An Iterative Approach to Multisensor Sea Ice Classification

Quinn P. Remund, David G. Long, and Mark R. Drinkwater

Abstract— **Characterizing the variability in sea ice in the polar regions is fundamental to an understanding of global climate and the geophysical processes which govern climate changes. Sea ice can be grouped into a number of general classes with different characteristics. Multisensor data from NSCAT, ERS-2, and SSM/I are reconstructed into enhanced resolution imagery for use in ice type classification. The resulting 12-dimensional data set is linearly transformed through principal component analysis to reduce data dimensionality and noise levels. An iterative statistical data segmentation algorithm is developed using maximum likelihood and maximum *a posteriori* techniques. For a given ice type, the conditional probability distributions of observed vectors are assumed to be Gaussian. The cluster centroids, covariance matrices, and *a priori* distributions are estimated from the classification of a previous temporal image set. An initial classification is produced using centroid training data and a weighted nearest-neighbor classifier. Though validation is limited, the algorithm results in an ice classification which is judged to be superior to a conventional k-means approach.**

I. Introduction

BECAUSE the polar regions of the earth play a critical role in the global climate, the remote sensing community has had a keen interest in the variability of polar sea ice characteristics. Sea ice influences heat transfer between the warmer ocean and cooler atmosphere. In this process, ice thickness and density are particularly important. The extent and surface characteristics of sea ice affect the global radiation budget by regulating the amount of solar radiation reflected back out into space. In addition, these regions influence the planetary water exchange cycle as well as local biota distributions. Sea ice is also considered a sensitive indicator of long term global climate change [1]. Hence, an accurate knowledge of important surface characteristics of sea ice is a valuable tool in acquiring an understanding of these geophysical processes.

Microwave remote sensing provides an excellent means for monitoring polar sea ice. Both active and passive microwave signatures are much less sensitive to atmospheric distortions than measurements collected at optical frequencies. This is particularly true in the Arctic and Antarctic where extensive cloud cover is common. Many research studies have shown that microwave signatures of sea ice are sensitive to surface parameters [2]. In addition, microwave sensors do not require solar illumination to collect measurements of the surface. However, these benefits often come at the expense of spatial resolution.

Fundamental sea ice characteristics can be grouped into a number of general sea ice classes or types. Various studies have been pursued to classify ice type from observed microwave signatures. A single-band classifier using 33.6 GHz passive high-resolution aerial measurements was used on Beaufort Sea data [3]. Kwok et al. developed a method for classifying high-resolution ERS-1 SAR imagery using ancillary data from meteorological databases [4]. Rignot and Drinkwater also performed a MAP classification on polarimetric airborne SAR data and compared results to high resolution passive microwave data [5]. Hara et al. proposed an unsupervised polarimetric SAR multi/first year ice classifier using a neural network followed by iterative maximum likelihood classification [6]. The primary strengths of these approaches lie in the high spatial resolution capability of the instruments. Consequently, image pixels are much less likely to contain a mixture of ice types. Lower resolution techniques have also been proposed. Wensnahan et al. proposed a classification method using passive radiometer data [7] to estimate the concentrations of first-year, multi-year, and thin ice in the Arctic. In [8], a classifier was developed that uses single channel 5.3 GHz ERS scatterometer data. Finally, a neural network classifier for sea ice type is given in [9]. These studies are representative of the different work that has been done in microwave sea ice classification.

This paper presents a multisensor sea ice classification approach which uses multispectral, dual-polarization data collected from both active and passive spaceborne instruments for the segmentation of Antarctic data. In section 2, important background information is given describing the instruments from which data is collected, the image reconstruction methodology, the ice extent mapping techniques, and the basic ice type signatures. Section 3 introduces the multivariate analysis techniques fundamental to the preprocessing stage of the algorithm including data fusion and principal component analysis. The sea ice classification algorithm is described in detail in section 4. A brief derivation of statistical measures as well as convergence metrics are given. Results of the application of the algorithm to actual data are presented in section 5. The final section contains the conclusions.

II. Background

The proposed ice classification scheme uses data from several different spaceborne instruments. This section provides a brief background of each of the data collecting instruments and the corresponding ice type signatures. In addition, the methodology for image reconstruction is described. Finally, an ice masking algorithm that removes open ocean pixels is summarized.

A. Spaceborne Microwave Sensors

Data from three different sensors are used in the classification approach which follows. The sensors are chosen for their temporal simultaneity of measurement collection during the target time frame of September-October 1996. In addition, all of the selected instruments have large-scale coverage capability. The

Q. P. Remund and D. G. Long are with the Microwave Earth Remote Sensing Laboratory, Brigham Young University, Provo, UT, USA. E-Mail: remundq@ee.byu.edu, long@ee.byu.edu.

M. R. Drinkwater is with the Polar Oceanography Group, Jet Propulsion Laboratory, Pasadena, CA, USA. E-mail: m.drinkwater@jpl.nasa.gov

first data set comes from the NASA scatterometer (NSCAT) that flew from August 1996 to June 1997. NSCAT is a dual-polarization, dual-swath, fan-beam scatterometer that collects measurements at multiple azimuth and incidence angles [10]. Doppler filtering is used to segment each fan-beam footprint into cells with resolution on the order of 25 km. Operating at Ku-Band (14 GHz), NSCAT has 6 vv-pol and 2 hh-pol beams. While NSCAT was originally designed to measure near surface wind vectors over the ocean surface, it has also found great utility in land and ice studies such as in [11] and [12].

The second sensor is the active microwave instrument (AMI) aboard the European remote sensing satellite (ERS-2) [13]. One mode of operation of the AMI is the wind scatterometer mode which measures the vv-pol normalized radar cross section (σ^o) at several azimuth and incidence angles. While similar to NSCAT in its fan beam configuration, the ERS-2 scatterometer has only a single side-looking swath. The instrument uses range filtering to resolve the measurements to an effective resolution of about 50 km.

Finally, passive radiometer data is used in concert with the active scatterometer data to produce a merged data set. The special sensor microwave imager (SSM/I) aboard the Defense Meteorological Satellite Program series of satellites is a total-power, seven channel, four frequency radiometer [14]. The channels are h- and v-pol at 19.35, 37.0, and 85.5 GHz and v-pol at 22.235 GHz. Brightness temperature (T_B) measurements are collected from each channel. The 3 dB antenna footprints range from about 15-70 km in the along-track direction and 13-43 km in the cross-track direction. The 3 dB antenna footprints, which are different for each frequency, generally have an elliptical shape on the surface of the earth due to the elevation angle of the radiometer [15].

B. Image Reconstruction

While the inherent resolutions of the various instruments are sufficient for the study of large-scale phenomena such as surface winds or atmospheric parameters, they can be too low for use in some studies. In an effort to ameliorate this problem and to place the data on compatible grids, the scatterometer image reconstruction (SIR) algorithm is used to enhance the spatial resolution of both scatterometer and radiometer data[16], [17]. SIR is an iterative block multiplicative algebraic reconstruction technique that increases the resolution of reconstructed imagery through the use of multiple passes of the satellite. SIR utilizes the increased sampling, though irregular in geometry and sample spacing, to raise the side lobes of the antenna pattern in the spatial frequency domain and thus increase resolution.

For scatterometers, σ^o (in dB) has a nearly linear incidence angle dependence over a limited range of incidence angles, $\theta \in [20°, 55°]$ given by

$$\sigma^\circ(dB) = A + B(\theta - 40°) \tag{1}$$

where A is σ^o normalized to 40° incidence and B is the incidence angle dependence of σ^o. SIR creates images of both A and B for each scatterometer. NSCAT images are reconstructed on a 4.45 x 4.45 km grid with an effective resolution on the order of 8–10 km. For NSCAT, the SIR with filtering (SIRF) algorithm is used [16]. For ERS-2, the median filter is not used. ERS-2 images are generated on a 8.9 x 8.9 km grid with an effective resolution of 20–25 km.

A univariate version of SIR can be applied to radiometer data such as SSM/I [17]. The lower side lobes of SSM/I make resolution enhancement more difficult. However, a clear improvement in resolution is observed in the reconstructed imagery. That is, surface features are more clearly defined in SIR imagery than in nonenhanced images on the same grid. SSM/I brightness temperature (T_B) SIR images are reconstructed on a 8.9 x 8.9 km grid for all channels except 85V and 85H which have an improved resolution and thus, a pixel spacing of 4.45 x 4.45 km.

All images are generated using six days of data with three days of overlap in consecutive images. While NSCAT v-pol and SSM/I can achieve full coverage of the Antarctic ice pack in much less time, ERS-2 and NSCAT h-pol require the full six days. For consistency in pixel spacing between the different images, the 8.9 km images are interpolated to the 4.45 km grid. All parameter images are used in the classification except for the ERS-2 B images which have relatively high noise levels and are thus discarded. The final merged data set consists of 12-dimensions with three A, two B, and seven T_B images. Sample images of all 12 types are shown in Figure 1 for 1996 JD 261–266. Figure 2 shows two zoomed versions of these images which illustrate the Weddell Sea quadrant of the NSCAT v-pol A and the SSM/I v-pol 37 GHz images. The imagery shown has been masked with an ice extent mapping algorithm discussed in the next section. A significant amount of detail is evident in the sea ice regime of these images. This is exploited in the proposed classification algorithm.

SIR enables comparison of sensors on compatible grids with similar resolution. While the SIR algorithm increases the resolution of the reconstructed imagery of a particular instrument, ice motion during the imaging period is a concern. In a six day period, sea ice can potentially move tens of kilometers. This represents several pixels in the reconstructed imagery. As a result, the classification results presented below are treated as average behavior during the imaging interval for each pixel.

C. Ice Masking

Open ocean pixels in the reconstructed imagery are masked out for two reasons. First, the sea ice classification algorithm presented below uses statistical preprocessing techniques which take advantage of the covariance structure of the data to reduce the dimensionality of the data space. Since ocean pixels have typically high covariance values in all of the active and passive signatures, undue weight would be given to ocean pixels in the new data space effectively reducing the classification potential. Second, a significant number of the image pixels are open ocean and the removal of these pixels reduces the size of the classification data set.

The ice extent mapping algorithm used in this study is described in [18]. The technique uses the NSCAT polarization ratio (A_v/A_h) as well as the NSCAT v-pol incidence angle dependence of σ^o (B_v) to discriminate between sea ice and ocean pixels. Linear and quadratic segmentation techniques are applied resulting in an estimate of sea ice extent. Since wind induced roughness of the sea surface causes ambiguities in the discrimination, a third parameter, the σ^o estimate error standard

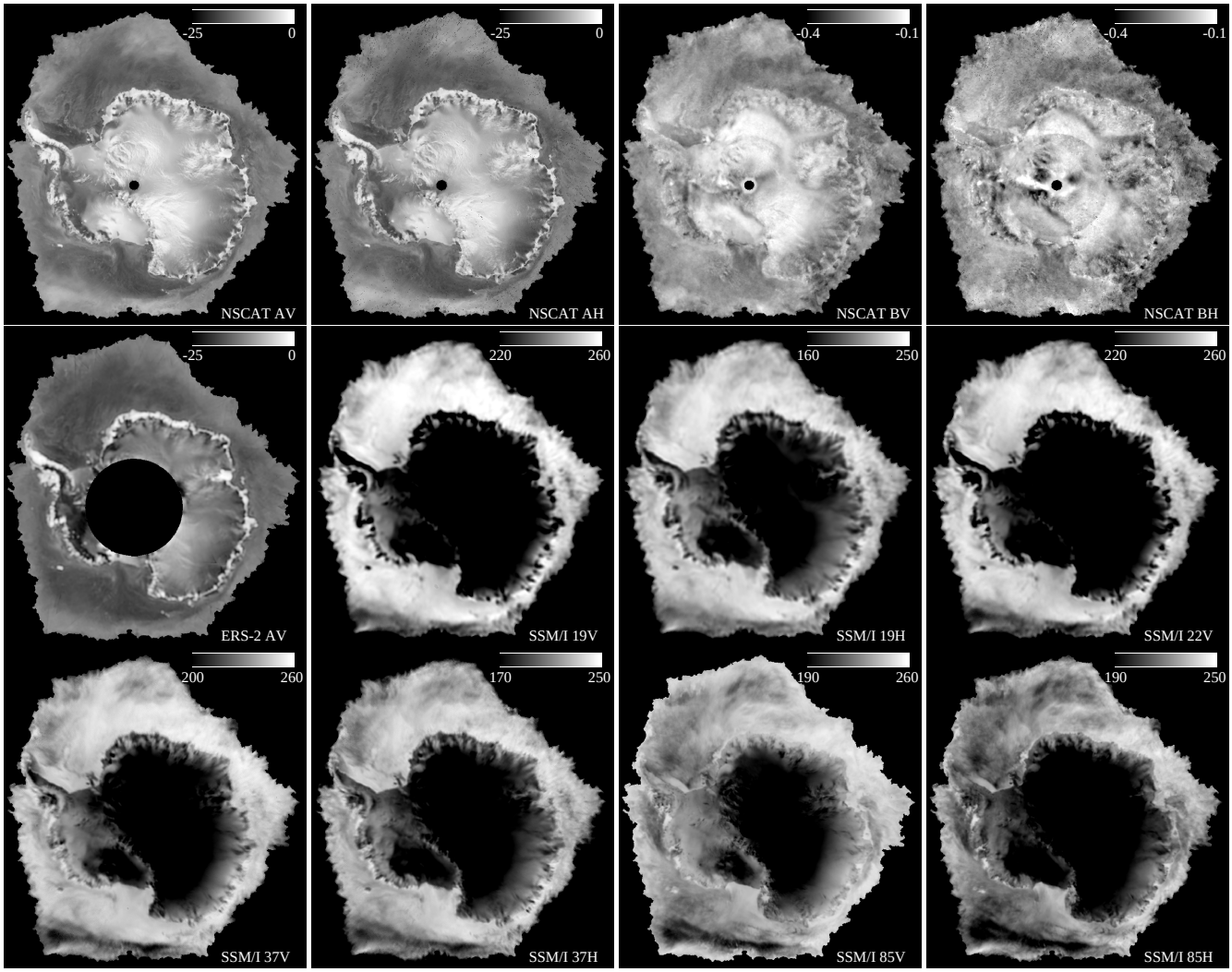

Fig. 1. Image set for 1996 JD 261-266.

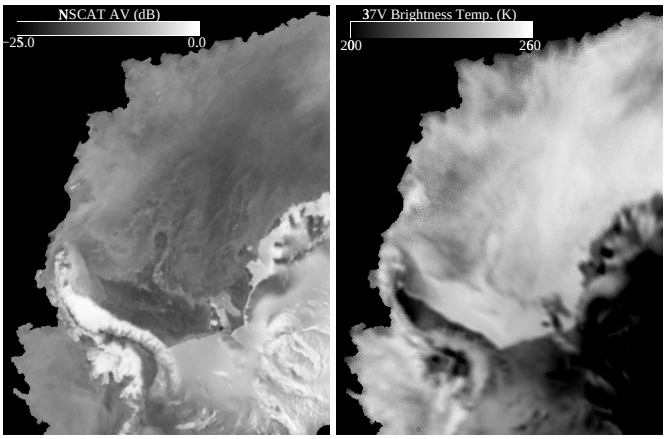

Fig. 2. Sample Weddell Sea quadrant images of NSCAT A_v (left) and SSM/I 37V (right).

deviation, is introduced. This metric is sensitive to temporal and azimuthal variations during the imaging period and is consequently quite large in regions of high winds. Through an empirically derived threshold on this parameter, many of the initial sea ice extent mapping errors are eliminated. When compared with NASA-Team SSM/I-derived ice concentration maps, the NSCAT ice extent edge most closely corresponds with the 30% ice concentration edge. A similar study conducted by Yueh and Kwok found that the NSCAT ice edge was close to the NASA-Team 25% edge in the Arctic [19]. While very similar, the discrepancies may be attributed to differences in the characteristics of Antarctic and Arctic sea ice or the analysis techniques. The NSCAT algorithm is used for this study because it provides an ice edge matching the image resolutions.

Figure 3 shows the NSCAT and NASA-Team ice edges plotted over a C-band hh-pol RADARSAT SAR image. The NASA-Team edge was generated by averaging six days of data to be consistent with the NSCAT imaging interval. While the RADARSAT image is not calibrated and some obvious geolocation errors exist, a clearly defined ice edge is observed. Although both edges are relatively good estimates of sea ice extent, the NSCAT curve most closely follows the actual ice edge in this particular case. For the purposes of this study, the NSCAT method was used to ice mask all parameter imagery during the preprocessing stage of the algorithm.

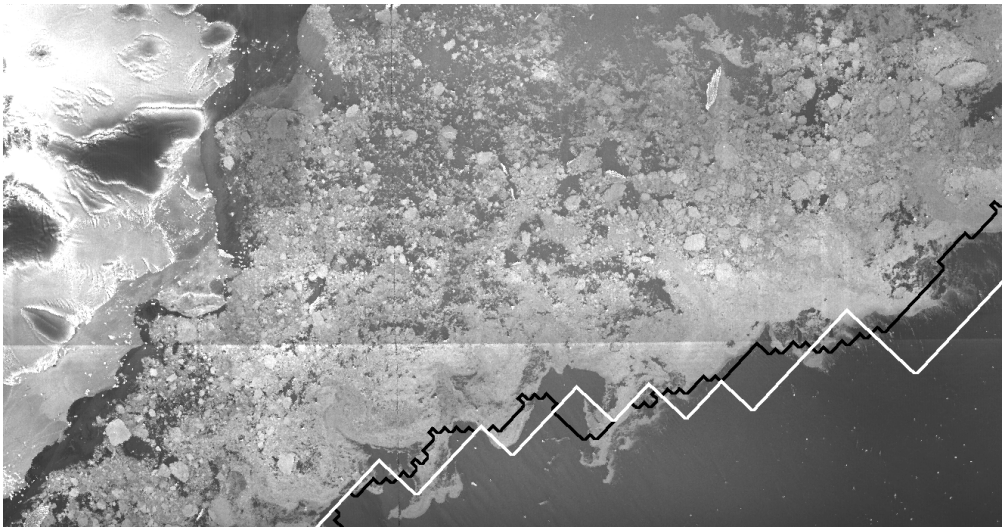

Fig. 3. Ice edge comparison with NSCAT edge (black) and NASA-Team 30% ice edge (white) plotted over a RADARSAT SAR mosaic (uncalibrated) of a sea ice region [Radarsat data © Canadian Space Agency, 1996]. A portion of Saunder's Coast is evident on the left of the image on 20 February 1997.

D. Ice Type Signatures

Data collected by NSCAT, ERS-2, and SSM/I are used to segment the data into six general ice types or classes. While the following discussion is based on the general behavior of scattering and emission from sea ice, *in situ* measurement averages for the various ice classes can be found in [8] for C-Band scatterometer Antarctic data and [9] for Arctic SSM/I data.

The first ice type to be considered is smooth first-year (SFY) ice. This class represents relatively young ice which has not been roughened by the differential motion and deformation of the ice pack. Ranging in thickness from 10 cm to 1 m, smooth first-year ice is highly saline with a high density of brine pockets caught within the ice crystal lattice. The high salinity causes this ice type to be very lossy and thus dominated by surface scattering and emission at virtually all frequencies used in the study. The active signatures exhibit low A and B values due to the strong incidence angle dependence of smooth surface scattering from level ice. T_B measurements are expected to be relatively high.

Like smooth first-year ice, rough first-year (RFY) ice is very saline and lossy. Surface scattering and emission dominate the signatures. Motion within the ice pack causes extensive roughening of this ice type. In general, the rough surface scattering cause A values to be higher than for smooth ice types and B values to rise (ie, have less incidence angle dependence) [20], [8]. While passive signatures are less sensitive to the difference in RFY and SFY ice classes, T_B values are radiometrically cooler for RFY when compared to SFY ice.

Perennial (PER) ice is another important Antarctic ice type. While multi-year ice is common in the Arctic, less Antarctic sea ice survives more than one summer's melt since the Antarctic continent limits the southern extent. Regardless, a small amount of perennial ice can be found and is included in the classification. Over time, brine drainage results in much lower salinity and hence lower electromagnetic absorption in this ice type. This leads to greater penetration depths and volume contribution to scattering and emission. A and B values are typically higher than those for RFY ice while T_B measurements are lower.

Another sea ice type to be considered in the classification is the iceberg class (IB) consisting of large floating plates of fresh water ice that have calved or broken off from an ice shelf. In the absence of surface melt, this ice class has very low loss resulting in a large contribution from volume scattering and emission especially at lower microwave frequencies Furthermore, these targets may extend vertically out of the water several tens of meters thereby acting as strong reflectors. As a result, A and B are very high and T_B values are very low compared to other ice classes. The volumetric scattering contribution also causes a depolarization resulting in similar response for both v- and h-pol measurements.

Pancake ice is also included in the classification effort because of its unique appearance over extremely large areas of the marginal ice zone during winter ice growth. This ice regime is normally found in the outer portions of the ice pack where wave action aggregates and deforms newly growing frazil ice into small floes called pancakes. The high roughness of this type results in a signature that is very similar to perennial ice in both active and passive signatures [8], [21].

The final ice type is the marginal ice zone (MIZ). This dynamic region of the ice pack consists of mixtures of ice and open water. The open water contribution drives T_B down. Wind roughening of the ocean surface in these regions causes A and B values to often be confused with other ice types. While pancake ice is typically found in the marginal ice zone, the two classes are considered separately in this study in order to discriminate between regions of low ice concentration in the MIZ and high concentration pancake regimes existing only under ice growth conditions.

One of the complicating factors in sea ice classification is the seasonal dependence of the ice type microwave signatures. The signatures are most distinct during the Austral winter months when ice types exhibit negligible surface melt. Hence, the differences due to subsurface contributions are strong. When tem-

peratures rise and water content increases (typically during the mid-November to March Austral summer months), the scattering and emission from lower ice layers become increasingly masked and surface scattering mechanisms begin to dominate. This causes ice type clusters to drift in the 12-dimensional data space with some clusters merging together. For this reason, an effective classification technique must have the ability to adapt to changing signatures in order to maintain a maximal degree of accuracy.

III. Multivariate Data Analysis

As previously discussed, the classification data set consists of 12 parameters from which sea ice type is to be extracted. Additional preprocessing is performed on the data to maximize classification accuracy and minimize required computational effort. Since the parameters are measured in very different units, data fusion techniques are used to give equal weighting to all of the data. In an effort to reduce the computational complexity and the noise levels, principal component analysis is implemented.

A. Data Fusion

The 12-dimensional data space consists of three basic types of data with differing units. The first data type, A, is measured in dB with a typical range of -30.0 to 0.0 dB. The incidence angle dependence of σ^o, given by B, contains dB/deg values ranging from -0.4 to -0.1 dB/deg. The last data type is T_B measured in degrees Kelvin with sea ice values from 150 K to 290 K depending on frequency and polarization. Since each data type is quite different from the others, standardization is required to ensure that each data type is given appropriate weight in the classification. The standard approach is to shift and scale the data so that each of the 12 parameters have zero mean and unit variance. However, this may remove some ice class information that exists between the mean responses of parameters that are within the same data type. In an effort to preserve the ice class dependent biases that exist in each data type, the following standardization technique is applied for a particular observation x

$$x_s = \frac{(x - \mu_{type})}{\sigma_{type}} \tag{2}$$

where μ_{type} and σ_{type} are the collective mean and standard deviation of all the parameters belonging to a particular data type (e.g., A, B, or T_B data) and x_s is the new standardized parameter value. Hence, the three general data types, A, B, and T_B are transformed such that they have zero mean and unit variance though specific parameters (e.g., A_v, A_h, etc.) may not have these characteristics. The resulting data resides in a 12-dimensional unitless space in which each data type has similar range and variance.

B. Principal Component Analysis

The high dimensionality of the classification data set equates with significant computational requirements. To reduce the number of required parameters, principal component analysis (PCA) is implemented. PCA is a powerful data analysis tool that effectively rotates the data space by projecting each observation onto a new orthonormal basis [22]. The resulting basis vectors are chosen such that the first spans the direction of maximum variance in the data. Successive vectors are chosen to span the maximum variance not accounted for by previous vectors.

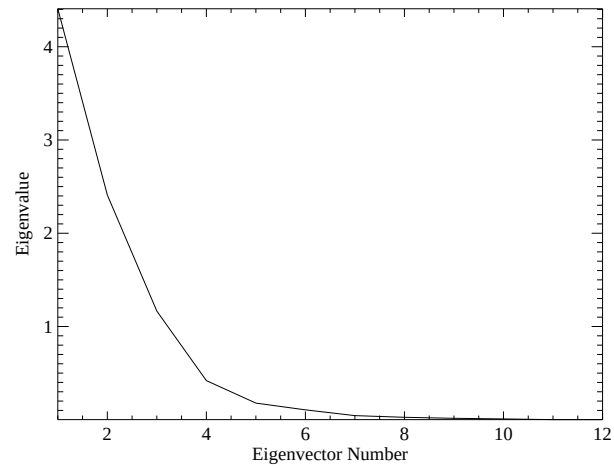

Fig. 4. Eigenvalue spectrum for the principal component data rotation during the imaging interval of 1996 JD 261-266. A large majority of the data variance is contained in the first few eigenvectors.

For the classification problem at hand, data vectors are composed of the 12 standardized values

$$\vec{y} = [y_1\, y_2 \ldots y_{12}]^T \tag{3}$$

where the y_i represent the standardized versions of the NSCAT, ERS-2, and SSM/I data values. PCA uses an eigenvalue/eigenvector decomposition of the data to construct the necessary orthonormal basis vectors. The eigenvalue/eigenvector equation is given by

$$K\Gamma = \Gamma\Lambda \tag{4}$$

where K is the 12 x 12 covariance matrix of the standardized data, Γ is a matrix with eigenvectors of K along the columns (which form a basis for the original 12-dimensional space), and Λ is a diagonal matrix with the eigenvalues of K along the diagonal (which represent the variances spanned by each eigenvector). Once these are obtained, a 12 x 1 data vector $\vec{y}$ containing standardized parameters is transformed through projection onto the new basis

$$\vec{z} = \Gamma^T \vec{y}. \tag{5}$$

The elements of $\vec{z}$ are called the principal component scores [22].

The analysis technique is used on land/ice masked imagery to produce 12 principal component images composed of a combination of information contained in the original parameters. The pixel values in individual PCA images represent coefficients of the eigenvector associated with that principal component score. The size of the corresponding eigenvalues determine the variance and informational content of each of the images. For example, the PCA transformation was performed for the microwave data set during the imaging interval 1996 JD 261-266. Figure 4

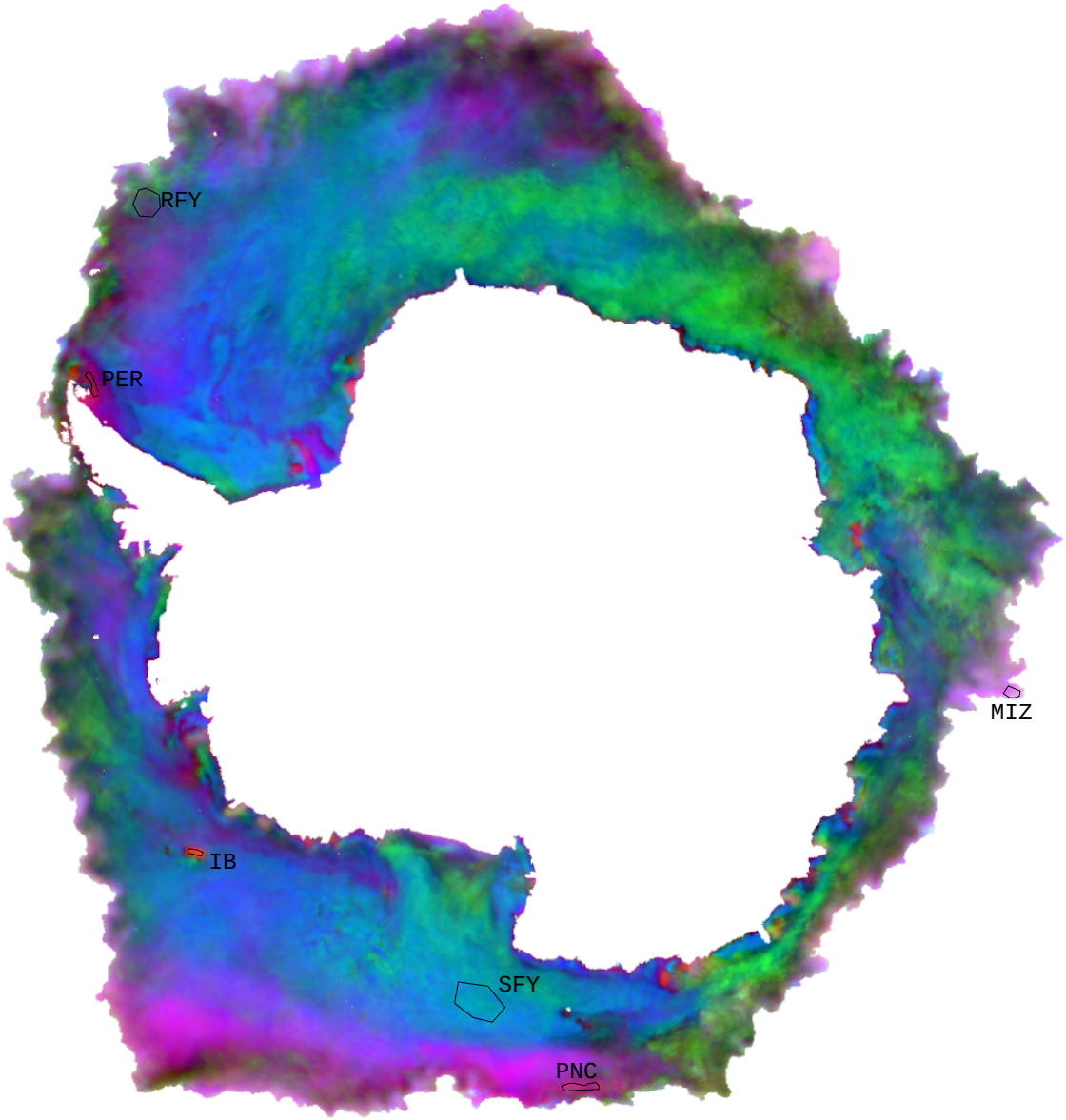

Fig. 5. RGB composite image of the first three principal components for 1996 JD 261-266. The red channel is the top principal component image, the green is second, and blue is third. This image is useful in evaluating ice type information contained in the top three PCA scores. The six training regions are also indicated.

illustrates the resulting eigenvalue spectrum. Clearly, a majority of the data variance is contained in the top few principal component images implying that lower PCA images can be neglected with minimal effect on the final classification. Wensnahan et al. suggest keeping only PCA parameters whose variance is much larger than measurement uncertainty (converted into principal component space) [7]. Such a choice of eigenvectors allows information to be separated from noise. Indeed, the lower principal component images used in this study appear very noisy with image reconstruction artifacts dominating the features. Hence, by ignoring these eigenvectors, we eliminate undesirable noise as well as reduce data dimensionality.

Another method for choosing principal component images is to keep the top N PCA transformed images that account for some predetermined percentage of the total variance in the data. For this classification project, the eigenvectors that span 90% of the variance are kept for use in the data segmentation. The 90% threshold was chosen to balance increased computational complexity and increased informational content when including additional individual components. That is, eigenvalues for the principal components beyond the first 90% are typically negligible in comparison. Only three PCA images must be retained in the case of the sample data set representing 1996 JD 261-266 Antarctic sea ice data. These principal component scores are shown in the composite RGB image of Figure 5 where the red channel contains the maximum variance principal component score, green contains the second highest variance, and blue contains the third highest. Features from all 12 original data value

images are evident in this figure. The training regions are also indicated and will be discussed later.

The composite image is composed of a variety of signatures in the sea-ice annulus around Antarctica. For example, pixels that contain known icebergs are bright orange in the image. Regions of old perennial ice are visible in locations where a residual amount of old ice have been observed in late winter, such as near to the north-eastern tip of the Antarctic peninsula [20], [23]. Signatures which appear close to that of perennial ice may be observed in other coastal zones. In the southern Weddell Sea, a patch of deformed old ice and fast ice may be found surrounding a number of grounded icebergs [8], with a resulting purple color. Further such regions of deformed ice may be found in the Amundsen and Eastern Ross Seas along the coast. Large portions of the central ice pack have blue and green hues. Blue regions of undeformed medium, snow-covered first year ice are found in the central Weddell and Ross Seas, while green appears more closely related to the younger regions of first-year ice. For instance, recently formed ice in the Ross Sea, just north of the Ross ice shelf, and relatively thinner, saline young ice formed around East Antarctica display larger areas of green hues. One factor that appears to confirm the relationship between principal component three and young ice is the appearance of green in known coastal polynya regions such as along the Ronne ice shelf front in the Southern Weddell Sea [8], and in the wake of the large drifting icebergs, such as those observed off the Terre Adelie Land coast, and those grounded off the Amery ice shelf [24].

Interesting mixtures of browns and cyans are observed predominantly at the outer ice margin. Bright cyan signatures appear to be extensive regions of pancake ice formation, as for instance in the region of maximum northern ice extent in the Amundsen Sea. Brown hues are more extensively found at the ice margin, and likely are associated with mixtures of deformed and wave fractured floes found in the marginal ice zone, together with mixtures of open water and ice signatures.

Although previous classification efforts have identified many of these primary cluster types [8], [25] in single channel datasets, the unique attribute of the top three principal components shown in Figure 5 is that they show mixtures of the primary end-members. Pure red may be thought of as the ice with the most typical volume-scattering signatures. This encompasses icebergs, old, thick snow-covered perennial ice, and fast ice - all with low salinity. In contrast, pure green appears to indicate the most different cluster of ice typifying signatures which have the greatest rate of incidence-angle signature change, and the least volume-scattering like signatures. Lastly, the pure blue appears to imply intermediate ice salinity, and the least amount of surface deformations.

PCA can be used not only to reduce the dimensionality and noise levels of the data, but to quantitatively assess the informational content of multisensor data. By observing the relative magnitudes of the elements of the first few eigenvectors, one can determine levels of informational content of the original parameters. An example is given in Figure 6 in which the coefficient magnitudes of the top three eigenvectors of the sample data are plotted. The first eigenvector gives very low weighting to the NSCAT *B* and SSM/I 85 GHz images while high weighting is given to the NSCAT *A* and SSM/I 19H and 37H images. The eigenvector plot can also be used to determine which parameters can be eliminated from the classification. For example, the first two eigenvectors have very low NSCAT B_v and B_h values indicating that the *B* data types do not contribute to the majority of data variance. We note also that the third principal component eigenvector has a much higher weight on B_h than B_v implying that one of the *B* parameters could be eliminated without major impact on the classification.

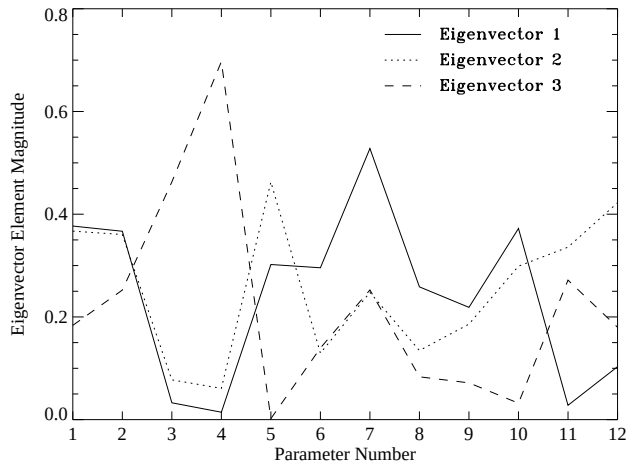

Fig. 6. Vector element magnitudes for the first three eigenvectors. Eigenvector 1 spans the most data variance, 2 spans the second highest variance, and 3 spans the third highest. The parameter numbers correspond with 1- NSCAT A_v, 2- NSCAT A_h, 3- NSCAT B_v, 4- NSCAT B_h, 5- ERS-2 A_v, 6- SSM/I 19V, 7- SSM/I 19H, 8- SSM/I 22V, 9- SSM/I 37V, 10- SSM/I 37H, 11- SSM/I 85V, and 12- SSM/I 85H.

IV. Classification Algorithm

Several techniques are available for classification of N-dimensional data sets. A nearest-neighbor approach is perhaps the simplest when centroids from training samples or electromagnetic models can be obtained. Iterative clustering algorithms such as k-means or ISODATA represent another methodology and search for natural clusters in the data. The task then remains to label the resulting clusters as different classes. In contrast, the proposed approach is a statistical classification scheme with the goal of maximizing the probability of correctly classifying sea ice type. This section presents the classification methodology through a development of an iterative maximum *a posteriori* algorithm.

A. Statistical Classification

The intrinsic value of statistical methods of classification stems from the ease of interpretation of results. That is, statistical classifiers attempt to maximize a probability measure given some level of knowledge of class distributions. Two primary branches have evolved in the field of statistical classification and estimation: maximum likelihood and Bayesian classification.

Maximum likelihood (ML) methods as applied to discrete classification problems such as the determination of sea ice type choose the solution that maximizes the conditional probability

of data vector observation over all possible sea ice types,

$$C_{ML} = argmax_c \; p(\vec{z}|C = c) \tag{6}$$

where C_{ML} is the chosen ice class, $\vec{z}$ is the principal component data vector, C is a discrete variable of different ice types, and $p(\vec{z}|C = c)$ is the probability of observing $\vec{z}$ given a particular ice type $C = c$. Hence, the maximum likelihood method can be implemented as long as the conditional distributions are known. Unfortunately, this is rarely the case. A weakness of this method lies in the fact that the occurrence of each sea ice type is effectively considered to be equal. Consequently, classes that occur infrequently such as icebergs are given equal weight in the data segmentation and may be chosen too often.

Bayesian methods represent another class of statistical approaches. This scheme requires the definition of a loss function which assigns a penalty for misclassifications. The Bayes solution then minimizes the expected loss which is also called the Bayes risk. Under a uniform loss function this reduces to a maximum *a posteriori* (MAP) classifier. The MAP technique treats the ice type C as a random variable and maximizes the probability of ice type given the observation vector $\vec{z}$

$$C_{MAP} = argmax_c \; p(C|\vec{z}) = argmax_c \; \frac{p(\vec{z}|C)p(C)}{p(\vec{z})} \tag{7}$$

where $p(C)$ is the *a priori* distribution. Since $p(\vec{z})$ is fixed for a particular observation, this reduces to

$$C_{MAP} = argmax_c \; p(\vec{z}|C)p(C). \tag{8}$$

MAP classification has an advantage over maximum likelihood techniques in that the probability of each class is included in the derivation ensuring that less likely ice types appear less frequently in the final classification. However, the *a priori* distribution and the conditional distributions are required.

Under a Gaussian assumption, the conditional distributions are

$$p(\vec{z}|C) = \frac{1}{(2\pi)^{n/2}|K_c|^{1/2}} e^{-\frac{1}{2}(\vec{z}-\vec{\mu}_c)^T K_c^{-1}(\vec{z}-\vec{\mu}_c)} \tag{9}$$

where $\vec{\mu}_c$ is the mean vector and K_c is the covariance matrix of ice type c, respectively. Hence, the statistical structure of the data for each ice class is completely determined by the mean vectors and covariance matrices. Even if the Gaussian assumption is not entirely correct, it is considered to be an improvement over the simple equal and isotropic distribution assumption inherent to a nearest-neighbor classifier since the Gaussian model can account for covariance between separate principal component scores. This allows the classifier to use cluster shapes in addition to the centroids to segment the data. For the maximum likelihood development, the maximization of Eq. (9) can be simplified. After taking the natural log (a monotonic function) and a little mathematical manipulation we obtain

$$\begin{aligned} C_{ML} &= argmax_c \; p(\vec{z}|C) \\ &= argmax_c \; [-(log|K_c| + (\vec{z}-\vec{\mu}_c)^T K_c^{-1}(\vec{z}-\vec{\mu}_c))] \end{aligned} \tag{10}$$

which is equivalent to

$$C_{ML} = argmin_c \; [log|K_c| + (\vec{z}-\vec{\mu}_c)^T K_c^{-1}(\vec{z}-\vec{\mu}_c)]. \tag{11}$$

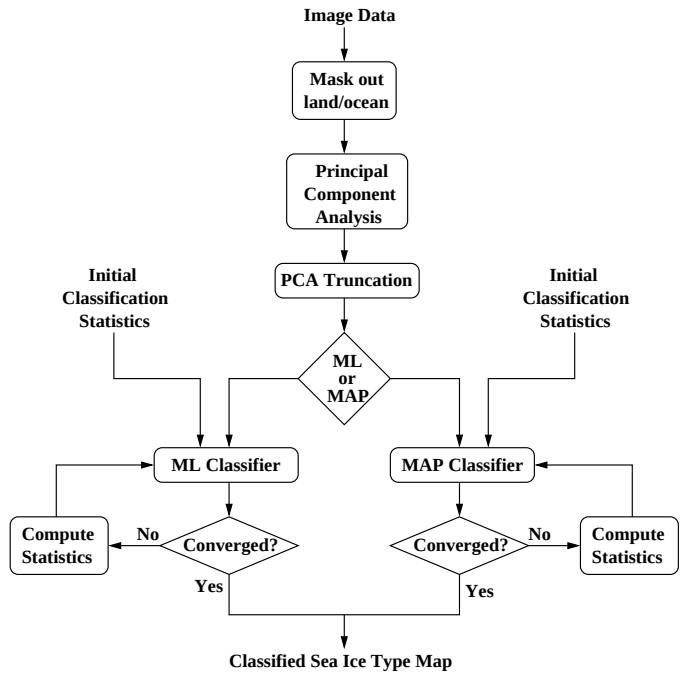

Fig. 7. Flowchart depicting the iterative ice classification algorithm for both ML and MAP methods.

We note that the second term in Eq. (11) is the Mahalanobis distance commonly used in Gaussian classification problems [26]. Thus, the ML classification can be interpreted as choosing the class centroid that minimizes a modified Mahalanobis distance.

A similar development applied to the MAP equations yields

$$\begin{aligned} C_{MAP} &= argmax_c \; p(\vec{z}|C)p(C) \\ &= argmax_c \; [-\frac{1}{2}(log|K_c| + (\vec{z}-\vec{\mu}_c)^T K_c^{-1}(\vec{z}-\vec{\mu}_c)) \\ &\quad + log(p(C))]. \end{aligned} \tag{12}$$

Both ML and MAP methods are separately used and compared in the sea ice classification given below.

B. Iterative Approach

In order to fully implement the ML and MAP techniques, the mean vectors $\vec{\mu}_c$ and covariance matrices K_c of the individual ice type clusters are required along with the *a priori* distribution $p(C)$. While a rough estimate of the cluster centroids can be generated from small homogeneous training regions, it is more difficult to obtain reasonable estimates of the K_c matrices. However, estimates can be obtained through an iterative procedure, assuming that the statistical measures converge to the correct values.

Figure 7 illustrates the complete process for the classification of a time series of image data. The initial SIR-derived images are first masked to remove all land and ocean pixels using the ice extent mapping procedure defined in an earlier section. The PCA linear transformation is then performed to rotate the coordinate space into ordered maximum variance axes. Next, the resulting 12-dimensional principal component space is truncated by choosing the top N eigenvectors that span 90% of the data variance.

TABLE I

TRAINING REGION SIGNATURES USED IN THE INITIAL NEAREST-NEIGHBOR CLASSIFICATION FOR 1996 JD 261-266.

	IB	PER	RFY	SFY	PNC	MIZ
NSCAT A_v	-3.19	-5.72	-11.02	-17.35	-6.91	-12.29
NSCAT A_h	-3.77	-5.79	-11.01	-17.34	-6.51	-13.65
NSCAT B_v	-0.20	-0.14	-0.23	-0.25	-0.20	-0.22
NSCAT B_h	-0.23	-0.14	-0.21	-0.24	-0.18	-0.32
ERS-2 A_v	-4.46	-9.25	-14.80	-17.96	-11.19	-13.58
19V T_B	224.6	235.7	249.2	256.6	233.4	191.4
19H T_B	187.5	219.1	218.7	241.4	199.3	126.1
22V T_B	225.5	233.2	247.5	254.2	232.4	199.2
37V T_B	231.2	224.1	242.2	249.6	221.7	213.7
37H T_B	204.0	209.6	218.3	237.2	199.0	160.3
85V T_B	236.9	228.2	239.6	234.3	223.3	240.8
85H T_B	219.5	218.2	224.3	224.2	214.2	206.2

After the preprocessing, an iterative maximum likelihood or maximum *a posteriori* classifier is implemented. The first iteration uses the $\vec{\mu}_c$, K_c, and $p(C)$ statistical measures computed from the classification of the previous image set. Thus, the preceding classification is treated as a training set to obtain initial sea ice type cluster centroids, covariance matrices, and the *a priori* distribution. Due to the seasonal nature of cluster characteristics, the approximated values are likely erroneous. However, they represent a good initial starting point for the iterative procedure. After the first iteration, the statistical measures are updated using the current classification. These are then used in a new classification. The process iterates until predefined convergence criteria are met. The result is a classified image which maps the spatial extent of each sea ice type.

One important issue regarding the MAP algorithm is the behavior of the $p(C)$ estimate during the iterations. It is conceivable that an element of this distribution could go to zero if the corresponding ice class becomes very scarce. If the ice type later becomes more abundant, the zero probability from a previous iteration precludes any pixels from being classified as this ice type. While this phenomenon is not observed in any of the realizations in this study, the problem can be solved by setting some very low value as a lower limit on values of $p(C)$ preventing any of them from becoming zero.

C. Convergence Metrics

Two metrics are used to determine algorithm convergence. Since the Gaussian clusters are completely defined by the centroid vectors and covariance matrices, appropriate norms are used to obtain scalar measures of individual cluster behavior as a function of iteration. The Euclidean norm is used to measure the behavior of the cluster centroid vectors. The matrix spectral norm of each covariance matrix K_c is computed as a measure of the overall variance structure of each cluster. The spectral norm is equivalent to the square root of the maximum eigenvalue of $K_c^T K_c$. Convergence of both metrics for a particular cluster is a good indication that the cluster remains unchanged from one iteration to the next.

D. Algorithm Initialization

The algorithm described above is a recursive method that uses the classification result from the previous imaging interval to compute the present sea ice type map. In order to obtain an initial classification result to start the process, the following procedure is used. Cluster centroid vectors are estimated from small homogeneous training regions derived from a basic knowledge of sea ice type spatial behavior and expected microwave signatures. For the ML classifier a simple nearest-neighbor (minimum distance) classification yields the needed initial classification result. For the MAP method, the data is segmented with a weighted nearest-neighbor technique in which the distances to each cluster are inversely weighted by an initial estimate of $p(C)$. While an accurate estimate of the *a priori* distribution is difficult to produce, an educated estimate can be made through a knowledge of sea ice type population in Antarctica. For example, a large majority of the Antarctic ice pack consists of various types of first-year ice. Other classes are much less prevalent. For either ML or MAP, the nearest-neighbor solution is used to compute the necessary statistics for the classifier and initiate the iterative algorithm.

Simulations of the algorithm for both the ML and MAP techniques are performed. The simulation data consists of four different two-dimensional Gaussian distributions with different mean vectors, covariance matrices, and cardinalities. The distributions are chosen to have significant overlap to increase the classification difficulty. Simulation results indicate that the iterative algorithms converge to solutions that are very close to the actual ML or MAP solutions given two conditions. First, the individual cluster centroids must be relatively close to the actual centroids. In the simulations, this means that the centroid estimate merely has to be closer to its actual centroid than any of the others. Second, for MAP classification, the initial distribution estimate of $p(C)$ must be a reasonable estimate of the actual *a priori* distribution.

V. RESULTS

The iterative algorithms are applied to the classification of Antarctic data during consecutive imaging periods in Septem-

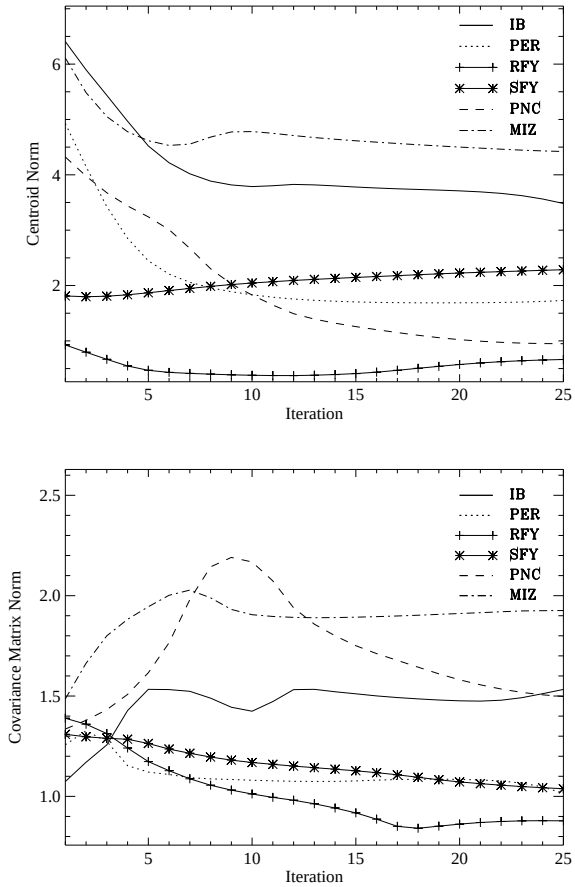

Fig. 8. Maximum likelihood classification cluster convergence metrics. (Top) The Euclidean norms of each ice type cluster as a function of iteration. (Bottom) The spectral norms of the covariance matrices.

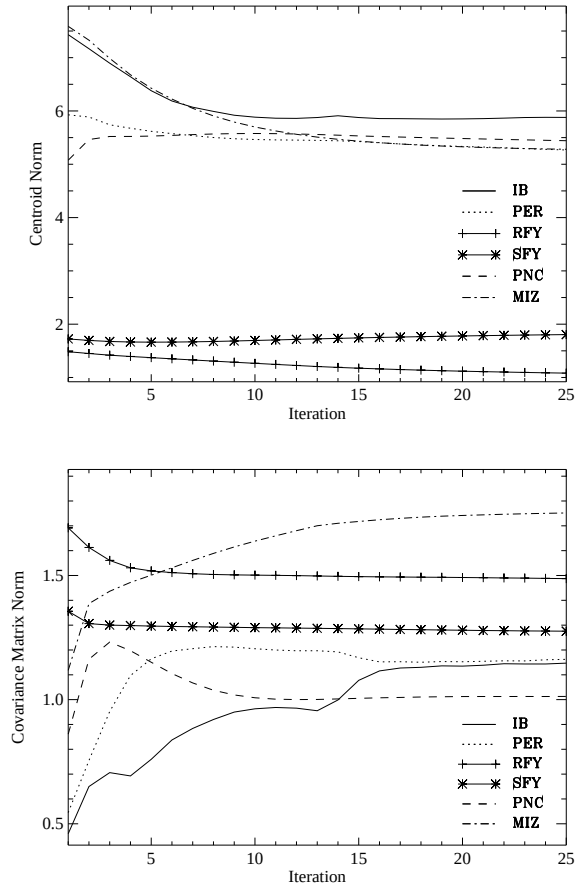

Fig. 9. Maximum *a posteriori* classification cluster convergence metrics. (Top) The Euclidean norms of each ice type cluster as a function of iteration. (Bottom) The spectral norms of the covariance matrices.

ber and October of 1996. The algorithm is initiated with multisensor data from JD 261-266. As noted in the previous section, the nearest-neighbor segmentation is required for the first image of the time series. As indicated in Figure 5, small homogeneous training regions are defined through a knowledge of sea ice dynamics and microwave signatures. Table I contains the cluster centroids obtained from these regions. In addition to the centroids, the MAP algorithm requires an initial estimate of the *a priori* distribution. For the 1996 JD 261-266 image classification we use $p(C) = [p(IB),\ p(PER),\ p(RFY),\ p(SFY),\ p(PNC),\ p(MIZ)] = [0.01,\ 0.02,\ 0.40,\ 0.45,\ 0.07,\ 0.05]$.

Figures 8-9 show the convergence metrics as a function of iteration for the ML and MAP classifications, respectively. After about 25 iterations, all metrics have converged relatively well. Some minimal centroid drift is still evident in the centroid norm trends. Most of the ML centroid norms shift significantly (and erroneously) during the iterations. On the other hand, only two of the MAP centroid norms move significantly indicating that the original centroids are reasonable estimates of the true values. Figure 10 illustrates the centroid drifts in the plane of the top two principal components for the MAP implementation using data for 1996 JD 261-266. The starting points for each centroid are denoted with triangles while the final centroid locations are represented by squares. The effect of the algorithm's iterative nature is evident as each of the points move varying amounts. Isoprobability elliptical contours have also been plotted surrounding each centroid convergence point using the sample covariance matrices from the final classification. The major and minor axes for each ellipse are 2-σ wide. This figure not only shows iterative migration of centroids, but gives a feel for how much separation exists between different ice classes in the plane of the top two principal components. Table II contains the final MAP cluster centroids in the normal parameter space. Table II can be compared with Table I for an indication of how much the initial training signatures were modified by the algorithm.

Figure 11 depicts the final ML and MAP sea ice type images. Since the ML image has large regions classified as icebergs and perennial ice, we conclude that the ML algorithm performs poorly. The primary source of the error is the ML assumption that all sea ice types are equally likely. This causes clusters that should have low cardinality to grow to sizes similar to more common ice types. This effect is responsible for the undesirable centroid drift discussed above.

In contrast, the MAP result exhibits a much more reasonable

TABLE II

ICE TYPE CENTROID SIGNATURES AFTER 25 ITERATIONS OF THE MAP ALGORITHM FOR 1996 JD 261-266.

	IB	PER	RFY	SFY	PNC	MIZ
NSCAT A_v	-5.35	-6.63	-11.66	-15.17	-7.85	-11.66
NSCAT A_h	-5.85	-6.67	-11.77	-15.27	-7.41	-12.39
NSCAT B_v	-0.201	-0.154	-0.224	-0.229	-0.183	-0.257
NSCAT B_h	-0.218	-0.152	-0.221	-0.225	-0.164	-0.286
ERS-2 A_v	-6.69	-9.47	-13.77	-16.72	-12.18	-12.99
19V T_B	230.3	239.6	246.9	256.2	232.0	213.8
19H T_B	195.8	219.6	217.5	235.8	196.5	160.8
22V T_B	230.6	237.9	245.9	253.9	229.1	219.2
37V T_B	234.1	231.2	241.5	248.3	215.4	226.6
37H T_B	208.2	214.4	219.0	230.8	192.6	184.6
85V T_B	238.4	229.0	236.8	234.5	211.5	243.3
85H T_B	220.5	217.3	223.4	222.3	200.7	215.8

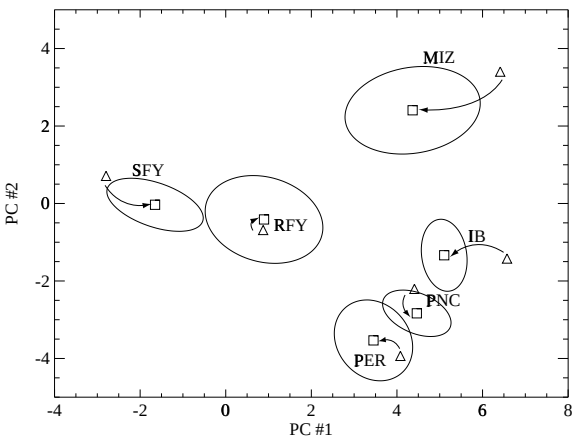

Fig. 10. Centroid locations in the plane of the top two principal components for the MAP classification of 1996 JD 261-266 data. The initial centroids (triangles) as well as the final converged centroid points (squares) are shown. Also plotted are the isoprobability contours according to the sample covariance matrices obtained from the final classification. The ellipses are 2-σ wide.

TABLE III

MAHALANOBIS DISTANCES BETWEEN CENTROIDS FOR 1996 JD 261-266 PROVIDING A MEASURE OF DISSIMILARITY BETWEEN DIFFERENT CLUSTERS IN THE CLASSIFICATION. THE COVARIANCE MATRIX OF THE REFERENCE CENTROID IS USED IN EACH COMPUTATION.

		Test Centroid					
		IB	PER	RFY	SFY	PNC	MIZ
Ref. Centroid	IB	0.0	24.5	25.0	59.1	70.8	111.0
	PER	15.5	0.0	79.1	138.3	46.3	350.2
	RFY	27.0	17.6	0.0	8.0	11.5	21.5
	SFY	119.2	67.0	17.0	0.0	87.9	180.3
	PNC	23.0	12.9	37.9	120.5	0.0	56.6
	MIZ	23.3	16.5	12.6	44.6	11.6	0.0

spatial distribution. Several icebergs known from the National Ice Center iceberg inventory are classified correctly (see URL www.natice.noaa.gov). The largest concentration of perennial ice is found just off the tip of the Antarctic Peninsula. The ice here has survived the previous melt season by avoiding being swept out to sea by the Weddell Gyre. Rough first-year ice in the classification surrounds smooth first-year ice which is located primarily in the inner portion of the ice pack. This is consistent with the classification results in [8]. In addition, the marginal ice zone exists on the perimeter of the ice pack as expected.

In order to gain an understanding of possible cross-confusion that may occur between classes using this algorithm, the Mahalanobis distance is computed between the final cluster centroids in principal component space. Since the Mahalanobis distance requires a cluster covariance matrix and each centroid comes from a different cluster, one of the centroids is treated as the reference and the other is considered the test vector. In the computation, the covariance matrix of the reference vector is used. The results are given in Table III for the 1996 JD 261-266 classification. The table contents may be interpreted by observing individual columns corresponding to a test ice type cluster. Each row value within a particular column is a measure of dissimilarity between the reference and test vectors. Thus, lower values correspond with higher probability that an ice type will be misclassified as the reference type. For example, the PER column implies that perennial ice is much more likely to be misclassified as pancake ice than smooth first-year ice.

As previously stated, the MAP algorithm requires an initial estimate of the *a priori* distribution $p(C)$. The technique is designed to use the $p(C)$ resulting from the classification of the previous image set. However, the first classification in the series requires the user to provide an approximate $p(C)$ for initialization. In an effort to determine the sensitivity of the final ice classification to the this parameter, a Monte Carlo analysis is performed. Several random realizations within a neighborhood of a nominal $p(C)$ are used in the 1996 JD 261-266 image segmentation. The study showed that the final spatial distribution of ice types is not particularly sensitive to the original *a priori* distribution.

Figure 12 shows classification images generated using two

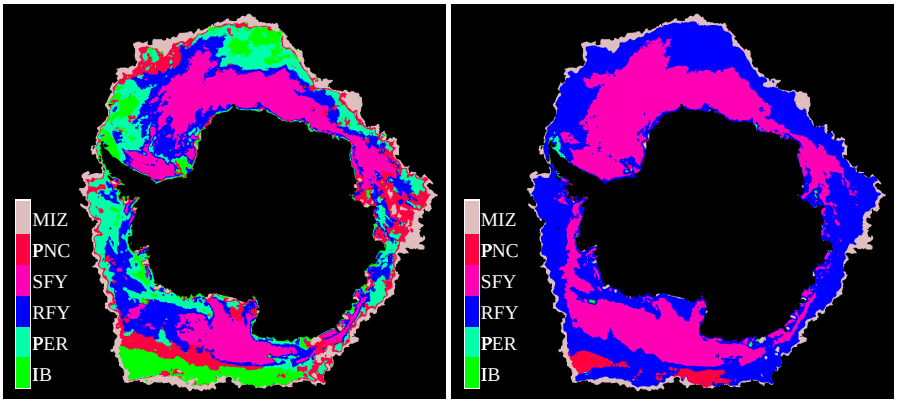

Fig. 11. ML (left) and MAP (right) sea ice classifications of 1996 JD 261–266 Antarctic data. The ML result is in error because of the assumption that each ice type is equally likely.

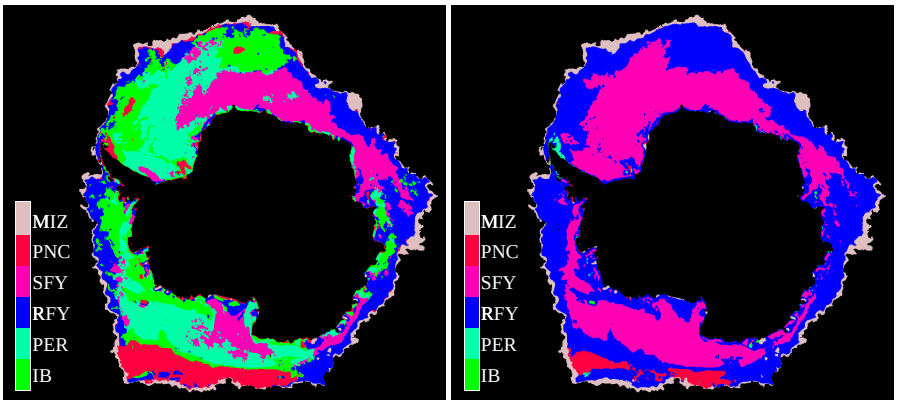

Fig. 12. K-means clustering classification results of 1996 JD 261–266 Antarctic data using different distance metrics. The left image is the conventional k-means result with the Euclidean distance metric. The right is k-means with the "MAP distance" metric as defined in the text.

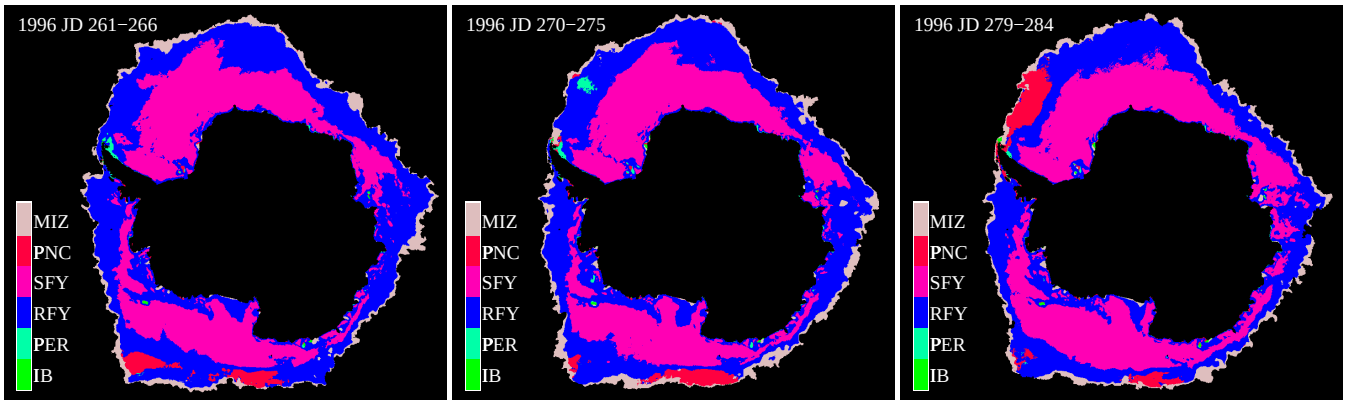

Fig. 13. MAP ice classification of the image series with day ranges 1996 JD 261–266, 270–275, and 279–284.

other methods for comparison. Both were implemented using the same training data for initial cluster centroids. The left image is the classification result of the standard k-means clustering algorithm. The k-means approach yields a solution that minimizes the within cluster sum of squared distances under the Euclidean distance metric. Since no regard is given to the probability of ice type, the k-means result has problems similar to the ML image. The second image was generated using a modified form of k-means in which a "MAP distance metric" measured the similarity between data samples and the centroids. The MAP distance metric is the negative of the *argmax* argument in Eq. (12). The resulting ice type map is nearly identical to the MAP classification. In fact, the two agree for 96.5% of the image sea ice pixels. The differences primarily occur in the number of pixels classified as pancake ice. Consequently, the modified k-means classifier result is similar to the maximum *a posteriori* technique result.

An obvious error in the MAP classification in Figure 11 is

the RFY labeled tongue extending from the Ross Ice Shelf. The perimeter of the ice shelf is actually a region of new ice formation and divergence. Consequently, the ice in this regime should have been identified as SFY. The source of the discrepancy is likely due to frost flower formation on the surface of smooth ice. Drinkwater and Crocker [27] found that frost flower formation can yield microwave signatures that are similar to RFY ice. The proposed classifier did not include a separate classification cluster for this ice type. A useful line of future research would include a study of the potential of segmenting frost flower covered ice from RFY ice using the results of Wensnahan [7] and Ulander [28].

Figure 13 shows the algorithm results when applied to a time series of images. The original images are generated with three days of overlap between consecutive intervals. The three day spaced classification maps illustrate stability in the ice classes between contiguous images. Since geophysical variability is greater on longer time-scales, we show results from images separated by three day gaps. The ice maps in Figure 13 reveal a number of interesting features. First, the temporal continuity between the spatial distributions of classes such as RFY and SFY ice types indicates that the algorithm is stable. However, some misclassification does occur such as the region of MY ice that appears in the outer ice margin of the Weddell Sea in the JD 270–275 image. In this case, pixels are exchanged between the RFY and MY categories.

Classified imagery can be used to better understand certain geophysical processes. For example, the tongue of "RFY" ice extending from the Ross Ice Shelf exhibits some interesting temporal behavior. As discussed previously, this region is likely not RFY ice, but SFY ice covered with frost flowers formed as off ice-shelf winds drive the ice pack northwards. Hence, temporal changes in the direction of the tongue relate to changes in wind direction over the ice pack through dynamic adjustments to the ice drift direction.

Another region of interest is in the outer Weddell Sea where a large region classified as pancake ice appears in the last frame. Initially, this feature appears to be an error in the classification. However, examination of the original data set images reveals that this classification relates to an actual physical event. From the first image in Figure 13 to the last, the scatterometer σ^o values increase several dB and the radiometer T_B values drop significantly. For example, the average NSCAT A_v value in the area rises from -11.2 to -7.4 dB. The SSM/I 37V T_B average decreases from 236 to 210 K. The final signatures are typical of pancake ice. A possible cause for this event is the occurrence of storm induced swells penetrating the ice margin. Under such conditions, intense ice floe fracturing and wave-washing of floes are observed in the field. Subsequent return of the signatures to values more typical of the marginal ice zone, and rough first-year ice confirm that this is an event of transitory nature. If this is indeed the case, it may be possible to relate pancake ice "blooms" in ice margin signatures to increased wave radiation stresses, or to surface flooding events caused by increased swells. Further analysis of this phenomenon could compare the occurrence of such signatures with the significant wave heights observed by altimeters in the region of the Southern Ocean directly off the ice margin at those times. Similarly, these events may be classified as "flooding" in future extensions of the algorithm.

VI. Conclusion

This study has demonstrated the utility of a multisensor, iterative maximum *a posteriori* sea ice type classification algorithm for Antarctic sea ice. The use of data collected from multispectral, dual-polarization, active, and passive instruments increases the level of information that can be exploited in segmenting the data. Through the use of principal component analysis, not only is the data dimensionality minimized, but the effects of noise and imaging artifacts are reduced. The resulting data set is classified in an iterative manner that utilizes MAP statistical techniques. The MAP classifier performs better than ML and the standard k-means and is very similar to a modified version of k-means with a different distance metric.

The iterative classification algorithm yields ice maps with spatial ice type distributions that are reasonable when general Antarctic sea ice dynamics are considered. However, while the algorithm appears to function well, a more detailed validation study is needed. Unfortunately, Antarctic validation data is difficult to obtain during this period of sensor overlap. Though SAR data exists for continental Antarctica, sea ice SAR imagery during the period spanned by our multisensor data set is scarce. Future research will apply the algorithm to Arctic data where validation data is much more abundant both spatially and temporally, and where current efforts are underway to plot ice drift and dynamics on a Lagrangian grid.

Several implications must be considered in a medium-scale classification such as the method presented in this study. First, the six day imaging period may introduce blurring in the images due to sea-ice motion resulting in ambiguous signatures and misclassification. The limiting factors for this data set are the scatterometers which need more time to achieve full coverage of the Antarctic ice pack. In the future, similar algorithms may be applied using instruments with wider swaths such as the SeaWinds scatterometer on board the QuikSCAT and ADEOS II spacecrafts. Furthermore, AMSR in conjunction with SeaWinds aboard ADEOS II will provide temporally and spatially coregistered active and passive data. This provides many of the channels required for such a method to be applied in the future. Both SeaWinds missions will reach full coverage in one to two days rather than six days. SeaWinds on QuikSCAT is currently in flight while ADEOS II is scheduled for launch before the end of 2001. Also, the relatively low resolution, even in the reconstructed imagery, implies that some pixels may contain a mixture of ice types. Thus, the classification result for a particular pixel is considered the spatial and temporal average behavior of sea ice in that region. A promising line of future research is the extension of this algorithm from a hard to a fuzzy classifier. That is, for each pixel the concentration of each ice type may be estimated. It is conceivable that the MAP probabilities could be used to achieve this. However, a greater understanding of the effects of within-footprint mixtures on observed microwave signatures is first required. Nevertheless, the algorithm yields results consistent with historic ice distributions and expectations.

This study has demonstrated one method for the application of multisensor data sets to classification problems. The use of

multiple sensors appears to improve the ability to identify different classes by combining the inherent strengths of each instrument. The three sensors used each add unique information to assist in segmenting the images into separate ice types. The scatterometers, NSCAT and ERS-2, are sensitive to surface roughness, volume inhomogeneities, and other scattering mechanisms which vary across different ice types. In addition, these instruments collect measurements at multiple incidence angles. Incidence angle dependence varies over the spectrum of sea ice types justifying the value of this parameter. NSCAT in particular is valuable in that it collects dual polarization measurements over a wider swath at higher resolution. The primary strength of the C-Band ERS-2 lies in greater penetration depth due to its lower frequency of operation. Unfortunately, both of these sensors require several days of data to obtain complete coverage of the Antarctic. SSM/I also contributes a great deal to the classification. As a passive instrument, the SSM/I sea ice signatures are a more a function of surface emissivity and dielectric properties than their active counterparts. The wide spectrum of frequencies and dual polarization nature of the SSM/I channels offers sensitivity to a larger range of surface properties than single frequency/polarization instruments. Additionally, the SSM/I measurement collection geometry allows complete coverage of the Antarctic usually in one day though at a lower resolution than NSCAT.

ACKNOWLEDGMENTS

M.D. performed this work at the Jet Propulsion Laboratory, California Institute of Technology under contract to the National Aeronautics and Space Administration. Radarsat data were kindly provided through ADRO Project #164 and support was provided by NASA Code YS Project 665-21-02. NSCAT data were obtained through the PO.DAAC at the Jet Propulsion Laboratory. SSM/I data provided by the Global Hydrology Resource Center (GHRC) at the Global Hydrology and Climate Center, Huntsville, Alabama.

REFERENCES

[1] W.F. Budd, "Antarctic sea ice variations from satellite sensing in relation to climate," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 15, pp. 417–426, 1975.

[2] F.D. Carsey, Ed., *Microwave Remote Sensing of Sea Ice,* Washington, D.C.: American Geophysical Union, 1992.

[3] D.T. Eppler, L.D. Farmer, A.W. Lohanick, and M. Hoover, "Classification of Sea Ice Types With Single-Band (33.6 GHz) Airborne Passive Microwave Imagery," *J. of Geophys. Res.*, vol. 91, no. C9, pp. 10,611–10,694, 1986.

[4] R. Kwok, E. Rignot, B. Holt, and R. Onstott, "Identification of Sea Ice Types in Spaceborne Synthetic Aperture Radar Data," *J. of Geophys. Res.*, vol. 97, no. C2, pp. 2391–2402, 1992.

[5] E. Rignot and M.R. Drinkwater, "Winter Sea-Ice Mapping from Multiparameter Synthetic-Aperture Radar Data," *J. of Glaciology,* vol. 40, no. 134, pp. 31-45, 1994.

[6] H. Hra, R.G. Atkins, R.T. Shin, J.A. Kong, S.H. Yueh, and R. Kwok, "Application of Neural Networks for Sea Ice Classification in Polarimetric SAR Images," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 33, no. 3, pp. 740–748, May 1995.

[7] M. Wensnahan, G.A. Maykut, and T.C. Grenfell, "Passive Microwave Remote Sensing of Thin Sea Ice Using Principle Component Analysis," *J. of Geophys. Res.*, vol. 98, no. C7, pp. 12,453–12,467, 1993.

[8] M. Drinkwater, "Satellite Microwave Radar Observations of Antarctic Sea Ice," in *Analysis of SAR Data of the Polar Oceans,* C. Tsatsoulis and R. Kwok, Eds. Berlin: Springer-Verlag, 1998, pp. 145-188.

[9] M. Golden, D. Borup, M. Cheney, E. Cherkaeva, M.S. Dawson, K. Ding, A.K. Fung, D. Isaacson, S.A. Johnson, A.K. Jordan, J.A. Kong, R. Kwok, S.V. Nghiem, R.G. Onstott, J.Sylvester, D.P. Winebrenner, and I.H.H. Zabel. "Inverse Electromagnetic Scattering Models for Sea Ice," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 36, no. 5, pp. 1675–1704, 1998.

[10] F.M. Naderi, M.H. Freilich, and D.G. Long, "Spaceborne Radar Measurement of Wind Velocity Over the Ocean–An Overview of the NSCAT Scatterometer System," *Proceedings of the IEEE*, vol. 79, no. 6, pp. 850–866, June 1991.

[11] D.G. Long and M.R. Drinkwater, "Cryosphere Applications of NSCAT Data," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 37, no. 3, pp. 1671–1684, 1999.

[12] D.G. Long and M.R. Drinkwater, "Azimuth Variation in Microwave Scatterometer and Radiometer Data over Antarctica," *this issue.*

[13] E. Attema, "The Active Microwave Instrument Onboard the ERS-1 Satellite," *Proceedings of the IEEE*, vol. 79, no. 6, pp. 791–799, 1991.

[14] J.P. Hollinger, J.L. Pierce, and G.A. Poe, "SSM/I Instrument Evaluation," *IEEE Trans. Geosci. Remote Sensing,* Vol. 28, No. 5, pp. 781-790, Sept. 1990.

[15] J. Hollinger *et al.*, "Special Sensor Microwave/Imager User's Guide," Naval Research Laboratory, Washington, D.C., Dec. 14, 1987.

[16] D. Long, P. Hardin, and P. Whiting, "Resolution Enhancement of Spaceborne Scatterometer Data," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 31, pp. 700–715, 1993.

[17] D.G. Long and D.L. Daum, "Spatial Enhancement of SSM/I Data," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, pp. 407–417, 1997.

[18] Q. Remund and D. Long, "Sea-Ice Extent Mapping Using Ku-Band Scatterometer Data," *J. of Geophys. Res,* vol. 104, no. C4, pp. 11515-11527, 1999.

[19] S.H. Yueh and R. Kwok, "Arctic Sea Ice Extent and Melt Onset from NSCAT observations," *Geophys. Res. Letters*, vol. 25, no. 23, pp. 4369–4373, Dec. 1998.

[20] M.R. Drinkwater and C. Haas "1994, Snow, Sea-ice and Radar Observations during ANT X/4: Summary Data Report," *AWI Berichte aus dem Fachbereich Physik,* 53, Alfred Wegener Institut fr Polar- and Meeresforschung, July 1994, 58pp.

[21] D.S. Early and D.G. Long, "Azimuthal Modulation of C-Band Scatterometer σ^o over Southern Ocean Sea Ice," *IEEE Trans. on Geosci. and Rem. Sens.*, vol. 35, no. 5, pp. 1201-1209, 1997.

[22] W.J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective,* Claredon Press, Oxford, England, 1988.

[23] C.T. Haas and H. Eicken, "Sea Ice Conditions during the Winter Weddell Gyre Study 1992 ANT X/4 with R/V Polarstern: Shipboard Observations and AVHRR imagery," *AWI Berichte aus dem Fachbereich Physik,* 34, Alfred Wegener Institut fr Polar- and Meeresforschung, Dec. 1992.

[24] N.W. Young, "Antarctic Iceberg Drift and Ocean Currents Derived from Scatterometer Image Series," *ESA SP-424*, pp 125–132, Nov. 1998.

[25] M.R. Drinkwater, "Active Microwave Remote Sensing Observations of Weddell Sea Ice," in *Antarctic Sea Ice: Physical Processes, Interactions and Variability, Antarctic Research Series* M.O. Jeffries, Ed. Washington, D.C.: American Geophysical Union, 1998, pp. 187–212.

[26] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis,* John Wiley and Sons, New York, New York, 1973.

[27] M. Drinkwater and G. Crocker "Modelling changes in the dielectric and scattering properties of young snow-covered sea ice at GHz frequencies," *J. of Glaciology,* vol. 34, no. 118, pp. 274–282, 1988.

[28] L. Ulander, A. Carlstrom, and J. Askne, "Effect of Frost Flowers, rough saline snow and slush on the ERS-1 SAR Backscatter Signatures of Thin Arctic Sea Ice," *Int. J. Remote Sens.*, vol. 16, no. 17, pp. 3287–3305, 1995.