.1.5.1 Approach (including long-term calibration considerations). Periodically (approximately every 3 months in the first two years), ASTER scenes obtained over polar regions will be randomly selected for the extraction of new labeled pixel samples. If algorithm testing results indicate degraded performance, the algorithm will be retrained and tested again. If the results are still degraded, the data will be partitioned into the determining geographical and/or seasonal conditions and multiple versions of the algorithm will be implemented at the DAAC. (Note: This does not imply more than one algorithm, only multiple table lookups which the algorithm will select automatically.)

Algorithm performance will continue to be checked against available surface-based, air-based, and satellite-based data when available.

4.1.5.2 Role of EOSDIS. The product resident within EOSDIS will have a pointer to the developer of the data product. This will provide a feedback mechanism in which users can alert the developer of possible errors or discrepancies. User feedback will provide an additional independent validation of the product that could potentially stimulate investigations into methods for product improvement.

4.1.5.3 Plans for archival of validation data. The validation results will be archived by the developer at his site. Subsampled versions of the classification masks generated for most of the scenes in the validation process will be maintained at an ftp site. Users will be able to request the full resolution product through the same site. The results from the validation efforts will be published in the peer-reviewed literature.

4.1.6 Summary

The fundamental basis for development of the ASTER Polar Cloud Mask algorithm is the extraction of labeled pixel samples primarily from Landsat TM and MAS data. These labeled samples are used to train the algorithm and define the distribution of the feature space for each class. They also provide the basis for validation of the algorithm. As long as the classification accuracy of the algorithm on the labeled samples is > 95%, the algorithm provides adequate cloud masking results (on the average). Appropriate data sets continue to be acquired and tested during the prelaunch phase and will continue to be acquired during the post-launch phase to ensure the classification accuracy does not drop below 95%. This does not imply that the algorithm always classifies at least 95% of the pixels in each scene correctly for cloud and clear, but that it will do so at an accuracy of approximately 95% on the average. This strategy for validating cloud masking accuracy is augmented by two other methodologies described in 4.2.1. This product will be derived from the ASTER products for radiance/temperature at the sensor. The assumption is that these inputs are valid (unless otherwise indicated, in which case the pixel will remain unclassified and passed through as bad/marginal/etc.).

4.2   Exception Handling and Missing Band

The algorithm will be applied only to "good" data. It is assumed that the data will be screened at the EDC DAAC before standard products are produced. If there are missing channels, then the following channels will be substituted:

ASTER Channels		If Not Available Then We Use	
(1)	0.56 μm	(2)	0.66 μm
(3n)	0.81 μm	(2)	0.66 μm
(4)	1.65 μm		
(5)	2.16 μm	(6)	2.20 μm
(10)	8.3 μm	(11)	8.65 μm
(13)	10.6 μm	(12)	9.1 μm
(14)	11.3 μm		

4.3   Limitations

For the present algorithm development, the Landsat TM calibration is adequate for channels 1-5 and 7. However, the infrared channel 6 is essentially uncalibrated. This has caused difficulties in our algorithm development. For example, ocean surfaces often are retrieved with surface temperatures on the order of – 12°C. However, unless they are covered with ice, the ocean temperatures cannot be below –1.8°C. Another problem is that most of the present 82 Landsat scenes do not have water in the scene. We are forced to take the average of these rescaling values found in the water scenes and apply it to the non-water scenes. On the other hand, MAS data sets do not suffer from these LANDSAT problems, but they do seem to be somewhat inconsistent in their calibration from one flight series to the next; i.e., ALASKA and ARMCAS.

4.4   Quality Assurance (QA)

This plan describes all the information that is to be included in the ASTER Polar Cloud Mask product that a user will have available to assess the overall quality of the cloud mask and/or the specific quality of the classification of each pixel.

4.4.1 *Carry-Through QA Information*.

Carry-through refers to QA information available from the header records of input dependencies. For this product input dependencies include ASTER registered radiance (AST03) and temperature (AST04) at the sensor (see Figure 5).

a. Level 1B input dependencies

All QA information present in the headers of the ASTER level 1B input dependencies are "carried- through" to (duplicated in) the Level 2 ASTER Polar Cloud Mask product header. This information includes:
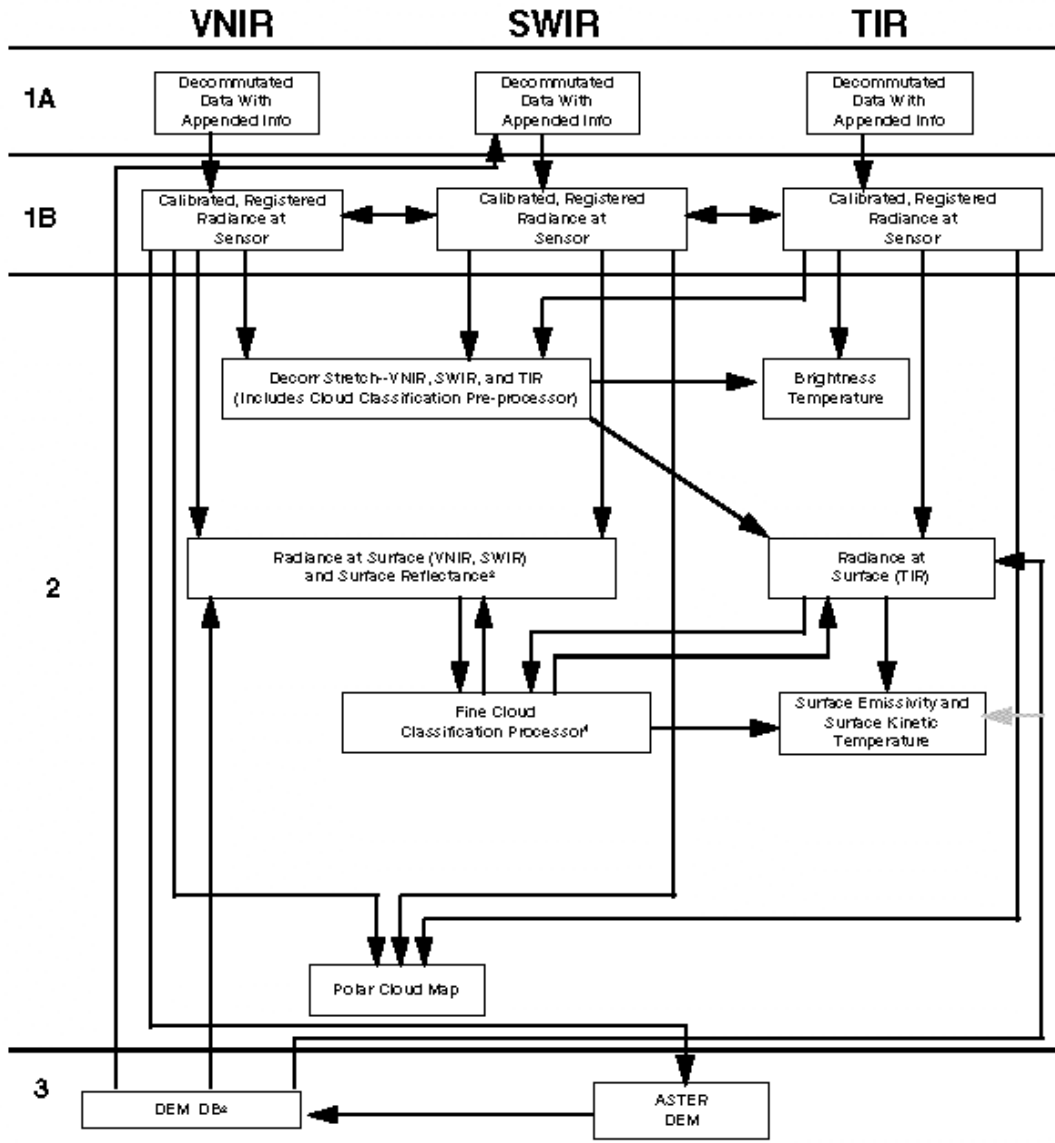
1) Percent of bad pixels in the full scene and each data plane
2) Bad pixel list indicating
   a) Data plane/band
   b) Line and sample or range/region of line/sample
   c) Reason the pixel or range/region of pixels is bad
3) Cloud coverage (percent full scene and per quad) (Note: This is fractional cloud coverage as determined by the Level 1B input dependency. A fractional cloud cover value generated by this product is provided as a separate estimate in the content for this product.)
4) Inter-telescope registration information
5) SWIR parallax correction information
6) Calibration degradation metric

b. Other input dependencies.

Currently the only other input dependencies intrinsic to the derived pixel classification in the Polar Cloud Mask product are the World Data Bank II coastline system, a digital elevation model, and the EPA Global Ecosystems database. These inputs are relatively static and will be complete when incorporated into the Product Generation System. That is, these inputs are not comprised of bad/missing/erroneous values; therefore, there is no QA information to pass through.

Other input dependencies such as sea ice and snow coverage from MODIS (MOD 29 and MOD10, respectively) and surface temperature analysis from DAO are not applied directly in the Polar Cloud Mask Algorithm in the derivation of the pixel classification. They are only used to augment the certainty measure of the pixel classification which is described in more detail in section 4.5a. Bad data from these input dependencies will be ignored in the determination of the certainty measure.

ASTER Product Inter-Dependencies

VNIR   **SWIR**   **TIR**

1A
Decommutated Data With Appended Info
Decommutated Data With Appended Info
Decommutated Data With Appended Info

1B
Calibrated, Registered Radiance at Sensor
Calibrated, Registered Radiance at Sensor
Calibrated, Registered Radiance at Sensor

Decorr Stretch--VNIR, SWIR, and TIR (Includes Cloud Classification Pre-processor)

Brightness Temperature

2
Radiance at Surface (VNIR, SWIR) and Surface Reflectance*

Radiance at Surface (TIR)

Fine Cloud Classification Processor¹

Surface Emissivity and Surface Kinetic Temperature

Polar Cloud Map

3
DEM DB³

ASTER DEM

¹Produces a cloud mask that is incorporated into other products
²Computed simultaneously with Radiance at Surface
³Refers to a database of DEM data regardless of the source

6 Dec 95

Figure 5. *Diagram of ASTER product inter-dependencies.*

40

4.4.2 Bad Data

This section addresses the encoding of pixels corresponding to bad data present in the AS-TER input dependencies (i.e., registered radiance and/or temperature at the sensor). These pixels are encoded with a 255 with the following exception. In the Polar Cloud Mask algorithm, during a final process, a spatial consistency test is performed to locate isolated pixels. Isolated pixels are defined as those which are classified differently than 6 or more of their nearest neighbors if those 6 nearest neighbors are classified the same. Isolated pixels are classified and/or reclassified the same as (or spatially consistent with) their like nearest neighbors. The QA data plane incorporates a certainty of correct classification for each pixel (see section 4.5a). The certainty of the classification of the isolated pixel is derived from the certainty values of its like nearest neighbors The encoding of the isolated pixels will not indicate whether the pixel was derived from a bad pixel, was the result of an unknown classification from the algorithm, or was classified but was inconsistent spatially with its nearest neighbors.

4.4.3 Suspect/Unknown Data

The algorithm derives a crude measure of the certainty of the classification for each pixel and outputs an unknown classification if necessary. This measure is incorporated into the encoding scheme and is briefly described in section 4.5.a. The algorithm does not provide any other mechanism for identifying suspect/unknown data.

4.4.4 Cloudy Pixels

The purpose of the Polar Cloud Mask Algorithm is to classify each (good) pixel into one of 10 classes whether they be cloud or not. There is no special handling for cloudy pixels beyond classifying a pixel as cloudy if the algorithm so determines.

4.4.5 Product-Unique Information.

a. Encoding of classification certainty

The Polar Cloud Mask algorithm classifies each good pixel into one of ten classes or as unknown, if necessary. During the classification process it also derives an approximate measure of certainty or confidence in the pixel classification. This certainty is coded into the QA data plane and classification mask as follows:

4 most significant bits of QA data plane

Dec	Binary	Meaning
8-14	1001	This pixel has been discovered to be bad during the production of this product due to out-of-range feature value/s (e.g., VNIR or SWIR band reflectance values less than 0 or greater than 2.0) Note: Other six codes not used.
1-6	0001	After processing, this pixel is now deemed suspect due to marginally out-of-range feature value/s (e.g., Band 14 brightness temperature greater than 310 K for a particular geographic region and season) Note: Other five codes not used.

bits 7 and 8 of the classification mask

Dec	Binary	Meaning
0	00	Certainty measures derived from the classification algorithm indicate a high confidence in the result (>90%)
1	01	Certainty measures derived from the classification algorithm indicate moderate confidence in the result (50%>90%)
2	10	Certainty measures derived from classification algorithm indicate low to moderate confidence in the result (10%>50%)
3	11	Certainty measures derived from the classification algorithm indicate a relatively low confidence in the result (i.e., the feature vector is highly ambiguous between 2 or more classes); however, the result is possibly correct. (<10%)

b. Composite certainty values

Mean certainty values for each class, for cloud/no cloud, and for all classified pixels is also included in the product header. These measures provide an overall indication of the degree of confidence in the classification. For example, a scene comprised of a large fraction of transparent cloud might result in a low overall scene/class classification confidence values since the feature vectors of transparent clouds are highly ambiguous with those of the underlying features. On the other hand a cloud free scene with low atmospheric turbidity and water vapor loading might result in a high overall scene/class certainty values if the feature vectors for the scene were unambiguous with other classes.

ACKNOWLEDGMENTS

REFERENCES

Allen, A. G., R. M. Hanson and J. W. Erismen, 1989: Field measurement of the dissociation of ammonium nitrate and ammonium chloride aerosols. *Atmos. Env.*, **23**, 1591-1600.

Barry, R. G., A. Henderson-Sellers and K. P. Shine, 1984: Climate sensitivity and the marginal crysosphere. In *Climate Processes and Climate Sensitivity (Geophysics Monograph 29)*, ed. J. E. Hansen and T. Takahashi, Am. Geophys. Union. 221 pp.

Berendes, T.A., R.M. Welch, Q. Trepte, S. Schaaf, and B.A. Baum, 1996: The EOS CERES Global Cloud Mask. *Proceedings Eighth Conference on Satellite Meteorology and Oceanography*, Atlanta, GA, AMS, --.

Coakley, J. A., and F. P. Bretherton, 1982: Cloud cover from high-resolution scanner data: Detecting and allowing for partially filled fields of view. *J. Geophys. Res.*, **87**, 4917-4932.

Crane, R. G., and M. R. Anderson, 1984: Satellite discrimination of snow/cloud surfaces. *Intnl. J. Rem. Sens.*, **5**, 213-221.

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

Ebert, E., 1989: Analysis of polar clouds from satellite imagery using pattern recognition and a statistical cloud analysis scheme. *J. Appl. Meteor.*, **28**, 382-399.

Ebert, E., 1992: Pattern recognition analysis of polar clouds during summer and winter. *Intnl. J. Rem. Sens.*, **13**, 97-109.

Key, J., and R. G. Barry, 1989: Cloud cover analysis with Arctic AVHRR data. 1. Cloud detection. *J. Geophys. Res.*, **94**, 8521-8535.

Key, J., 1990: Cloud cover analysis with Arctic AVHRR data. Part II: Classification with spectral and textural measures. *J. Geophys. Res.*, **95**, D6, 7661-7675.

King, M. D., and S. C. Tsay, 1993: Theoretical basis of cloud retrieval algorithms for MODIS: Cloud cover, thermodynamic phase, optical thickness, and effective particle radius. MODIS Algorithm Theoretical Basis Document, NASA Goddard Space Flight Center.

Kuo, K-S., R. M. Welch and S. K. Sengupta, 1988: Structural and textural characteristics of cirrus clouds observed using high spatial resolution LANDSAT imagery. *J. Appl. Meteor.*, **27**, 1242-1260.

Li, Z., and H. G. Leighton, 1991: Scene identification and its effect on cloud radiative forcing in the Arctic. *J. Geophys. Res.*, **96**, 9175-9188.

Logar, A.M., E.M. Corwin, and W.J. Oldham, 1992: An efficient gradient descent learning scheme using weight trajectory projections. *Proceedings Oklahoma Artificial Intelligence Conference*.

Logar, A.M., E.M. Corwin, and W.J. Oldham, 1993: A comparison of recurrent neural network learning algorithms. *Proceedings I.E.E.E. Conference on Neural Networks,* Vol. II, 1129-1134.

Logar, A. M., D. E. Lloyd, E. M. Corwin, M. L. Penaloza, R. E. Feind, T. A. Berendes, K.S. Kuo and R. M. Welch, 1998: The ASTER Polar Cloud Mask, IEEE Trans Geosci. Remote Sens, 36, 1302-1312.

Markham, B. L., and J. L. Barker, 1986: LANDSAT MSS and TM post calibration dynamic ranges, exoatmospheric reflectances and at satellite temperatures. LANDSAT Technical Notes, No. 1, EROS Data Center, Sioux Falls, SD.

McGuffie, K., R. G. Barry, A. Schweiger, D. A. Robinson and J. Newell, 1988: Intercomparison of satellite derived cloud analyses for the Arctic ocean in spring and summer. *Intnl. J. Rem. Sens.,* **9**, 447-467.

Menzel, P., and K. Strabala, 1993: Cloud top properties and cloud phase algorithm theoretical basis document. Cooperative Institute for Meteorological Satellite Studies, Space Science and Engineering Center, University of Wisconsin-Madison.

Minnis, P., and E. F. Harrison, 1984: Diurnal variability of regional cloud surface radiative parameters derived from GOES data: Part I: Analysis method; Part II: Cloud results; Part III: Radiative budget. *J. Clim. Appl. Meteor.,* **23**, 993-1051.

Ormsby, J. P., and D. D. Hall, 1991: Spectral properties of fog over the Malospina Glacier, Alaska, in comparison to snow, ice, and clouds. *Photogrammetric Engineering and Rem. Sens.,* **Vol. 57,** No. 2, 179-185.

Paola, J.D., and Schowengerdt, R.A., 1995: A review and analysis of back propagation neural networks for classification of remotely-sensed multi-spectral imagery. *International Journal of Remote Sensing*, **16**, 16, 3033-3058.

Parkinson, C. L., J. C. Comiso, H. J. Zwally, D. J. Cavalieri, P. Gloersen and W. J. Campbell, 1987: Arctic sea ice, 1973-1976: Satellite passive microwave observations, NASA SP-489, Scientific and Technical Branch, Washington, DC. 296 pp.

Rabindra, P., S. K. Sengupta and R. M. Welch, 1992: An interactive hybrid expert system for polar cloud and surface classification. *EnvironMetrics*, **3**(2), 121-147.

Raschke, E., P. Bawer and H. J. Lutz, 1992: Remote sensing of clouds and surface radiation budget over polar regions. *Intnl. J. Rem. Sens.,* **13**, 13-22.

Richards, J. A., 1993: *Remote Sensing Digital Image Analysis.* 2nd Edition, Springer-Verlag, New York. 340 pp.

Rossow, W. B., 1989: Measuring cloud properties from space: A review. *J. Climate*, **2**, 201-213.

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

Rossow, W. B., C. L. Brest and L. C. Garder, 1989b: Global, seasonal surface variations from satellite radiance measurements. *J. Climate*, **2**, 214-247.

Rumelhart, D., G. Hinton, and R. Williams, 1986: Learning internal representations through error propagation. Parallel Distributed Processing: Exploration in the Microstructure of Cognition, edited by D. Rumelhart and J. McClelland, MIT Press, Cambridge, MA, 318-362.

Sakellariou, N. K, and H. Leighton, 1988: Identification of cloud-free pixels in inhomogeneous surfaces from AVHRR radiances. *J. Geophys. Res.,* **93,** D5, 5287-5293.

Saunders, R. W., and K. T. Kriebel, 1988: An improved method for detecting clear sky and cloudy radiances from AVHRR data. *Intnl. J. Rem. Sens.,* **9**, 123-150.

Schlesinger, M. E., and J. F. B. Mitchell, 1987: Climate model simulations of the equilibrium climatic response to increase carbon dioxide. *Rev. Geophys.,* **25**, 760-798.

Serra, J., 1982: *Image Analysis and Mathematical Morphology.* Academic Press, London. 610 pp.

Sèze, G., and W. B. Rossow, 1991a: Time-cumulated visible and infrared radiance histograms used as descriptors of surface and cloud variations. *Intnl. J. Rem. Sens.,* **12**, 921-952.

Sèze, G., and W. B. Rossow, 1991b: Effects of satellite data resolution on measuring the space-time variations of surfaces and clouds. *Intnl. J. Rem. Sens.*, **12**, 130-132.

Shine, K. P., and R. G. Crane, 1984: The sensitivity of a one-dimensional thermodynamic sea ice model to a change in cloudiness. *J. Geophys. Res.,* **C89**, 10,615-10,622.

Simpson, J. J., and C. Humphrey, 1990: An automated cloud screening algorithm for daytime AVHRR imagery. *J. Geophys. Res.*, **C8**, 13,459-13,481.

Steffen, K., and J. E. Lewis, 1988: Surface temperatures and sea ice typing for northern Baffin Bay. *Intnl. J. Rem. Sens.*, **9**, 409-422.

Steffen, K., R. Bindschadler, G. Casassa, *et al.*, 1993: Snow and ice applications of AVHRR in polar regions: Report of a workshop held in Boulder, CO, 20 May 1992. *Ann. Glaciol.*, **7**, 1-16.

Stowe, L. L., H. Y. M. Yeh, T. F. Eck, C. G. Wellemeyer, H. L. Kyle and the NIMBUS-7 Cloud Data Processing Team, 1989: Nimbus-7 global cloud climatology. Part II: First year results. *J. Climate*, **2**, 671-709.

Stowe, L. L., E. P. McClain, R. Carey, P. Pellegrino, G. Gutman, P. Davis, C. Long and S. Hart, 1991: Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Adv. Space Res.,* **11**, 51-54.

Strabala, K. I., S. A. Ackerman and W. P. Menzel, 1994: Cloud properties inferred from 8-12 mm data. *J. Appl. Meteor.,* **33**, 212-229.

Thome, K., K. Arai, S. Hook, H. Kieffer, H. Lang, T. Matsunaga, A. Ono, F. Palluconi, H. Sakuma, P. Slater, T. Takashima, H. Tonooka, S. Tsuchida, R. M. Welch and E. Zalewski, 1998: ASTER Preflight and Inflight Calibration and the Validation of Level 2 Products, IEEE Trans Geosci. Remote Sens 36, 1161- 1172.

Tovinkere, V. R., M. Penaloza, A. Logar, J. Lee, R. C. Weger, T. A. Berendes, and R. M. Welch, 1993: An intercomparison of artificial intelligence approaches for polar scene identification. *J. Geophys. Res.,* **98**(D3), 5001-5016.

Welch, R. M., S. K. Sengupta and D. W. Chen, 1988a: Cloud field classification based upon high spatial resolution textural features. 1. Gray level co-occurrence matrix approach. *J. Geophys. Res.*, **93**, 12,663-12,681.

Welch, R. M., S. K. Sengupta and K. S. Kuo, 1988b Marine stratocumulus cloud fields off the coast of southern California observed using Landsat imagery. Part II: Textural analysis. *J. Appl. Meteor.,* **27**, 362-378.

Welch, R. M., K. S. Kuo and S. K. Sengupta, 1990: Cloud and surface textural features in polar regions. *IEEE Trans. Geosci. and Remote Sens.*, **28**, 520-528.

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

Wetherald, R. T., and S. Manabe, 1980: Cloud cover and climate sensitivity. *J. Atmos. Sci.*, **37**, 1485-1510.

THE ASTER POLAR CLOUD MASK

RESPONSE TO REVIEWER COMMENTS
TO THE ALGORITHM THEORETICAL BASIS DOCUMENT

DECEMBER 1998

Overall Evaluation was "C;" This algorithm is scientifically important; however, it contains minor deficiencies that must be rectified prior to development as an operational data product and early post-launch availability.

1. Will the products addressed in this ATBD help meet the objectives of the NASA MTPE program and the global change research agenda?
REVIEWER: I think it may be difficult at this point in time to monitor climate change in the polar regions since we currently have a pretty poor idea of what the current climate is or has been. However, the polar cloud and surface observations can be useful in understanding the relevant climate processes.
RESPONSE: The polar regions are receiving a great deal of attention. This includes the recent SHEBA program, the new ARM site, and MAS flights, to name a few. Extensive observations of the polar regions coupled with new surface measurements will help us understand the processes taking place in these regions. Long-term reliable monitoring of these regions is a first step.

2. Is the theoretical basis of the algorithm described in this document sound? Please list strengths, weaknesses, and recommendations for improvement below.
REVIEWER: The proposed approach seems fairly complex to me. There may be no way around this, since the simpler classification techniques have difficulty in polar regions.
RESPONSE: We have made progress in recent years in simplifying the algorithm.

2a. Major strengths of the approach.
REVIEWER: The algorithm uses practically every know trick to try to identify cloud and surface types in polar regions. Where one approach fails, hopefully another approach can succeed. Use of ancillary information (navigation, elevation, ecosystem, USN sea ice, etc) is good. (However, I wasn't clear on whether these would be used inside the classification procedure, or just as checks on sensible analyses).
RESPONSE: We have studied a wide range of different approaches for classification in polar regions, intercomparing both accuracy and cpu requirements. This has been accomplished using AVHRR, LANDSAT TM and MAS data sets. We now feel that we can achieve approximately 95% accuracy in the polar regions during daylight hours for solar zenith angles less than 80 degrees, and about 90% accuracy for solar zenith angles between 80 to 85 degrees. The nighttime accuracy is 80% to 85% accurate. These values are based on our most recent MAS data set analyses. Furthermore, the neural network approach has been shown to have the best overall performance in terms of accuracy and cpu (speed).
2b. Major problems with the approach.
REVIEWER: While the algorithm should work well during daytime, I would have liked to see more information and preliminary results for a nighttime algorithm, which will be needed full-time for 6 months at a time. Perhaps even tougher are the very low sun angle swaths, where albedo may be unreliable, and NIR (2.4 um for example) is sun-contaminated, making it difficult to use as IR temperatures. Another serious problem will be the training of the algorithm.
RESPONSE: It was impossible to develop a useable nighttime algorithm using the LANDSAT TM data; only one IR channel was available. However, the MAS ARMCAS and Alaska data sets have been used to create a new algorithm for both day and night. The results suggest that 80% to 85% accuracy can be achieved at night. However, this is based solely on the MAS results, and the data

probably are not representative for the broader range of conditions that will be encountered. A separate high zenith angle set of results have been developed for angles above 80 degrees and less than 85 degrees. We do not have a solution for solar zenith angles greater than 85 degrees.

2c. Recommendations to improve the approach.
REVIEWER: First, without having actual ASTER data, the imagery from other instruments (LANDSAT, AVHRR, etc) will only approximate what ASTER will see. Perhaps it would be worthwhile to synthesize some ASTER data using other instruments' data. It may be necessary to do quite a lot of retraining once ASTER is making measurements.
The other training problem I foresee is knowing the correct identification of a scene. Again this is more of a problem during polar night. It is not always clear, even to an expert using the most so-phisticated image processing/display systems. Some coincident satellite and surface observations are needed during polar night (perhaps during the SHEBA experiment).
RESPONSE: It is indeed difficult to develop the algorithm using only LANDSAT and AVHRR data. However, the MAS data with the 50 channels have been very useful. While exact matches to ASTER channels are not possible, they are reasonably close. Adjustments and retraining will have to be made after launch to create the final version. We have used the daytime algorithm first. Then we have used the nighttime channels to create the nighttime algorithm. In this way we can make a positive identification of the clouds, sea ice, land, water, etc. A set of validations is planned using the STAR sites.

3. Is it feasible to generate an operational product using this algorithm? If not, please explain why not.
REVIEWER: Yes. The proposed product is a classification image, with the value of each pixel containing information on each of the 8 classes (limited by the # of bits). The algorithm uses some ancillary information, namely the NAVY/NOAA sea ice product and NOAA snow data product, which are available on a weekly basis. How close to "real time" do you need to be in order to be operational?
RESPONSE: The newer version of the algorithm does not use the sea ice and snow data products. It is unlikely that the product needs to be used in a real-time basis. It will be produced as the data arrives from Japan.

4. What are the critical areas needing further research and development prior to development as an operational data product?
REVIEWER: More theoretical work on the expected signal at 8.5 um from various polar surfaces and clouds. This of channels will probably be critical for the nighttime algorithm.
More testing of the sensitivity of the results to solar irradiance, geometry of sun and viewing. How many sets of weights will the neural network require? What percent of the data will be unus-able due to calibration and/or viewing limitations? The fuzzy expert may need more refining.
RESPONSE: We are using the MAS data which contains the 8.5 um channel. This is the only data set that we have found with this channel for application to the polar regions. The 8.5 um channel is used in the nighttime algorithm.
The neural network has been refined considerably. It is being optimized for polar data. An inter-comparison of methods showed that the fuzzy expert system was not competitive in terms of cpu requirements, and it was slightly less accurate. High solar zenith angle test cases have been made, at angles greater than 80 degrees.

5. What are the strengths and weaknesses or the approach adopted for product validation?
REVIEWER: Data from the polar FIRE III IFO will be useful for algorithm testing and validation, but it certainly will be limited to those situations which occurred during the IFO, which may not be representative of the wide variety of cloud and surface types. Otherwise, only human experts looking at satellite imagery will be used to validate the algorithm's analyses. As previously men-tioned, the human expert may have trouble identifying nighttime scenes from IR data alone. Sur-face observations aren't terribly reliable in the polar regions in winter, so there may be no solution.

47

RESPONSE: The FIRE III IFO has passed. We plan to use an extensive set of surface sites for the validation effort, as listed in our STAR requests. Some of these sites have lidar and cloud radar instruments, which should be far more reliable.

6. Is the planned schedule for generating the product (i.e., at launch/post-launch) appropriate?

REVIEWER: No schedule was given. I think Welch, Penaloza and Feind has a lot of work to do before the algorithm is ready to be used, but there should be plenty of time to get the most important things done.

RESPONSE: The algorithm has been extensively tested for the last three years with different data sets. It has matured and become much more robust. However, no exact matches in channels is possible, so that the algorithm will have to be adjusted after launch to obtain the most accurate results. Nevertheless, accuracy of 85% to 90% should be obtained during the daytime and 75% to 80% at night with the newer algorithm using MAS data. Accuracy of about 95% should be obtained during the daytime using the actual ASTER data, and 85% at night.