ELSEVIER

Classification of Boreal Forest Cover Types Using SAR Images

Sasan S. Saatchi and Eric Rignot**

Mapping forest cover types in the boreal ecosystem is important for understanding the processes governing the interaction of the surface with the atmosphere. In this paper, we report the results of the land-cover classification of the SAR (synthetic aperture radar) data acquired during the Boreal Ecosystem Atmospheric Study's intensive field campaigns over the southern study area near Prince Albert, Canada. A Bayesian maximum a posteriori classifier was applied on the National Aeronautics and Space Administration/Jet Propulsion Laboratory airborne SAR images covering the region during the peak of the growing season in July 1994. The approach is supervised in the sense that a combination of field data and existing land-cover maps are used to develop training areas for the desired classes. The images acquired were first radiometrically and absolutely calibrated, the incidence angle effect in airborne images was corrected to an acceptable accuracy, and the images were used in a mosaic form and geocoded and georeferenced with an existing land-cover map for validation purposes. The results show that SAR images can be classified into dominant forest types such as jack pine, black spruce, trembling aspen, clearing, open water, and three categories of mixed strands with better than 90% accuracy. The unispecies stands such as jack pine and black spruce are separated with 98% accuracy, but the accuracy of mixed coniferous and deciduous stands suffers from confusing factors such as varying species composition, surface moisture, and understory effects. To satisfy the requirements of process models, the number of cover types was reduced from eight to five general classes of conifer wet, conifer dry, mixed deciduous, disturbed, and open water. Reduction of classes improved the overall accuracy of the classification over the entire region from 77% to 92%. ©*Elsevier Science Inc., 1997*

INTRODUCTION

One of the major challenges of developing Earth system process models both on global and on regional scales is the accurate representation of the terrestrial vegetation. These process models work at a variety of spatial scales ranging from meters to kilometers. Depending on the application of the processes and their scales, the definition of categories of vegetation types may change. For example, for global land-atmosphere models such as Biosphere-Atmosphere Transfer Scheme (BATS), 18 general land-cover types are defined that are often inferred from maps, atlases, and national databases (Dickenson, 1994). For finer-scale process models, the availability of high-resolution land-cover maps can improve the parameterization of landscape to functionally different strata. Currently, there are several approaches under investigation to statistically aggregate the high-resolution maps derived from remote-sensing techniques to a desired process model grid scale (Hall et al., 1995). These techniques are primarily focused on exploiting optical remote-sensing data such as that of the advanced very high resolution radiometer (AVHRR) and Landsat (Sellers et al., 1994; Townshend et al., 1991).

As a complement to optical remote-sensing techniques, land-cover maps derived from multipolarization, multifrequency synthetic aperture radar (SAR) systems are an important tool for terrestrial ecologists and process modelers. Independence of SAR data of solar irradiance and cloud cover is one significant reason for using this technique for land-cover classification, especially in northern latitude boreal forest and tropical rainforest where the acquisition of optical data is hindered by frequent cloud cover and fire smoke. In addition, the sensi-

* Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

Address correspondence to Dr. Sasan S. Saatchi, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109.

Received 27 November 1995; revised 12 August 1996.

REMOTE SENS. ENVIRON. 60:270–281 (1997)

655 Avenue of the Americas, New York, NY 10010

0034-4257/97/$17.00
PII S0034-4257(96)00181-2

tivity of the radar signal to moisture content and structural properties of vegetation may separate forest types, particularly when optical sensors are saturated over dense vegetation. Several studies, using a variety of classification approaches, have used SAR images for land-cover-type classification in forested regions (Saatchi et al., 1996; Rignot et al., 1994; Ranson and Sun, 1994; Dobson et al., 1994; Cimino et al., 1986). The application of these high-resolution maps in process models has not been thoroughly explored. For example, Bonan (1993) has used the SAR-derived land-cover map over the boreal forest of interior Alaska to improve the estimation of forest assimilation.

Our classification of land cover is used to address the specific requirements of the BOREAS (Boreal Ecosystem Atmospheric Study) modeling activities. Separating functionally important land-cover types for modeling the exchange of trace gases between the land surface and the atmosphere is the ultimate goal of this study. The area has been under intensive study during the BOREAS project for its important role in biogeochemical cycles between land and atmosphere at northern latitude (Sellers et al., 1995). The land-cover types in this region can be characterized by only a few dominant tree species. For example, the separability of conifer and deciduous stands and the dry and wet conditions in this region are important for estimating the rates of photosynthesis, respiration, carbon assimilation, and nitrogen concentration. Therefore, the process of classification with the use of SAR imagery is, first, to illustrate the capability of the instrument to identify these classes, and, second, to show the spatial pattern of these classes over a region used for ecosystem processing models. In this study, we discuss the application of SAR data for mapping of forest types in the BOREAS area. A supervised classification approach using a maximum a posteriori Bayesian classifier is applied on the three-frequency polarimetric Jet Propulsion Laboratory airborne synthetic aperture radar (JPL AIRSAR) data to identify eight classes. Classification accuracies are computed first for the areas used for training the classifier, then over several homogeneous sites examined during the field observation, and finally by comparing the results with a digital vegetation map assembled from infrared aerial photointerpretation performed in 1984.

BOREAS EXPERIMENT

The Boreal Ecosystem Atmospheric Study is a cooperative field experiment integrating land surface climatology, tropospheric chemistry, and terrestrial ecology. In general, the experiment was designed to extend the findings of FIFE over grass prairie to boreal forests, one of the earth's largest and complex biomes, where coniferous species dominate. The biome has upland forests, extensive wetlands, some deciduous species, and many lakes, and it is a major storage of organic carbon, mostly in the soil (BOREAS Science Steering Committee, 1990). Among the primary objectives of the experiment, 1) improving understanding of the processes that govern the exchange of energy, water, heat, carbon, and trace gases between the boreal forest ecosystem and the atmosphere and 2) developing and validating remote-sensing techniques to transfer our knowledge of these processes from local to regional scales are of great importance. A relevant scientific issue is the sensitivity of the boreal forest biome to changes in physical climate and vice versa. Mapping vegetation-cover types, changes in land use, and the species composition in the region can contribute to long-term climatological research studies.

Site Description

The focus of our paper is the BOREAS southern study area (SSA), which covers an area about 130 km in the east–west direction and 90 km from north to south (Fig. 1). The southern boundary is located approximately 40 km north of the town of Prince Albert, Saskatchewan, Canada. The SSA topography is gentle, with local elevations ranging from 550 to 730 m. Soils range from gray wooded to degraded black and are classified as brunisolic, gleysolic, chenozemic, luvisolic, and organic soil orders (Anderson and Ellis, 1976). Glacial deposits vary in thickness from 100 to 1000 m on the top of the Cretaceous bedrock. The western part of the SSA is in the Prince Albert National Park (PANP), and the eastern region falls within and around the Narrow Hills Provincial Forest.

The SSA is near the southern limit of the boreal forest and the transition to natural prairie grassland and agricultural land is 15 km to the southeast. The image data discussed in this paper are east of PANP in the area of the Narrow Hills Provincial Park. The image area also coincides with the BOREAS modeling grid (50×50 km) used mainly for verifying remote-sensing algorithms and ecosystem modeling results. The vegetation in this area is classified as mixed boreal forest. On well-drained and sandy soil, the predominant species is jack pine (*Pinus banksiana*). Poorly drained sites support black spruce (*Picea mariana*). Mixed stands of aspen (*Populus tremuloids*), balsam poplar (*Populus balsamifera*), and white spruce (*Picea glauca*) are found on well-drained glacial deposits. In poorly drained areas throughout the study area, bogs support black spruce with tamarack (*Larix laricina*). The fen areas are composed mostly of sedge (*Carex spp.*) with discontinuous cover of tamarack or swamp birch (*Betula pumila*). Localized logging for paper pulp and fence posts is common along Highways 106 and 120 and along Harding Road (see Fig. 1). The northeastern part of the study area encompasses a part of the Fishing Lakes burn that occurred in 1977 and 1978. Stands of small (<5 cm) jack pine regrowth now cover most of the burn areas.

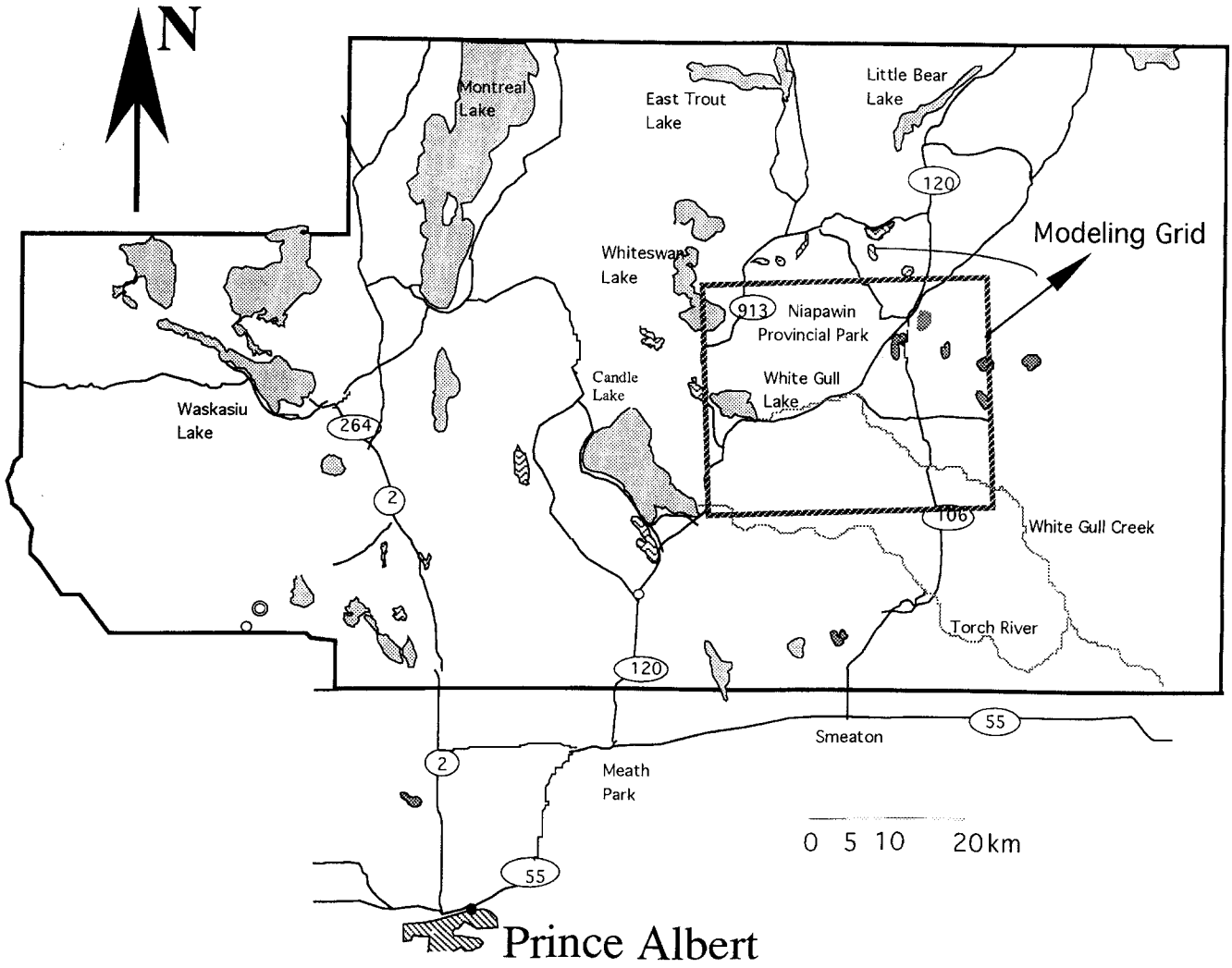

Figure 1. Map of the BOREAS southern study area and the modeling subgrid.

Major land-cover types are identified according to the needs of the BOREAS scientific applications. These land-cover types are chosen on the basis of their dominant species, canopy closure, soil organic properties, and their roles in determining the physics of the interaction of land surface and atmosphere. The land-cover categories consist of dry conifers (e.g., jack pine), wet conifers (e.g., black spruce), deciduous (trembling aspen), clear cut, open water (lakes and river), brushland, treed muskeg, mixed coniferous and deciduous trees, and regrowth (e.g., young jack pine). Between summer of 1993 and fall of 1994, forest stands of major land cover were sampled to measure tree species composition, stand geometry, biomass density, and several other forest canopy attributes. Data collections on the ground were performed for many applications and are available for all the flux tower and auxiliary sites. The flux tower sites are mainly single species stands.

In addition, there exists a digital vegetation map of SSA that was assembled from 1:12,500 scale infrared aerial photography and field reconnaissance notes in 1984. This vegetation map has been verified on the ground, but no accuracies are provided. The map consists of 40 different classes, regrouped to simplify the representation of vegetation types for dominant classes (Fig. 2). The map does not show recent changes due to tree logging, regrowth, and transformation of treed muskeg to predominantly black spruce stands.

AIRSAR DATA

The JPL airborne synthetic aperture radar (AIRSAR) was flown aboard a National Aeronautics and Space Administration DC-8 during all the intensive field campaigns (IFC) in summer of 1993, in April 1994 during the thaw period of the boreal forest, and in summer and fall of 1994. The AIRSAR operates at three frequency bands, P-band (68-cm wavelength), L-band (24-cm), and C-band (5.6-cm), with fully polarimetric capability. The incidence angle of the radar varied between approxi-

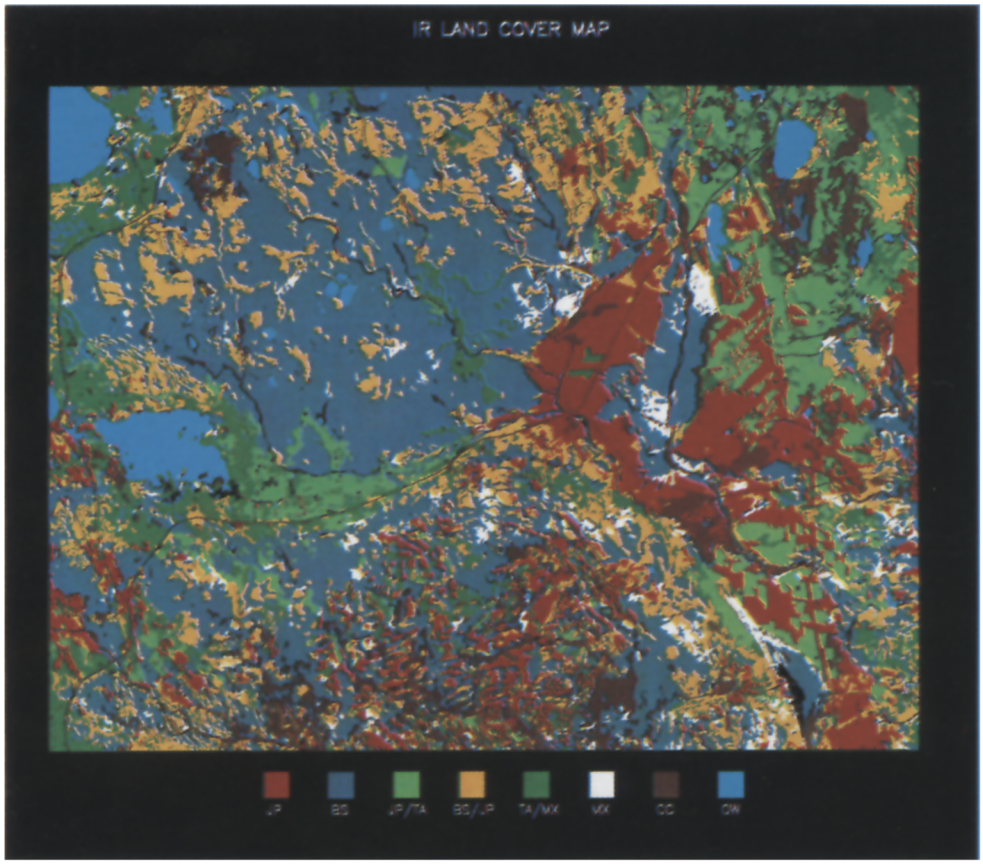

Figure 2. Digital vegetation map of the modeling subgrid assembled from visual interpretation of infrared aerial photography. The map includes eight different types of land cover that result from the regrouping of 40 original land-cover classes.

mately 20° and 60°. The radar data used for land-cover classification were acquired on 21 July 1994 and processed in synoptic mode (50-km swath). We chose this date to prevent possible errors in classification due to the partially frozen condition during the thaw period and to the leaf-off condition during the fall season. We used images from several parallel flight lines in a mosaic mode to create larger area coverage over the modeling grid. The calibration, radiometric correction, and mosaic of the images were performed in several steps.

Image Calibration

In this study, we made use of synoptic SAR images that were acquired with parallel flight lines in a "race track" trajectory. The synoptic images have larger coverage (approximately 50 km) but only three polarizations. These images are often processed for the purpose of surveying the area and are not absolutely calibrated. We processed a total of 15 synoptic images to cover all the bands and polarizations of the AIRSAR system. Calibration of images was performed by using fully polarimetric calibrated frame images processed over a part of the synoptic images. Absolute calibration constants were obtained by computing the ratios of backscattering coefficients from identical areas from both images and applying the calibration constants to all synoptic images. When compared with frame images, the synoptic images were absolutely calibrated with less than 0.1-dB error for all polarization

Figure 3. P-band polarimetric color overlay of the AIRSAR mosaic image of the modeling subgrid within the BOREAS southern study area acquired on 21 July 1994. P-band HH, HV, and VV polarizations are in red, green, and blue, respectively. The mosaic image is coregistered with the digital vegetation map and georeferenced to universal transverse Mercator coordinates with North being parallel to the side of the image.

channels. The frame images were calibrated both internally and externally by using data collected over an array of corner reflectors deployed over the Rosemond dry lake calibration site in California before and after the AIRSAR campaign. After the absolute calibration, the images were resampled to ground range to remove the distortions in the near-range and far-range pixels.

Incidence Angle Correction

One of the disadvantages of airborne SAR data, when used for land-cover classification, is the variation of the incidence angle along the range lines across the image (20–60°). Consequently, areas with similar land-cover types produce different backscatter signatures if they are imaged at different incidence angles; and, depending on the scene characteristics, the variation of the backscatter signature along each range line may be different. These effects can cause inaccuracies in a consistent class separation over the entire image. Correction of the image for incidence angle effects, therefore, becomes a necessary but impossible task to accomplish exactly. To have an optimal correction for incidence angle effects, several approaches have been suggested. Yueh et al. (1988) normalized the SAR data by the total power. This technique eliminated most of the incidence angle effects but at the same time changed some of the information in SAR backscatter signatures. The resulting normalized images were not able to discriminate all classes. Another method was proposed by Sader (1987), in which homogeneous areas of the same types were chosen along the range line, and the total image was calibrated such that these areas had equal backscattering signatures. This technique will not work in areas with complex land-cover types. Rignot and Drinkwater (1993) corrected the incidence angle effects in their classification of sea-ice types by first segmenting the image along the range line, performing range-dependent clustering of the image, regrouping the clusters, and employing a supervised classification to produce self-consistent classes across the image scene. However, their technique requires that similar class types be represented in each segment of the image. Over complex land-cover types, sometimes only a limited number of classes is present over each range segment, causing difficulty in regrouping the clusters and removing the range-dependent effects. Ranson and Sun (1994) used AIRSAR images over forested land surfaces, selected a part of each image line within sapwood areas, calculated the mean and standard deviation of these pixels, and discarded all pixels falling outside of ± 2 standard deviations. The remaining pixels were used to estimate the mean values at each image row, then a linear regression was used to estimate the calibration ratio for each line, and, subsequently, the entire image was calibrated by using these ratios. When employing this technique, it was found that the linear regression method did not always compensate for the inhomogeneous scene characteristic along the range line.

The synoptic images used in this study were corrected for incidence angle variations with a technique slightly different from that of Ranson and Sun (1994). We plotted the incidence angle variations for each range line, and then a nonlinear regression in conjunction with a cubic spline smoothing algorithm was used to estimate the general behavior of the incidence angle variations along each range line. The regression curve was then normalized by the mean backscattering coefficient of the range line and then used to correct for the incidence angle effects of that range line. The entire image was then corrected line by line.

Image Mosaic

After calibration and incidence angle correction, the images from each frequency band and polarization were used in tandem to generate a mosaic image over almost the entire modeling subgrid. Figure 3 shows a color composite of the mosaic image at P-band (red, P-HH; green, P-HV; blue, P-VV). Because the images were acquired from flight lines with the same heading, they also had an area of overlap with adjacent images. A linear feathering technique was then employed to remove the tonal inconsistencies that existed at the areas of overlap. In some areas where incidence angle effects were not optimally corrected, the feathering technique guaranteed further smoothing at the edges of images. If the overlapping regions were near the lakes where there was a dramatic change in the radar backscatter signature, incidence angle effects could not be totally removed, and the edge effects were still obvious in the mosaic image.

CLASSIFICATION METHODOLOGY

When designing a classifier, it is important to define the mathematical basis of the classifier and, at the same time, to distinguish between the supervised and unsupervised learning procedures within the classifier. Here, we make use of a maximum-a-posteriori (MAP) Bayesian classifier developed for multifrequency polarimetric SAR data (Rignot and Chellappa, 1993). The MAP classifier models the SAR amplitudes as circular Gaussian distribution, which means that textural variations in radar backscatter from the surface are not considered to be significant enough to be incorporated into the classification scheme. In this method, the a priori distribution of image classes is modeled by using a Markov random field. From the models of the a priori distribution of classes, a model for the a posteriori distribution of the image classes is derived from the SAR image by using the Bayes' theorem. The optimal image classification of the SAR data is defined as that which maximizes the a posteriori distribution of classes and is called the maximum a posteriori estimate of the image classes.

The MAP method is inherently different from and superior to the *maximum likelihood* estimation (MLE)

procedure. The classifiers based on the maximum likelihood methods view the parameters (classes) as quantities whose values are fixed but unknown, and the best classification is defined to be the one that maximizes the probability of obtaining the samples actually observed (Duda and Hart, 1973). The MAP classifier views the parameters or classes as random variables with some a priori distribution. Iterative observation of the feature space converts this into an a posteriori density, thereby revising the decision about the true nature of classes. Because class distributions are analytical functions, the class characteristics obtained from training areas stay the same as in MLE. An advantage of this technique, besides its mathematical rigor, is that it is general enough and, with minor modification of the feature space, can be applied to both optical and SAR images, therefore creating the opportunity for both comparative and synergistic studies.

A number of other classifiers that have been used successfully but with limited capability for generalization also are available in the literature. Among them, Ranson and Sun (1994) used a combination of principle component analysis and MLE to come up with about 80% accuracy over Northern Experimental Forest near Howland, Maine. Pierce et al. (1994) introduced a knowledge-based classifier over a test site in northern Michigan. This technique depends on the absolute backscatter values derived over training areas and the texture information that may be degraded owing to multilook averaging in polarimetric SAR data in removing the speckle noise.

The learning procedure for the classifier is supervised in the sense that the state of the nature (class label) is known in advance, and training areas are chosen on the basis of a priori knowledge of the scene or the visual interpretation of the image. To implement the MAP classifier over the SAR mosaic image, we first define the a priori distribution of the SAR data for image classes by computing the average covariance matrix over the single training area. We concentrated on eight categories of training cover: 1) jack pine (JP), 2) black spruce (BS), 3) trembling aspen/mixed (TA/MX), 4) mixed jack pine and aspen (JP/TA), 5) mixed black spruce and jack pine (BS/JP), 6) mixed strands (MX), 7) clear cut, disturbed, and nonforest (CC), and 8) open water (OW). For each category, we selected large homogeneous stands from the knowledge acquired during the field observation and the existing land-cover map. The average covariance matrices are then computed over training areas for all three frequencies. Here, we used three training areas for jack pine stand, depending on the density and age, and two black spruce areas from a tower site and mature treed muskeg stand. The use of a limited number of training areas ensures realistic classification accuracy and the extrapolation of the results to the entire image. Table 1 lists the calibrated radar backscattering coefficients, copolarized phase difference in degrees, and the coefficient of correlation in the linear domain between the complex amplitudes at HH and VV polarizations. The radar characteristics are obtained from the frame images processed over 10×10 km areas within each synoptic image. The forest stands chosen for the training areas were imaged at nearly the same incidence angle (typically, about 45° incidence angle); thus the radar parameters for the image classes are assumed to be independent of the incidence angle. However, the SAR image was classified over all of its angle variations and, although the images were corrected for incidence angle variation along the range line, we can still expect some misclassification, particularly near the areas of overlap. Among the training areas, we encountered some difficulty in identifying aspen stands because of their small sizes within the BOREAS modeling grid and their vicinity to mixed stands. As a result, aspen class is labeled TA/MX to illustrate the aspen-dominated mixed stands.

RESULTS AND DISCUSSION

The map of forest types constructed from SAR data is shown in Figure 4. This result is obtained by using polarimetric data at P-band and the HH and HV polarizations at L- and C-bands. The choice of the frequency and polarization channels for achieving the optimum classification results was made by changing the dimensionality of the classification, or, equivalently, reducing the number of elements in the covariance matrix of each pixel that are used for classification. Consequently, the optimum classification accuracy was obtained by excluding only L- and C-band VV polarizations. In this process, it was also found that the contribution of P-band data was crucial in separating the classes. The reason for this combination is partly due to the calibration and radiometric inaccuracies at higher frequencies. In particular, the C-band VV-polarized synoptic mosaic data suffered from banding in the image, and inaccuracies resulted from incidence angle correction. In fact, when a radar channel does not separate two image classes, it adds as a noise source to the classification and increases the classification probability of error. The combination of polarizations and frequencies used to attain maximum separability differs from a similar technique applied on AIRSAR frame images over Alaskan boreal forest where the highest accuracy was obtained by only L-band and C-band HV polarizations (Rignot et al., 1994). We believe the reason for this difference resides in the poor radiometric accuracy of the synoptic images at high frequencies in our case and in the P-band interference problem in the data used over the Alaska region.

Classification accuracy for each class is determined by measuring the number of pixels correctly classified into the class divided by the total number of pixels in that class and is illustrated in the form of a confusion matrix. In assessing the total classification accuracy, we included open water and clear cut, though they are often separated with no difficulty within SAR images. The con-

Table 1. Radar Characteristics of the Training Sites of Forest-Type Classes at P-, L-, and C-Band Frequencies on 21 July 1995

Frequency	σ^0_{HH}	σ^0_{HV}	σ^0_{VV}	φ	ρ	*Pixels*	*Class*
P-band	−1.13	−12.40	−3.74	69	0.48	2814	JP
	−7.02	−14.65	−6.21	28	0.37	1097	BS
	−4.80	−13.99	−8.24	78	0.25	4920	TA/MX
	−6.66	−17.14	−5.75	51	0.24	1883	JP/TA
	−1.28	−12.33	−0.44	64	0.46	1850	MX
	−6.25	−14.28	−6.58	36	0.20	2006	BS/JP
	−10.24	−22.69	−14.78	−8	0.36	1097	CC
	−15.26	−27.71	−14.35	19	0.35	286	OW
L-band	−1.80	−12.35	−6.49	120	0.41	2814	JP
	−5.73	−13.62	−6.75	23	0.36	1097	BS
	−6.69	−12.16	−7.41	32	0.14	4920	TA/MX
	−5.87	−15.47	−7.39	49	0.30	1883	JP/TA
	−2.50	−10.83	−4.01	94	0.26	1850	MX
	−5.07	−13.23	−7.83	63	0.16	2006	BS/JP
	−9.20	−20.32	−12.60	−19	0.44	1097	CC
	−20.21	−29.77	−16.54	2	0.48	286	OW
C-band	−6.47	−13.03	−8.56	4	0.32	2814	JP
	−5.24	−13.05	−7.36	15	0.47	1097	BS
	−7.13	−12.84	−6.38	5	0.47	4920	TA/MX
	−5.98	−13.59	−9.16	11	0.35	1883	JP/TA
	−4.48	−12.06	−7.08	15	0.32	1850	MX
	−5.74	−12.88	−7.24	12	0.48	2006	BS/JP
	−8.69	−18.38	−11.51	17.5	0.58	1097	CC
	−22.70	−29.75	−21.19	10	0.34	286	OW

σ^0_{pq} is the backscattering coefficient at pq polarization averaged over the number of pixels within each training site, and φ and ρ are the phase difference in degrees and the coefficient of correlation of HH and VV polarizations, respectively. The classes are: JP (jack pine), BS (black spruce), TA/MX (aspen dominant mixed), JP/TA (jack pine dominant mixed with aspen), MX (mixed), BS/JP (black spruce dominant with jack pine), CC (clear cut), OW (open water).

tribution of each frequency in the total classification was assessed qualitatively when the classifier results were examined during the dimensionality test. The results indicate that the HV polarizations contribute the most for forest-type mapping at all frequencies. As shown in Table 1, the HV channels at L-band and P-band show the highest variability over the range of forest types because they are mainly related to the volume scattering within forest canopy and in turn sensitive to the forest biomass density. Furthermore, over forested areas, the HV backscatter is less sensitive to the incidence angle variation, and therefore the channels are less contaminated by the correction errors that may have remained over the image mosaic. The copolarized backscatter is less variable over different stands; but, because the calibration of copolarized channels is usually better than that of cross-polarized channels, their role in separating classes is significant. For example, over low vegetation, clear cut, and open water, the HV-polarized backscatter is very low, and the copolarized backscatter signatures are the primary source for separating these classes.

Table 2 shows the confusion matrix computed from the results of MLE and MAP classifiers over the training areas with 90% and 96% accuracies, respectively. For the mixed aspen and jack pine (JP/TA) class, pixels over the training area were classified with only 72% accuracy. The reason for this is the similar average copolarized backscatter values at all three bands. In general, for mixed stands, the choice of the training areas is poor compared with the monospecies homogeneous stands and, as a result, the mean backscatter returns for these sites are not very distinctive. Therefore, we expected poor accuracies over TA/MX sites because, over this region, most of the aspen stands are mixed with conifer trees. Jack pine and black spruce stands were classified with 100% and 99% accuracy, respectively. P-band and L-band HH polarizations are the main channels for separating these two classes. In jack pine stands, the trees are taller with less foliage and the ground surface is dry and smooth, which collectively contributes to high double-bounce return at HH polarization (Moghaddam and Saatchi, 1995). The black spruce stands, on the other hand, have shorter trees, more foliage, and a thick and wet moss layer and thus lower returns at HH polarization because of the absorption of the electromagnetic energy by the underlying moss layer.

To examine the ability of the classifier in separating coniferous and deciduous stands, we applied the classifier, without any changes in its current configuration, on an AIRSAR frame image acquired over the aspen tower site in Prince Albert National Park on the same date. The image covers the area south of Halkett Lake and

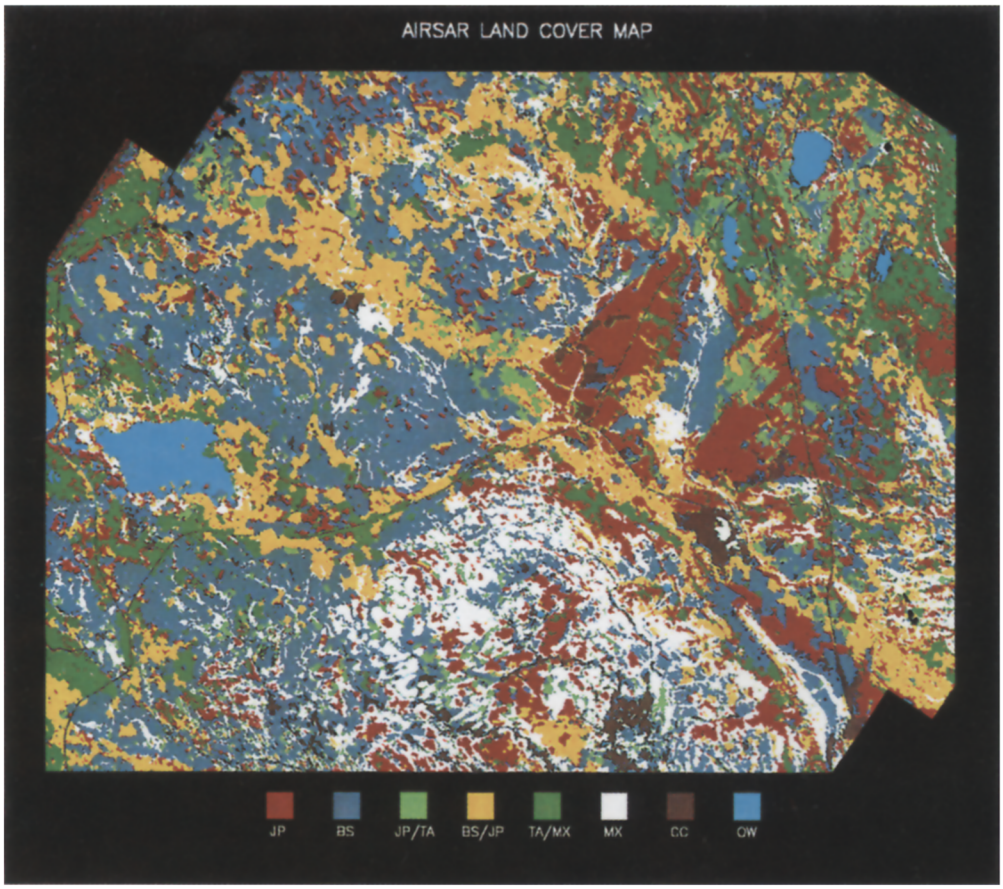

Figure 4. Map of forest types obtained from P-band polarimetric and L- and C-band HH and HV polarizations. The map includes eight classes similar to those in Figure 2. The classes are JP (jack pine), BS (black spruce), JP/TA (dominant jack pine mixed with trembling aspen), BS/JP (dominant black spruce mixed with jack pine), TA/MX (dominant trembling aspen mixed with other conifers), MX (mixed conifer), CC (clear cut, bogs, and disturbed), OW (open water).

north of dirt road Rt. 240 and is centered at the aspen tower site at almost 45° incidence angle. The area is covered mainly with aspen trees and with small scattered patches of balsam poplars (*Populus balsamifera*) that are similar in structure to aspen. The result of the classification is shown in Figure 5. From a visual interpretation of the map, it appears that the classifier separates the aspen stands with no difficulty. Over the tower site, the

Table 2. Confusion Matrices of Forest types Derived from MLE and MAP Classifiers

Species	*JP*	*BS*	*TA/MX*	*TA/JP*	*MX*	*BS/JP*	*CC*	*OW*
MLE								
JP	98	0	1	0	1	0	0	0
BS	0	95	0	4	0	1	0	0
TA/MX	0	0	98	0	1	0	0	0
JP/TA	1	26	2	58	0	13	0	0
MX	6	0	0	0	94	0	0	0
BS/JP	0	13	3	1	0	82	0	0
CC	0	0	0	0	0	0	100	0
OW	0	0	0	0	0	0	2	98
MAP								
JP	100	0	0	0	0	0	0	0
BS	0	99	0	0	0	1	0	0
TA/MX	0	0	100	0	0	0	0	0
JP/TA	0	20	0	72	0	7	0	0
MX	0	0	0	0	100	0	0	0
BS/JP	0	0	0	0	0	100	0	0
CC	0	0	0	0	0	0	100	0
OW	0	0	0	0	0	0	2	98

Abbreviations: MAP=maximum a posteriori classifier; MLE=maximum likelihood classifier.
The diagonal elements of confusion matrices define the percentage of those pixels classified into the correct class.

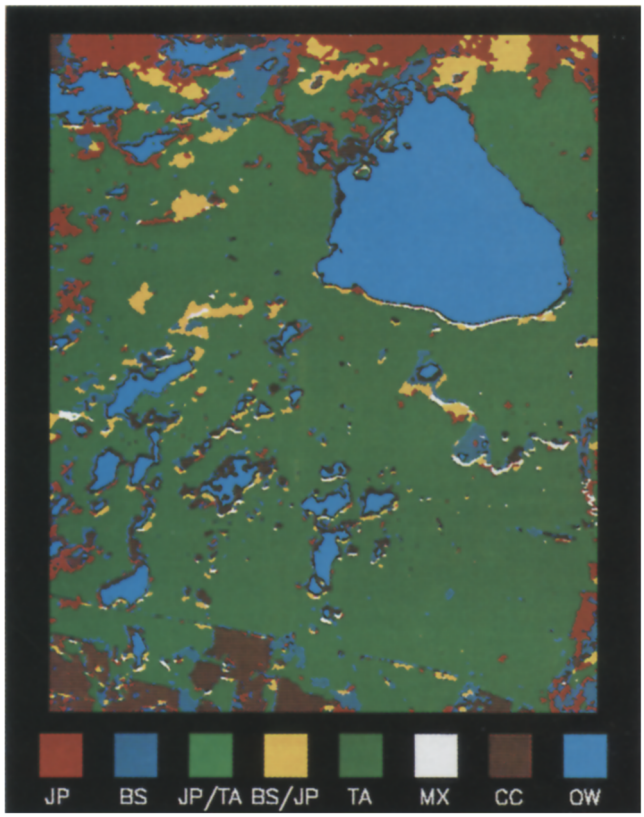

Figure 5. Map of forest types obtained over the aspen tower site in Prince Albert National Park. The map was obtained from a frame image covering an image of approximately 10×10 km. The cover types are similar to those in Figure 4, with aspen being separated from other land-cover types.

Figure 6. (a) Reduced SAR map over the modeling subgrid derived from the original SAR map shown in Figure 4. The map includes five classes of conifer-wet, conifer-dry, mixed deciduous and conifer, clear cut, and open water. (b) Reduced digital vegetation map derived from the original vegetation map by regrouping the land-cover classes.

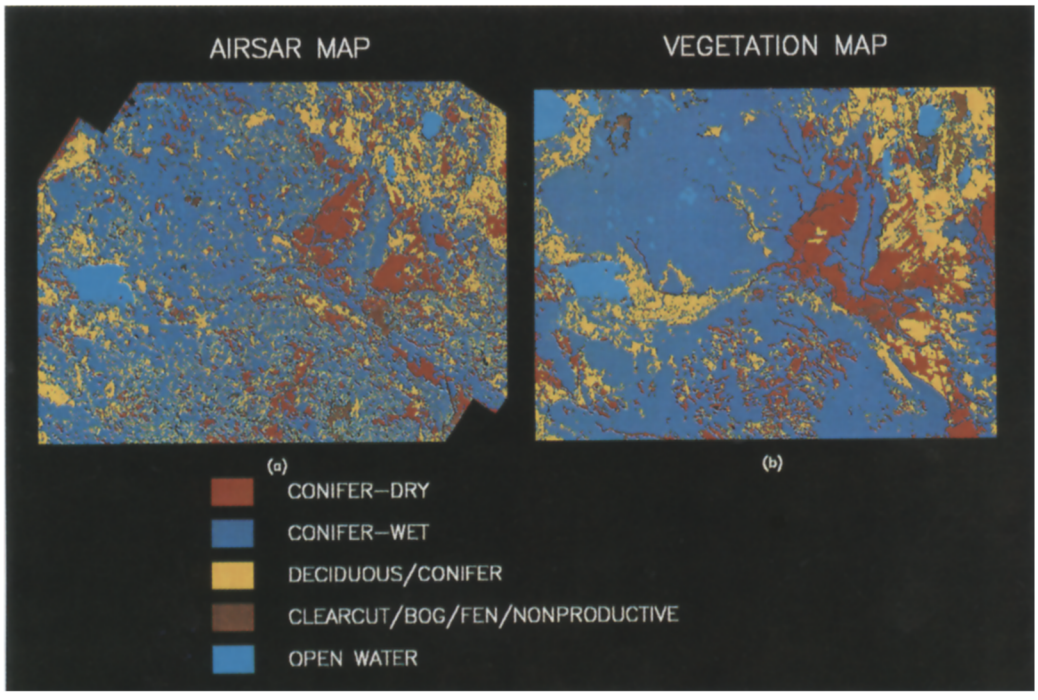

Table 3. Tree Species Composition of 19 Forest Stands in the Modeling Subgrid Area

Site	*Stand Composition 1 (Ground Truth) (%)*	*Stand Composition 2 (Cover Map) (%)*	*Stand Composition (SAR MAP) (%)*
G2L7S	BS: 40, Lala: 60	BS: 99, JP/BS: 1	BS: 64, MX: 36
G9I4S	BS: 99, Salix: 1	BS: 100	BS: 100
G2I4S	BS: 92, TA: 2, Salix: 2, Bepa: 3	BS: 100	BS: 96, JP: 4
G6K8S	BS: 94, Lala: 6	BS: 100	BS: 100
F1N0M	TA: 71, WS: 25, Abba: 3	TA/JP: 92, BS: 8	TA/MX: 96, JP: 4
G8L6P	JP: 91, TA: 9	JP/TA: 100	JP: 76, JP/TA: 8, BS/JP: 16
F5I6P	JP: 100	JP: 100	JP: 92, TA/MX: 8
G9L0P	JP: 100	JP: 100	JP: 100
F7J0P	JP: 17, WS: 4, BS: 74, TA: 4	BS: 100	BS: 96, JP: 4
F7J1P	JP: 59, BS: 17, TA: 23	JP/TA: 72, TA: 16, BS: 12	JP: 60, TA/MX: 24, MX: 16
G4K8P	JP: 100	JP: 100	JP: 100
G1K9P	JP: 100	JP: 100	JP: 96, BS/JP: 4
G7K8P	JP: 100	JP: 100	JP: 100
G4I3M	TA: 49, WS: 43, Bepa: 6, Abba: 3	TA/WS: 96, WS/TA: 4	TA/MX: 20, BS/JP: 68, JP: 12
TE-OBS	BS: 100	BS: 100	BS: 100
TE-OJP	JP: 100	JP: 100	JP: 100
TF-YJP	JP: 100	JP: 84, JP/TA: 16	JP: 64, JP/TA: 36
ADM-3		TA/JP: 92, JP/TA: 8	TA/MX: 92, JP/TA: 8
BDH-4		JP/TA: 76, BS: 24	TA/MX: 20, JP/TA: 68, MX: 12

The ground-truth data were taken from the field notes of TE-6 investigators (Sellers et al., 1995). WS is white spruce (*Picea glauca*), Lala is *Larix laricina*, Abba is Balsam Fir (*Abies balsamea*), Bepa is paper birch (*Belula papyrifera*). Numbers in ground-truth column indicate the percentage of each tree species based on the number of stems within the test plots. In vegetation cover and SAR maps, the numbers indicate the percentage of image pixels of each stand classified in type of forest.

classification accuracy reached 100%. This is one of the striking results of the SAR classification because, in general, the separation of coniferous and deciduous stands in boreal forests is considered one of the most challenging problems in any land-cover classification. This result also indicates that, over homogeneous stands, the structural information of the forest embedded in the SAR backscatter data becomes one of the key discriminants in the forest-type classification.

To analyze the accuracy of the SAR-derived cover map further, we compared the map with the field data and the existing land-cover map derived from infrared aerial photography. Table 3 shows the tree species composition of 19 test sites within the modeling grid obtained from actual measurements for each site, the vegetation map, and the SAR map. The ground measurements were conducted during the intensive field campaigns in summer of 1994 and coincide with the time frame in which the SAR data were acquired. The species composition was measured on small plots within each stand and was not designed to address the species composition at the SAR pixel scale. The vegetation map is almost 10 years old and may be inaccurate because it is based on a visual interpretation of the aerial photography and does not include the changes that have occurred since then. However, we included the map as an extra source for evaluating the accuracy of the SAR map. Moreover, the classifier was used to label each pixel by the dominant forest type and was never intended to estimate the species composition. Nevertheless, by performing this comparison, we are able to examine the general performance of the classifier and the capability of SAR to identify species composition.

The SAR map was georeferenced and coregistered with the vegetation map with less than one pixel (30 m) accuracy. The center locations of the sites were identified on the images by using the GPS (ground positioning system) data. Stand compositions of 19 sites were computed over 5×5 pixels from SAR and vegetation maps. The results in Table 3 indicate that classifications of auxiliary sites and tower sites are in good agreement with the field data and the vegetation map. Over 13 forest stands, errors in percentage of each species represented in the classification are less than 8%. The remaining six sites are mixed and contain species that are not included in the SAR classification. Over these sites, the errors in estimating species composition can increase to 20% with the exception of auxiliary site G4I3M, where the error exceeded 50%. These errors stem from several factors: 1) the spatial variability of species composition within the mixed stands is not compatible with the pixel sizes of the SAR map, 2) the location of the sites on the SAR map can be wrong owing to errors in the GPS measurements that may be larger than 100 m, and 3) the number and size of plots used in the field measurements may not be adequate for the mixed stands. Furthermore, because a combination of tree geometry, biomass, and surface conditions contributes to changes in SAR backscatter, the presence of several tree species within one SAR pixel will add to the confusion of the classifier in separating stands. These results suggest that the SAR map can be

Table 4. Area Percentage of Land-Cover Types over 875 km² of Modeling Subgrid Derived from Original SAR and Vegetation-Cover Maps

Cover Type	*SAR Map (%)*	*Cover Map (%)*	*Difference (%)*
JP	18.26	17.73	+0.53
BS	30.27	37.79	−7.52
TA/MX	8.17	7.86	+0.31
JP/TA	5.73	10.01	−4.28
MX	16.36	8.39	+7.97
BS/JP	15.57	12.91	−2.66
CC	3.62	3.39	−0.23
OW	1.23	1.59	−0.36

used to determine the species composition on the scale of the SAR image data. Given that SAR classification is performed by assuming that the training areas are purely single species stands, we expect that the SAR data would have better accuracy in mapping stand composition if the training areas were chosen differently.

Next, we examine the accuracy of the SAR map over the entire modeling grid by computing the percentage of area covered by each forest type in the region. An area of approximately 25×35 km is taken from the middle of the SAR map, and the number of pixels of each forest type is counted and divided by the total number of pixels. In Table 4, the percentage of area covered by each type in the SAR map and in the vegetation-cover map are compared. The difference between the two maps represented by the percentage of change implies a combination of errors in both maps and changes in the land cover between the times of the two data takes. If the vegetation-cover map is considered accurate at the time of the SAR data take, then the difference can mean that 23% of the total area was classified inaccurately. Field observations during BOREAS campaigns showed that certain parts of the land cover have been altered. For example, some logged and burned areas have been forested, and some forest areas have been recently cut. Because there is no accurate information about land-cover types on a regional scale, the assessment of the accuracy of the SAR map can be difficult. Given the uncertainties in the vegetation-cover map, we expect that, on a regional scale, more than 77% of the total area can be classified accurately with SAR imagery.

Process Modeling Requirements

Land-cover maps can be used as one of the parametric inputs to ecosystem process models. The requirements for accuracy and spatial scale of the map depend on the ecosystem model and the application. For example, general circulation models (GCM) have incorporated 1° by 1° global land-cover classification maps (Sellers et al., 1994). Recently an AVHRR/NDVI (normalized difference vegetation index) based global land-cover map also has become available as an input to GCMs (DeFries and Townshend, 1994). For modeling the net canopy assimilation in boreal or tropical forest, ecosystem models may require much finer resolution data over local or regional scales (Bonan, 1993). The BOREAS process models require five major land-cover types for the region. These are conifer wet, conifer dry, deciduous, mixed conifer and deciduous, and fen and disturbed. As an attempt to produce maps that can be readily used as input parameters to these models, we combined classes and modified the SAR and vegetation-cover maps to represent these five classes. Because pure deciduous and fen sites are rare over the modeling subgrid mapped by SAR, we chose conifer wet, conifer dry, mixed deciduous and conifer, clearing/disturbed, and open water as typical cover types for the region. The new classes were formed by grouping BS, BS/JP, and MX (mixed wet) into the conifer-wet class; JP into the conifer-dry class, and TA/MX and JP/TA into the mixed conifer-deciduous class. The clear cut and disturbed and the open water classes were not changed. The results are shown in Figure 6. By preserving the original pixel size (30 m), the new maps can be used in future for the accurate estimation of land-use change due to environmental and anthropogenic forces. The modified SAR and vegetation-cover maps show similar patterns of land-cover types in the region. A comparison of the two images over a 25×35 km subarea is given in Table 5. Results indicate that the accuracy of the SAR image can improve when fewer classes are used. The difference between the two maps has reduced to only 7.3% of the total area. With a reduction in the number of classes to functionally significant land-cover types, SAR data can provide maps with greater than 92% accuracy over the modeling grid.

Table 5. Area Percentage of Land-Cover Types over 875 km² of Modeling Subgrid Derived from Reduced SAR and Vegetation-Cover Maps

Cover Type	*SAR Map (%)*	*Cover Map (%)*	*Difference (%)*
Conifer-wet	62.22	59.34	+2.88
Conifer-dry	18.27	17.73	+0.54
Deciduous/conifer	13.91	17.22	−3.31
Disturbed/bog	3.62	3.39	+0.23
Open water	1.23	1.59	−0.36

SUMMARY AND CONCLUSIONS

This work summarizes the approach and the results of mapping forest types in the southern study area of the BOREAS project in the boreal forest of Canada by using SAR imagery. The images were collected by the JPL AIRSAR system and combined in a mosaic to cover the ecosystem process modeling subgrid. Eight classes were separated in the SAR image, and the classification accuracy was performed at several levels. Over 19 forest stands surveyed during the BOREAS field campaigns, the SAR map exhibited an accuracy of about 90%. The

analysis showed that the map was also able to correctly predict tree species composition on the SAR pixel scale. At a larger scale, an area of 25×35 km from the SAR map was compared with a digital vegetation map based on infrared aerial photography, and more than 77% of the total area was classified accurately. Finally, the number of classes were reduced to produce a map compatible with the requirements of the BOREAS land surface process models. The reduced map had five classes and, when compared with the vegetation map, showed similar land-cover patterns with greater than 92% accuracy over the total area. It is important to note that the classification accuracies performed in this study were highly dependent on the accuracy of the image calibration and impediments caused by errors due to incidence angle effects, aircraft motion compensation, and the image mosaic procedure. Furthermore, the results were obtained by using data from a single date. Multitemporal data can provide information about the seasonal and environmental states of the boreal forest and enhance the characteristics of the feature space for the classifier. Therefore, we believe that the accuracy obtained in this study is conservative and can be improved by incorporating multitemporal data and spaceborne systems with better image fidelity.

Some of the important results of the SAR classification were the separation of black spruce and jack pine stands and of coniferous and deciduous trees with close to 100% accuracy. These forest types are considered the dominant coniferous and deciduous stands covering large patches throughout the entire region of the boreal forest. The results also have a significant effect on modeling the canopy assimilation and biogeochemical processes for the region. Deciduous trees, because of their phenological, understory, and seasonal characteristics, represent different functional forms in ecosystem process models. Among conifers, jack pine and black spruce trees are also treated differently in process models. Unlike the dry and sandy soils of jack pine stands, the soils of black spruce patches are often covered by a thick moss layer and are poorly drained; black spruce patches also have different characteristics due to the release of trace gases from the soil surface and canopy.

The authors wish to thank Jaime Nickeson and the BOREAS staff at NASA/GSFC for providing the digital vegetation map and ground-truth data, respectively, and the AIRSAR group at JPL for collecting and processing the images analyzed in this paper. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Anderson, D. W., and Ellis, J. G. (1976), The soils of the provincial forest reserves in the Prince Albert map area. Extension Publication No. 261, Saskatchewan Institute of Pedology, University of Saskatchewan, Saskatoon, Canada.

Bonan, G. (1993), Importance of leaf area index and forest type when estimating photosynthesis in boreal forests. *Remote Sens. Environ.* 43:303–314.

BOREAS Science Steering Committee (1990), Charting the boreal forest's role in global change. *EOS* 72:33–35.

Cimino, J., Brandani, A., Casey, D., Rabass, J., and Wall, S. D. (1986), Multiple incidence angle SIR-B experiment over Argentina: mapping of forest units. *IEEE Trans. Geosci. Remote Sens.* 27:36–45.

DeFries, R. S., and Townshend, J. R. G. (1994), NDVI-derived land cover classification at a global scale. *Int. J. Remote Sens.* 17:3567–3586.

Dickenson, R. E. (1994), Land processes in climate models. *Remote Sens. Environ.* 51:27–38.

Dobson, M. G., Ulaby, F. T., and Pierce, L. E. (1994), Land-cover classification and estimation of terrain attributes using synthetic aperture radar. *Remote Sens. Environ.* 51:199–214.

Duda, R. O., and Hart, P. E. (1973), *Pattern Classification and Scene Analysis*. Wiley-Interscience, New York.

Hall, F. G., Townshend, J. R., and Engman, E. T. (1995), Status of remote sensing algorithms for estimation of land surface state parameters. *Remote Sens. Environ.* 51:138–156.

Moghaddam, M., and Saatchi, S. S. (1995), Analysis of scattering mechanism in SAR imagery over boreal forest: results from BOREAS'93. *IEEE Trans. Geosci. Remote Sens.* 33:1290–1296.

Pierce, L. E., Sarabandi, K., and Ulaby, F. T. (1994), Knowledge-based classification of polarimetric SAR images. *IEEE Trans. Geosci. Remote Sens.* 31:1081–1086.

Ranson, K. J., and Sun, G. (1994), Northern forest classification using temporal multifrequency and multipolarimetric SAR images. *Remote Sens. Environ.* 47:142–153.

Rignot, E., and Chellappa, R. (1993), Maximum *a posteriori* classification of multifrequency, multilook, synthetic aperture radar intensity data. *J. Opt. Soc. Am.* 10:573–582.

Rignot, E., and Drinkwater, M. R. (1994), Winter sea-ice mapping from multi-parameter synthetic-aperture radar data. *J. Glaciol.* 40:31–46.

Rignot, E., Williams, C., Way, J., and Viereck, L. A. (1994), Mapping of forest types in Alaskan boreal forest using SAR imagery. *IEEE Trans. Geosci. Remote Sens.* 32:1051–1059.

Saatchi, S. S., Soares, J. V., and Alves, D. S. (1996), Mapping deforestation and land use in Amazon rainforest using SIR-C Imagery. *Remote Sens. Environ.* 59:191–202.

Sader, S. A. (1987), Forest biomass, canopy structure and species relationships with multipolarization L-band synthetic aperture radar data. *Photogramm. Eng. Remote Sens.* 53:193–202.

Sellers, P., et al. (1995), The boreal-ecosystem atmosphere study (BOREAS): an overview and early results from the 1994 field year. *Bull. Am. Meteorol. Soc.* 76:1549–1577.

Sellers, P. J., Tucker, C. J., Collatz, G. J., Los, S. O., Justice, C. O., Dazlich, D. A., and Randall, D. A. (1994), A global 1° by 1° NDVI data set for climate studies, 2: the generation of global fields of terrestrial biophysical parameters from the NDVI. *Int. J. Remote Sens.* 15:3519–3545.

Townshend, J. R. G., Justice, C. O., Li, W., Gurney, C., McManus, J. (1991), Global land cover classification by remote sensing: present capabilities and future possibilities. *Remote Sens. Environ.* 35:243–256.

Yueh, H. A., Swartz, A. A., Kong, J. A., Shin, R. T., and Novak, L. M. (1988), Bayes classification of terrain cover using normalized polarimetric data. *J. Geophys. Res.* 93:15261–15267.