Discriminating between Tropical Vegetation Formations Using Reconstructed High-Resolution Seasat-A Scatterometer Data

Perry J. Hardin and David G. Long

Abstract

Satellite wind scatterometers are microwave radars designed to measure nearsurface wind speed and direction over the oceans. This was the primary mission for the Seasat-A Scatterometer (SASS), which acquired 14.6-Ghz data from 50-km resolution cells during its three-month mission in 1978. However, an image reconstruction technique utilizing overlap in resolution cells from successive satellite orbits can improve that spatial resolution to 5 km over land. An experiment conducted on reconstructed imagery of central South America illustrates the potential of this imagery for tropical vegetation analysis. In a supervised discrimination experiment, radar backscatter coefficients from the reconstructed imagery were found adequately diverse to distinguish between humid forest, subhumid woodland, climatic savanna, and edaphic savanna with a high degree of confidence. The potential for deriving geophysical information from reconstructed scatterometer imagery of Earth's land surface is discussed.

Introduction

Satellite scatterometers are active microwave radar instruments originally designed to measure the radar backscatter of the ocean's surface under all-weather conditions. The first spaceborne scatterometer, designated S-193, flew as a part of the Skylab missions in 1973. Acquiring both ocean and land data at a frequency of 13.9 Ghz, the S-193 experiments demonstrated scatterometers were feasible instruments for spaceborne platforms (Sobti *et al.*, 1975). Between June and October of 1978, the Seasat-A Satellite Scatterometer (SASS) was able to obtain nearly continuous global coverage at a spatial unit cell resolution of 50 km until a catastrophic short-circuit in the satellite's electrical system terminated subsequent data acquisition. Although Seasat-A operated for only 99 days, the wind vector information derived from SASS data conclusively showed accurate wind velocity measurements could be made from scatterometers on polar orbiting platforms (Jones *et al.*, 1982).

Without exception, the primary mission for all satellite scatterometers has been the determination of ocean surface wind direction and velocity. This is not surprising, because most land imaging applications require spatial resolutions finer than the 50 or 25 km appropriate for large ocean expanses. However, this paper will present results of a discrimination experiment designed to determine whether SASS imagery reconstructed to a 5-km resolution possesses adequate information for discriminating between several types of tropical vegetation in central South America. As shown in Figure 1, the study area covers most of the central and south-central Brazilian states and territories, portions of Paraguay's *Gran Chaco*, eastern Bolivia, and the Guiana Highlands of Venezuela, Guyana, and Suriname. While the experiment focuses on the existing vegetation cover of this region, the goal of the paper is to encourage others to consider reconstructed scatterometer imagery as a data source for any medium-scale land imaging applications where backscatter at microwave frequencies is an important geophysical parameter.

Although the data used in this experiment originated with Seasat-A, several additional spaceborne scatterometers are either currently active or in various planning stages. Like Seasat-A, these instruments will acquire backscatter data over land. The European Space Agency (ESA) successfully launched its ERS-1 satellite into a quasi-polar mission-adjustable orbit in the summer of 1991. The instrument payload included the Active Microwave Instrument (AMI) which was capable of operating in a wind scatterometer mode (5.3 Ghz) for the production of wind vector products. ESA is also planning additional follow-on missions.

The NASA Scatterometer (NSCAT), designed for acquisition of data at 14 Ghz, is planned as part of the payload aboard the 1995 Advanced Earth Orbiting Satellite (ADEOS). As discussed by Naderi *et al.*(1992), NSCAT is heavily based on SASS with significant improvements to enhance backscatter measurement accuracy at a spatial resolution of 25 km. In order to support the Earth Observing System (EoS) mission, a 14.6 Ghz scatterometer (EoS Scatterometer) based on the NSCAT design has been selected for flight in the late 1990s.

The next section of the paper presents a very brief history of scatterometer imaging of land surfaces. The subsequent section provides some background information concerning the ground processing concept for high resolution

P.J. Hardin is with the Department of Geography, 690 H SWKT, Brigham Young University, Provo, UT 84602.

D.G. Long is with the Department of Electrical and Computer Engineering, 448 CG, Brigham Young University, Provo, UT 84602.

Photogrammetric Engineering & Remote Sensing, Vol. 60, No. 12, December 1994, pp. 1453–1462.

0099-1112/94/6012-1453$3.00/0

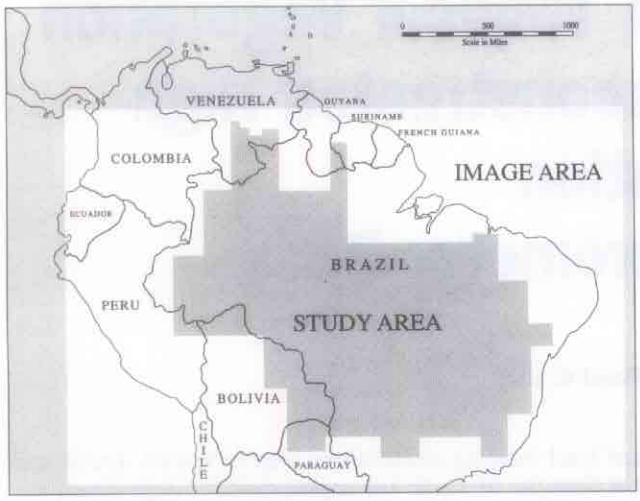

Figure 1. The study area.

scatterometry and the production of backscatter coefficients used in the investigation of tropical land cover types. Because the reconstruction process is described intricately elsewhere by Long *et al.* (1993), this section will be limited to concepts required to appreciate the sections following. The fourth section will report the methodology used to capture and compare backscatter signatures of several tropical land-cover classes from the high resolution data. The goal of this investigation was to assess the potential for utilizing reconstructed Ku-band scatterometry to discriminate between broad tropical vegetation classes. The results of the discrimination experiments will then be reported. Finally, conclusions will be drawn, and implications for future research will be suggested.

Scatterometers and Land Applications

Although primarily designed for oceanic wind studies, SASS data were also collected over global regions of ice and land during the ill-fated three-month flight of Seasat-A. Recently, Kennett and Li (1989a) were successful in mapping global SASS backscatter with a 111-km resolution (at the equator) from the original SASS 50-km resolution data. From this mapping, a rough, imperfect, but obvious agreement existed between global land cover and 14.6-Ghz backscatter. This was particularly evident in desert and equatorial forest regions where the resemblance of backscatter maps to published land-cover and geologic maps was striking.

While the potential of SASS data for broad land-cover mapping was apparent, most land imaging with scatterometers has historically focused only on identifying land-cover targets suitably large, homogeneous, diurnally invariant, and seasonally unchanging to calibrate future spaceborne scatterometers. In exploratory work using Skylab S-193 scatterometer data from several portions of the Earth, Sobti *et al.* (1975) found that, while 13.9-Ghz radar has very little surface penetration, backscatter at this frequency is very sensitive to vegetation cover, surface water, soil moisture, and physiography. These authors also concluded that targets defined *a priori* according to physiographic and land-use descriptions rather than expected microwave response suffered from a certain amount of confusion. As a result of studying incidence angle dependence and diurnal variation, Birrer *et al.* (1982) found the Amazon rain forest a homogeneous target which was azimuthally isotropic, insensitive to polarization, and which appeared to change only slightly from morning to afternoon. Bracalente and Sweet (1984) confirmed these earlier findings, and suggested important corrections to improve the estimation of backscatter from SASS data.

These earlier studies have all found support from Kennett and Li's (1989b) examination of 33 broad global land-cover types in a exhaustive effort to locate suitable scatterometer calibration sites. In conducting that effort, Kennett and Li (1989b) found a certain amount of overlap existing among the several land-cover categories, and substantial variability existing within the categories. Furthermore, while the early demise of Seasat-A precluded a conclusive examination of seasonal change in global land-cover backscatter, the three-month data set provided evidence that areas of greatest seasonality (e.g., arctic, deciduous forests) would exhibit the greatest variance in backscatter throughout the year. The converse would also likely hold true—backscatter coefficients in areas of little seasonality such as the central Amazon rain forest would probably deviate little through the months. In conclusion, because the Amazon rain forest was large, appeared azimuthally isotropic, was insensitive to polarization, presented little diurnal change, and could be expected to change less throughout the year than other large mid-latitude areas, it was deemed the most suitable target for future calibration of scatterometers in support of their ocean imaging mission.

A Concept for High Resolution Scatterometry

The SASS Design

Spaceborne scatterometers transmit microwave pulses to the ocean surface and measure backscattered power received at the instrument, allowing estimation of the normalized radar cross section (σ^o) of the surface. In simple terms, the radar cross section can be considered to be target albedo at radar frequencies. From each illuminated location on the Earth, the total power received by a radar is the sum of the power backscattered by the target, noise from the frequency-specific natural emissivity of the Earth-atmosphere system, and noise from the instrument itself. Once noise is subtracted from the total received power, σ^o can then be calculated using the basic radar equation (see Fung and Ulaby, 1983).

A simplified diagram of the SASS antenna illumination pattern is illustrated in Figure 2. As the satellite revolves in orbit, two sets of antennas are used to illuminate 475-km swaths on both sides of the satellite track. Moment by moment, each antenna produces an instantaneous footprint which is several hundred kilometres long but only a few kilometres wide. The 475-km swath swept by these antennas is then further resolved into smaller observation cells. A SASS-class scatterometer achieves along-track resolution by a combination of narrow antenna pattern and timing of transmit pulses. Cross track resolution is obtained by Doppler filtering, the narrow beam pattern, and the antenna illumination geometry (see Long *et al.*, 1988).

The antennas on each side of the instrument are arranged at two different azimuth angles. As the satellite orbits, a resolution element on the surface of the Earth is observed first by the forward looking antenna and then by the aft antenna a minute or two later. This produces two sets of co-located, nearly simultaneous observation pairs from

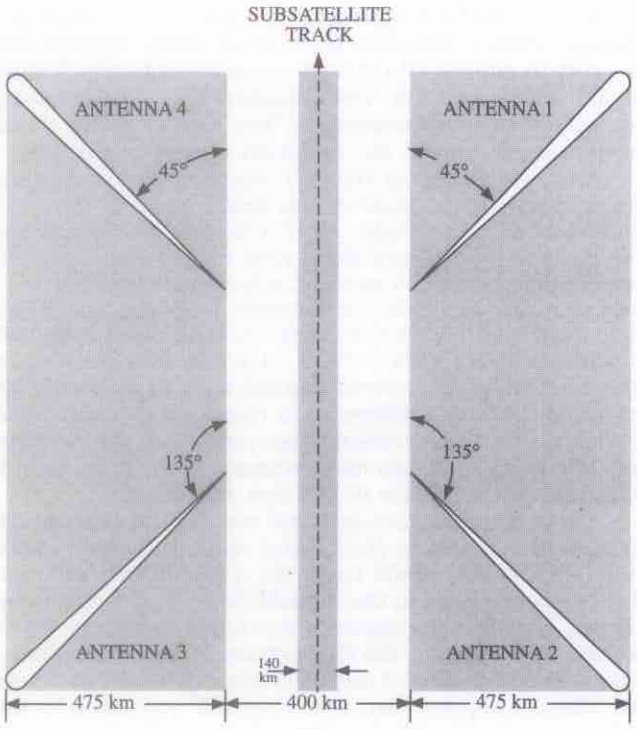

Figure 2. A simplified diagram of the SASS antenna illumination pattern. The four antennas are arranged so the same location on the Earth is observed first by the forward looking antenna and then by the aft antenna a few minutes later. Data from the near-nadir swath, consisting of three additional doppler cells, were not used in this study.

two different azimuth angles. The result of this arrangement is a set of vectors—multiple σ^0 measurements made of each cell over an irregular 50-km by 50-km along-track/cross-track sampling grid. The purpose of these multiple measurements is to derive surface wind speed and direction using a geophysical model relating ocean surface roughness to wind velocity (see Ulaby *et al.*, 1982; 1986).

Processing for High Resolution

To understand the process for obtaining higher resolution from the scatterometer imagery, consider a grid of high resolution elements on the Earth's surface. Each element of this grid has a cross-track dimension of s_c km and an along-track dimension of s_a km relative to the nadir track. Imperfectly imposed on this grid is a second irregular grid of the σ^0 measurements. Each cell of this second grid has corresponding dimensions greater than the first (i.e., it is lower resolution). The goal of high resolution processing is to estimate parameters of the smaller cells from the σ^0 measurements of the larger cells.

The observed σ^0 of a given resolution element depends on (1) the target character such as roughness, conductivity, etc.; and (2) the incidence angle and azimuth angle of the observation. As discussed by Birrer *et al.* (1982), when it can be assumed there is no azimuthal modulation of σ^0, a commonly accepted model for the incidence angle dependence of σ^0 is

$$10 \log_{10} \sigma^0(\theta) = A + B(\theta - 40°)$$

or alternatively

$$\sigma^0(\theta) = \alpha^A[\beta(\theta)]^B$$

where A and B are independent of the incidence angle θ and dependent on the target character, and where β is dependent on θ. It should be noted that A is the value of σ^0 where θ equals 40°. The constant α is requisite for the exponential expression of the logarithmic relationships and corresponds to the value $10^{1/10}$.

A given σ^0 measurement cell for the i^{th} cross-track cell and the k^{th} antenna beam, i.e., σ_k (i) will partially or completely cover a number of the smaller resolution elements. In the noise-free case, the total power σ_k received by the scatterometer for this σ^0 cell will be the sum of the echo power returned from each theoretical high resolution element. This can be written

$$\sigma_k = \sum_{c=c_{1,k}}^{c_{2,k}} \sum_{a=a_{1,k}}^{a_{2,k}} h(c,a;k)\sigma^0(c,a),$$

where the weighting function $h(c,a;k)$ is a function of the radar parameters (e.g., antenna gain, slant range) computed for each resolution element for the particular antenna beam k. These parameters are also assumed constant over any single high resolution element. Ignoring uncertainties in calibration and spacecraft attitude, $h(c,a;k)$ is a known constant and this previous equation can be written in terms of A and B as

$$\sigma_k = \sum_{c=c_{1,k}}^{c_{2,k}} \sum_{a=a_{1,k}}^{a_{2,k}} h(c,a;k)\alpha^{A(c,a)}[\beta(c,a;k)]^{B(c,a)}$$

Multiple measurements of σ_k which overlap the desired cell in the high-resolution grid can be used to estimate A and B. Two images of high-resolution cells are produced in the process, one each for A and B. The estimation problem for a single cell in the high-resolution grid can be summarized by the following equations. Given that $\beta(\theta_k(c,a))$ and $h(c,a;k)$ are constants for a single σ^0 measurement cell, assuming a reasonable approximation that $\theta_k(c,a) = \theta_k$, and being provided with multiple overlapping low-resolution SASS measurements $\sigma(i)$, the problem is to estimate A and B where

$$\sigma_k^0(c,a) = \alpha^{A(c,a)}[\beta(\theta_k(c,a))]^{B(c,a)}, \text{ with}$$

$$\sigma_k = \sum_{c=c_{1,k}}^{c_{2,k}} \sum_{a=a_{1,k}}^{a_{2,k}} h(c,a;k)\sigma_k^0(c,a).$$

The interested reader can consult Long *et al.* (1992; 1993) for a complete discussion of the numerical approaches to the actual calculation of the A and B for each element, discussions concerning noise, and equations describing variance in the estimates.

Applying the High Resolution Concept to SASS Data

The key to achieving high resolution by signal processing is overlap in the SASS σ^0 measurements. However, the SASS σ^0 measurements made by a single antenna on a single beam during one orbit have no overlap, and the forward and aft beam measurements are overlapped only a small amount. Furthermore, when this lack of overlap is combined with the problem of small gaps in the surface coverage for a single orbit, the result is insufficient data to attempt the resolution

improvement scheme with data from a single SASS pass. However, due to precession in the spacecraft orbit, a point on the Earth's surface was observed from the same beam with slightly different azimuth and incidence angles about once every three days. Data were also acquired by the instrument on both ascending and descending orbital nodes. The resulting set of criss-crossing measurement swaths in a single orbit combined from several orbits provides ample overlap to facilitate accurate estimation of A and B over much of the Earth.

The combination of multiple-pass co-located SASS data into a single set was pioneered by Bracalente and Sweet (1984) and utilized by Kennett and Li (1989a; 1989b), although neither group of researchers attempted to process the data for improved resolution. Instead, the goal of combining several orbits of data into a single set of σ^o measurements was to calculate various statistics for sensor calibration purposes.

In order to obtain a high-resolution reconstructed image of backscatter coefficients suitable for land cover analysis, the target should satisfy several criteria.

- The dependence of σ^o on the time of day should be small or correctable.
- The dependence of σ^o on the azimuth angle should be small or known. In the experiment described later, we have assumed there is no azimuthal modulation in σ^o.
- The A and B coefficients in the equation which summarize the incidence angle dependence of σ^o (i.e., $10\log_{10}\sigma^o(\theta) = A + B\,(\theta - 40°)$) should remain constant for the data acquisition period. In practical terms, this implies that the landscape should not change during the acquisition period.
- Heavy continuous rainfall should not persist for the data acquisition period.

Meeting these conditions is essential for generating a meaningful high-resolution image with minimum variation within land-cover classes and maximum statistical distance between the classes. Two more constraints must be added which are peculiar to the image reconstruction itself.

- The topography should be relatively flat. For reasons described below, we have assumed the topography effects are negligible.
- Over the range of incidence angles used for the reconstruction, the linear model $10\log_{10}\sigma^o(\theta) = A + B\,(\theta - 40°)$ adequately describes the incidence angle dependence of σ^o for the target.

While some of these issues have been explored by other researchers, the exact nature of these dependencies at 14.6 Ghz is only poorly understood for most land-cover types. Birrer *et al.* (1982) found that a small diurnal change did occur in the Amazon rain forest, with sunrise measurements approximately 1 dB higher than measurements acquired at different times of the day. During the 99 days of the Seasat mission, no change in the diurnal relations was noted.

Changes in global backscatter over the 99-day image acquisition period were assessed by Kennett and Li (1989a) by (1) dividing the globe into a grid of 1° by 1° mapping cells, (2) calculating a 99-day mean σ^o for each of the cells, (3) calculating mean σ^o values for each cell during successive two-week periods, and (4) mapping the deviation from the 99-day mean for each two-week segment of the mission life. Based on these maps, the period spanning 13 July to 30 July, Kennett and Li (1989) found some extremely small scattered savanna areas in central South America with σ^o values above the 99-day mean by 1.5 dB. Given the change in the savanna which occurs as the dry season progresses, this is not surprising.

It is common knowledge that rainfall can attenuate backscatter at 14.6-Ghz frequencies. In a preliminary analysis of Amazon region SASS data, Birrer *et al.* (1982) utilized GOES imagery to remove all the SASS measurements which showed significant cloud cover. The procedure was quickly deemed unnecessary and abandoned, as "few measurements should have been corrupted" due to the dry season acquisition.

Our own empirical evidence supports the conclusions of Birrer *et al.* (1982). Seasat-A was launched on 27 June 1978 and operated successfully until 10 October 1978. SASS measurements of the Amazon Basin were made twice daily. During the first part of the mission, acquisitions were made between 0500 and 0630 and between 1600 and 1830. During the last part of the mission, acquisition occurred from 0900 to 1200 and from 2100 to 2400. Given the late forenoon and afternoon nature of convective precipitation, we would have expected the largest difference in measured σ^o from rainfall contamination to be evident when comparing the nighttime and afternoon SASS data measurements of the same area. We found no such evidence in any area examined.

We would also have expected any rainfall contamination to be most apparent in the tropical rain forest areas. Statistically, this rainfall would lower the A coefficients and produce high variances in the A coefficients for the rain forest classes. However, the opposite was found to be true. As described in the results, the forest classes have the smallest variances and highest A means of any vegetation group studied. Thus, it appears that, even if rainfall contaminates the σ^o measurements, the effect on the A imagery is minor.

To understand how only minor contamination from rain is detectable in an equatorial region renown for its *rain* forest, it is important to remember that the reconstruction is primarily a weighted averaging function utilizing several measurements. Because multiple passes of data are averaged, rainfall would need to persist for dozens of days during the Seasat-A mission over the same high-resolution pixel being reconstructed. In addition, as detailed above, Seasat-A's orbit changed and data were collected at different solar times. For any significant contamination to occur, the rain would be required to fall at dawn and late afternoon during the first part of the mission, and then switch to late morning and late night until the mission concluded. These preceding two conditions are unlikely during the driest season of the year where convective precipitation is a sporadic episode, spatially localized, and fundamentally a high-sun event on any specified day. It is a common misconception that rain falls daily in the Amazon Basin. For example, a weather station in Manaus, Brazil, located at a latitude of 3°S in the heart of the Amazon Basin, records a yearly normal rainfall of 200 cm. Only 16 cm (8 percent) of this falls between 1 July and 31 September. Other stations within the study area show similar dry season patterns during the lifetime of Seasat-A.

Trying to eliminate topography's effect is somewhat problematic because, as mentioned before, the interaction between azimuth angle, incidence angle, and topographic factors (e.g., elevation, slope, relative relief) are very poorly understood, especially at several-kilometre resolutions and 14.6-Ghz frequencies. While elevations in the study area extend from nearly sea level to 900 metres in the Venezuelan Highlands, the primary topographic problem in the reconstruction is terrain with severe local relative relief over large portions of the 50-km by 50-km SASS σ^o measurements.

As mentioned previously, A and B coefficient images are estimated in the reconstruction process. While A is the incidence angle normalized σ^o, the B coefficient describes the in-

cidence angle dependence of σ^o. In areas where terrain is very steep, the actual measurement incidence angle may be slightly different when viewed from different azimuth angles. This will result in errors in the A and B estimates. However, noting that B within the study area is very small (< 0.15), the resulting A errors from topography would be minimal and appear as variance in any calculated statistical moments.

Precedent, our own evidence, and computational constraints guided our selection of the linear model $10\log_{10} \sigma^o(\theta) = A + B\,(\theta - 40°)$ in preference to higher order models to describe the incidence angle dependence of σ^o for the central South American study area. Birrer *et al.* (1982) stated that the simple linear fit described the incidence angle dependence of σ^o extremely well in the Amazon rain forest. Our own graphical analysis of σ^o versus incidence angle (θ) demonstrates that the linear relationship is particularly valid over the narrow incidence angle range between 23° and 55°. As described below, we utilized only the SASS measurements within that extremely narrow incidence angle range for the reconstruction. As a practical matter, models of higher order would require estimation of more parameter images beyond the A and B images produced by the simple linear model. The computational complexity and resources required to estimate A and B are already formidable — the cost of reconstructing higher order coefficient images would likely prove prohibitive. Furthermore, given our experience, it is likely that coefficient images of higher order would become progressively degraded in quality.

A Cover Discrimination Experiment

Figures 3 and 4 are A coefficient images which illustrate the resolution improvement of the new reconstruction method when applied to central South America. Figure 3 is a 1/2° by 1/2° resolution image (similar to those generated by Kennett and Li (1989a)) illustrating the inherent spatial resolution of the SASS radar. Figure 4 is a reconstructed A image with a resolution element size of 0.04° by 0.04° generated using the SIRF algorithm described in detail by Long *et al.* (1993). The B image contains very little information and is not shown. Three months of data (1 July through 10 October 1978) were used to generate both these images. While SASS was capable of making both vertically and horizontally polarized measurements of the radar backscatter, in these images only vertically polarized data were used. Additionally, orbits for which the spacecraft attitude determination was in error (see Bracalente and Sweet, 1984) were excluded. Furthermore, as mentioned above, only σ^o measurements in the incidence angle range from 23° to 55° with noise below a certain predetermined level were used as data for the reconstruction.

Figure 3. The A coefficient image with a resolution element size of 1/2° by 1/2° illustrating the inherent spatial resolution of the SASS radar.

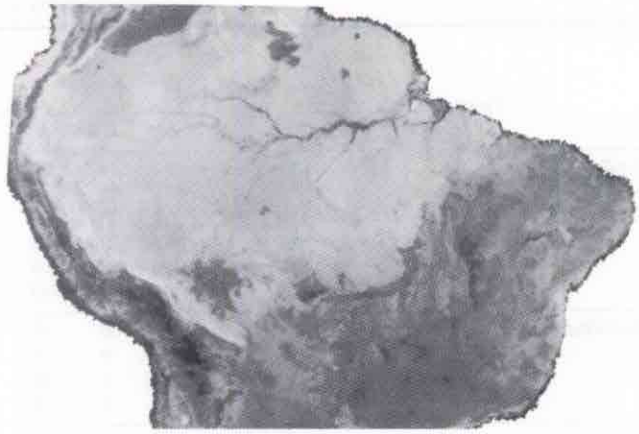

Figure 4. The reconstructed A coefficient image with a resolution element size of 0.04° by 0.04°. Note the finer detail visible along the course of the Amazon River when compared with Figure 3.

A simple experiment was devised to test the utility of the reconstructed imagery for discriminating between broad vegetation classes within the study area of central South America. Using the 1:5,000,000-scale *Vegetation Map of South America* (UNESCO, 1980), large homogeneous rectangular regions for 18 different vegetation formations were delimited. Realizing the limitations of the map source material, the inexact nature of cartographic classification, unavoidable generalizations, and the difficulties in drawing exact boundaries for vegetation classes which likely blend in transition zones, the rectangular areas were delimited only for the larger vegetation map polygons. Thirty-two rectangular regions were delineated, totaling over 740,000 km². The smallest solitary region (3500 km²) consisted of degraded tropical forest formations, while the largest single region (81,200 km²) consisted of Brazilian, Paraguayan, and Bolivian *Chaco*. The average region size was 23,136 km² — this corresponds to approximately 925 high resolution pixels. Once these regions were defined, the high resolution A and B images for central South America were constructed, and pixel values were then extracted for each of them and saved for further analysis. In the discrimination tests described below, only one half of the pixels for each of the 18 formations were used for training. The remaining half were reserved for testing the discrimination.

Discrimination Methodology

To establish a baseline estimate of discrimination ability, there was an initial analysis to classify the 18 vegetation formations in the test set using quadratic discriminant functions generated from the training set classes. If we designate the A and B coefficients of a single unassigned pixel (j) as the vec-

TABLE 1. CONFUSION MATRIX FOR THE 18-GROUP SUPERVISED DISCRIMINATION EXPERIMENT. FORMATION NAMES WERE OMITTED FROM THE TABLE TO SAVE SPACE, BUT THE IDENTITY OF EACH IS GIVEN IN TABLE 3.

		Actual Class																	
		Forest							Woodland			Climatic Savanna			Edaphic Savanna				
	Predicted Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Forest	1	126	149	89	34	41	73	0	0	0	0	0	0	0	98	0	0	0	0
	2	90	1120	52	719	62	110	2	0	0	0	0	0	0	54	7	0	0	0
	3	5	0	92	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	4	11	767	17	4149	0	5	14	0	0	0	0	0	60	3	0	0	0	0
	5	6	35	4	0	8	14	1	0	0	0	0	0	0	13	4	3	0	0
	6	0	22	4	35	0	33	17	0	17	0	0	0	0	3	0	6	0	0
Woodland	7	0	7	9	16	8	27	398	0	89	19	79	0	0	96	29	63	68	4
	8	0	0	0	9	0	3	0	2507	116	2	0	0	0	0	0	0	0	12
	9	0	0	0	2	0	35	5	58	335	14	11	0	0	0	0	41	0	1
	10	0	0	0	0	0	0	0	21	194	103	92	0	30	8	0	12	0	72
Climatic Savanna	11	0	0	17	0	3	0	43	0	133	28	1220	7	92	96	34	473	372	227
	12	0	0	0	0	0	0	0	0	0	0	18	280	0	171	85	17	158	88
	13	0	0	0	0	0	0	0	0	0	0	30	0	40	4	7	62	12	77
Edaphic Savanna	14	0	0	0	0	0	0	20	0	31	0	0	7	0	174	20	46	21	141
	15	0	0	0	0	0	0	0	0	2	0	0	0	0	13	1	0	0	11
	16	0	0	0	0	0	0	0	0	0	0	28	0	41	2	29	42	65	36
	17	0	0	8	0	0	0	0	0	13	0	301	48	27	123	85	83	346	138
	18	0	0	0	0	0	0	0	6	3	2	182	0	104	157	21	210	230	211

tor $\mathbf{x}_j$, and the individual vegetation group mean vectors, group covariance matrices, and group covariance matrix determinants as μ, ζ, and d, respectively, the equation for the quadratic discriminant function distance Q between unlabeled pixel j and vegetation group h is given by

$$Q_{jh} = (\mathbf{x}_j - \mu_h)^T \zeta_h^{-1} (\mathbf{x}_j - \mu_h) + \ln(d_h)$$

Like other classifiers of this genre, once the function calculations have been made for all groups, the pixel is assigned to the group which generates the smallest distance.

After this baseline classification was established, the original 18 categories were continually regrouped into a much smaller set until the classification accuracy produced by the quadratic discriminant functions reached an acceptable level. However, potential groupings which would require tenuous interpretations were not allowed. For example, the floristic formations designated as forest types on the UNESCO map were never mixed with members of the grassland category. Conversely, some of members in the woodland category might logically be grouped with scrub formations without insurmountable interpretation problems.

In the ultimate discriminant analysis, the goal was to determine whether sufficient statistical distance in the A and B coefficient data existed between the final supervised land-cover groupings for successful land-cover mapping. As in all the previous analyses, quadratic discriminant functions were calculated for each of the final groups using the training data set alone. These functions were then used to classify each pixel from the testing set.

The agreement between each pixel's predicted and actual group was summarized in confusion matrix form for accuracy assessment. The quality of the supervised classification was assessed by calculating both an overall Cohen's (1960) kappa (κ) measure and a simple percentage.

TABLE 2. CONFUSION MATRIX PRODUCED FROM THE FOUR-GROUP SUPERVISED DISCRIMINATION EXPERIMENT.

	Actual Class			
Predicted Class	Forest	Woodland	Climatic Savanna	Edaphic Savanna
Forest	7903 98%	108 3%	0 0%	225 5%
Woodland	95 1%	3599 86%	117 4%	151 3%
Climatic Savanna	9 <1%	201 5%	1064 40%	935 20%
Edaphic Savanna	23 <1%	285 7%	1456 55%	3274 71%
Total	8030 100%	4193 100%	2637 100%	4585 100%

Findings and Discussion

Supervised Discrimination: Membership and Accuracy

Attempts to classify the original 18 categories without any regrouping produced a classification accuracy of 57 percent ($\kappa = 0.52$). From the confusion matrix (Table 1) and graphical analysis, it was clear that substantial backscatter similarity existed within groups roughly considered forest, woodland, and savanna. However, there was apparently good discrimination between these broad groups. Logically regrouping the 18 formations into four groups improved the classification accuracy to 81 percent ($\kappa = 0.74$). In this case the four categories consisted of humid forest, tropical woodlands, climatic savannas, and edaphic savannas (Table 2). Table 3 summarizes the A and B information for each of the original 18 cover categories as subdivided into these four categories. Combining the edaphic and climatic savanna groups into a single savanna category generated a marked improvement (94 percent, $\kappa = 0.91$). Table 4 presents the confusion

TABLE 3. SUMMARY INFORMATION FOR THE 18 VEGETATION FORMATIONS SAMPLED FROM THE VEGETATION MAP OF SOUTH AMERICA. THE NUMBERS IN PARENTHESIS AFTER EACH FORMATION NAME MAY BE USED TO CROSS-REFERENCE THE GROUP TO THE APPROPRIATE ROW AND COLUMN OF TABLE 1.

	Vegetation Formation	Class *n*	Mean A	Stdv. A	Mean B	Stdv. B
Forest	Extremely moist ombrophilous forest (1)	238	−7.19	0.193	−0.124	0.002
	Moist evergreen seasonal forest (2)	2100	−7.54	0.203	−0.122	0.003
	Ombrophilous submontane forest (3)	293	−7.50	1.006	−0.121	0.002
	Very moist ombrophilous forest (4)	4978	−7.67	0.203	−0.118	0.002
	Degraded forest formations (5)	121	−7.54	0.400	−0.125	0.004
	Tropical evergreen seasonal lowland forest (6)	300	−7.84	0.589	−0.123	0.003
Woodland	Degraded woodland formations (7)	499	−8.59	0.373	−0.124	0.002
	Chaco (8)	2592	−9.46	0.494	−0.106	0.003
	Caatinga (9)	933	−9.30	0.696	−0.118	0.007
	Degraded caatinga woodland formations (10)	168	−9.82	0.472	−0.116	0.003
Climatic Savanna	Campo cerrado N (11)	1961	−10.04	0.681	−0.122	0.004
	Degraded subhumid campo cerrado and crops (12)	342	−11.18	0.491	−0.131	0.002
	Degraded lowland campo cerrado (13)	334	−10.65	0.563	0.119	0.002
Edaphic Savanna	Campos sujo and limpo (14)	1015	−10.14	1.641	−0.127	0.004
	Grassland with palms (15)	321	−10.77	1.126	−0.127	0.003
	Pantanal (16)	1058	−10.33	0.938	−0.122	0.002
	Cultivated crops (17)	1273	−10.43	0.903	−0.125	0.003
	Campo cerrado S (18)	918	−10.94	0.734	−0.121	0.004

matrix resulting from the three group discrimination experiment.

As mentioned previously, the first general group with the highest *A* coefficients was forest. It consisted of two moist tropical ombrophilous forest types, moist evergreen seasonal forest, tropical evergreen seasonal lowland forest, ombrophilous submontane forest, and associated degraded forest mosaics.

The second general group consisted of four woodland vegetation formations, associated mosaics, and degraded woodland landscapes. The specific woodland types included drought deciduous woodland known as *chaco*, thorn forest with succulents/*caatinga*, related degraded woodland formations/mosaics, and degraded *caatinga* formations.

The savanna group consisted of several separate shrub, scrub, and grassland categories. These included two floristic formations of *campo cerrado, campo sujo/limpo*, grassland with palms, and seasonally flooded grassland or *pantanal*. Two varieties of degraded formations and cultivated crops were also members of this category. With some loss of accuracy, the savannas could be further subdivided into edaphic and climatic savanna categories. In the four-group discrimination experiment, the climatic savannas consisted of degraded and non-degraded *campo cerrado* formations whereas the edaphic savanna formations consisted of primarily of grasslands, wet/flooded grasslands, and associated adjacent southern *campo cerrado*.

The first confusion matrix (Table 1) indicates the greatest confusion problems between the forest and woodland groups to be primarily a result of backscatter overlap between degraded woodland formations, *caatinga*, and tropical evergreen seasonal lowland forest. The small amount of confusion remaining in the table exists primarily between the woodland and savanna groups. While there is some backscatter resemblance between most members of the woodland and savanna groups, the two degraded woodland formations appear to be major culprits in several instances. This confusion reflects the common remote sensing task of distinguishing between woodland and shrubland vegetation when both are so heterogeneous with variable canopy densities, species compositions, plant spacing, tree heights, and canopy response to seasonal factors. For example, there is substantial confusion between the degraded woodland formation and the *campo sujo/limpo* savanna. It is highly likely that a woodland degraded through cattle grazing, fire, and agriculture would exhibit a great variance of backscatter dependent on the areal extent of the degradation, its severity, or the successional stage of any sub-climax regrowth. Thus, it is entirely reasonable that some regions of degraded woodland would exhibit similar backscatter characteristics to the shrubby grassland of *campo sujo* or the pure grassland of *campo limpo*.

Summary

When classified using supervised methods, it appears reconstructed SASS imagery has potential for discriminating between the broad tropical classes of humid forest, subhumid woodlands, and savanna. There is also some ability to distinguish between climatic and edaphic savanna formations.

In conducting this experiment, the limitations of using a

TABLE 4. CONFUSION MATRIX PRODUCED FROM THE THREE-GROUP SUPERVISED DISCRIMINATION EXPERIMENT.

	Actual Class		
Predicted Class	Forest	Woodland	Grass-Shrubland
Forest	7892 98%	88 2%	217 3%
Woodland	110 1%	3763 90%	377 5%
Grass-Shrubland	28 <1%	342 8%	6628 92%
Total	8030 100%	4193 100%	7222 100%

map as ground truth for understanding the interaction between 14.6-Ghz scatterometry and tropical vegetation became readily apparent. By its very nature, traditional vegetation mapping of large regions at small scales requires generalization, simplification, discreet boundaries, and thus hides the complexity of the heterogeneous mixtures that occur in reality. This mixture was readily detected in processing the reconstructed image. Because of this heterogeneity, only three or four broad classes were capable of being distinguished using an *a priori* classification scheme, and further subdivision of the classes was difficult.

Only a general statement of scattering differences between forest, woodland, and savanna can be offered because, unfortunately, 14.6-Ghz vegetation scattering has been studied very little. It has never been studied at 5-km pixel scales where mixed pixels present interpretation problems. Unlike lower frequency radars where volume scattering in vegetation dominates, Ku-band radar primarily scatters off the surface of heavily leafed canopies. The high density and random orientation and spacing of the leaves in the tropical forest produce diffuse scattering which results in bright signature at the narrow range of incidence angles used in the project. The extremely moist, heavily canopied, tropical ombrophilous forest serves as an excellent example of this strong backscattering response. On the other hand, substantial losses are imposed from multiple reflections in the vertically structured grass component of savannas. As a result, savannas dominated by grassland such as *campo limpo* will appear very dark. Park-like savannas and woodlands will demonstrate a mixture of the two scattering mechanisms and will have an intermediate backscatter response proportional to the relative amounts of canopied overstory, grassland, and bare ground within the pixel. Areas of flooded grassland, woodland, or forest will appear darker than their dry counterparts because of water's high dielectric constant.

Final Comments

Given the success of these discrimination experiments, further research into tropical land scatterometry may bring substantial payoffs. While the goal of this project was not to create a general 1978 vegetation map of central South America, it may be possible to do so with SASS data. In such a project, much of the confusion in the purely backscatter-based classification could be removed by incorporating topographic, climatic, pedologic, and other ancillary data. Inclusion of AVHRR data collected during the same time period is also possible. If the effect of topography on 14.6-Ghz backscatter can be controlled, it may be possible to extend the procedure using the global SASS data set to all the tropical regions of the world in order to estimate and map the 1978 extent of the global tropical forests into broad categories. With the launch of ADEOS in 1995 and the high resolution ground processing of NSCAT imagery using similar techniques, the potential may exist to assess change in the tropical forest character and extent over the 17-year interval at identical spatial resolutions and frequencies. Unlike the SASS data which covered only three months, NSCAT coverage of the equatorial and savanna regions for the entire year would also allow change in backscatter through the seasons to be used as a discriminator for more accurate vegetation mapping. It would also allow real-time monitoring of global tropical seasonal changes without interference from cloud cover which afflicts visible and near-infrared instruments.

Currently, there is a substantial amount of research being conducted to map the tropical rain forests of the world using NOAA AVHRR and ERS-1 SAR data (TREES, 1991). In this endeavor, reconstructed SASS imagery of the world's global rain forests can provide another important historical data source. Furthermore, when applied to current ERS-1 scatterometry, the reconstruction process may allow us to inexpensively extend high resolution ERS-1 SAR tropical forest canopy measurements from small test sites to larger continental regions.

However, the potential for utilizing reconstructed SASS (or other scatterometer) imagery goes beyond land-cover mapping or monitoring. The detection of several vegetation class gradations in the reconstructed imagery indicates the strong possibility of obtaining critical quantitative information about tropical vegetation community character from reconstructed scatterometer imagery. It is still not clear whether 14.6-Ghz backscatter coefficients change in response to canopy density, canopy vigor, another canopy structural characteristic, or more complex quantitative variables. Using higher resolution SAR, visible, thermal, and infrared imagery of carefully selected verification sites as primary data sources, it may be possible to describe backscatter-based classes not by simple name alone, but more precisely in terms of biomass, moisture content, transpiration rates, leaf area, or as a percentage of each species constituting the target. This information is as important to understanding global geophysical processes as simple nominal mapping of vegetation. As an example of this approach, our current research involves modeling AVHRR derived parameters (e.g., NDVI) as functions of reconstructed SASS A and B coefficients.

The SASS image reconstruction technique also holds great promise for land imaging of other regions of interest during the late summer months of 1978. These include the world's boreal forests, Greenland, and the Antarctic. The potential for imaging glacial ice is particularly fascinating, because scatterometer images provide multiple incidence angle observations. It may be possible to create several coregistered high-resolution images, each possessing a different range of incidence angles, to infer ice age, water content, snow cover, or other parameters of interest.

Given the potential for scatterometers to obtain useful information over land, it may be profitable for the scientific community to carefully consider both NSCAT and its follow-on EoS-era scatterometers as potential sources for land imaging. According to Long *et al.* (1993), minor modifications involving changes to the digital subsystem combined with the ground signal processing algorithm outlined above could potentially produce reconstructed image resolutions between 1 and 3 km. These changes would make NSCAT and the EoS-era scatterometers dual mission instruments which could provide ground imaging at 14 Ghz with a spatial resolution compatible with several other EoS instruments, including the Multifrequency Imaging Microwave Radiometer (MIMR), Multi-angle Imaging Spectro-Radiometer (MISR), and the Moderate Resolution Imaging Spectrometer (MODIS).

References

Birrer, I. J., E. M. Bracalente, G. J. Dome, J. Sweet, and G. Berthold, 1982. σ^0 Signature of the Amazon Rainforest Obtained from the Seasat Scatterometer, *IEEE Transactions on Geoscience and Remote Sensing*, GE-20(1):11–17.

Bourliere, F., 1983. *Tropical Savannas*, Vol. 13 of *Ecosystems of the World*, (D.W. Goodall, editor), Elsevier, Amsterdam.

Bracalente, E. M., and J. L. Sweet, 1984. *Analysis of Normalized Radar Cross Section (σ^o) Signature of Amazon Rain Forest Using Seasat Scatterometer Data.* NASA Technical Memorandum 85779, NASA NTIS.

Cohen, J., 1960. A Coefficient of Agreement of Nominal Scales, *Educational and Psychological Measurement*, 20:37–46.

Fung, A. K., and F. T. Ulaby, 1983. Matter-Energy Interaction in the Microwave Region, Chapter 4 in *Manual of Remote Sensing*, Vol 1. (R.N. Colwell, editor), American Society of Photogrammetry.

Jones, W. L., L. C. Schroeder, D. H. Boggs, E. M. Bracalente, R. A. Brown, G. J. Dame, W. J. Pierson, and F. J. Wentz, 1982. The Seasat-A Satellite Scatterometer: The Geophysical Evaluation of Remotely Sensed Winds over the Ocean. *Journal of Geophysical Research*, 87:3297–3317.

Kennett, R. G., and F. K. Li, 1989a. Seasat Over-Land Scatterometer Data, Part I: Global Overview of the Ku-Band Backscatter Coefficients, *IEEE Transactions on Geoscience and Remote Sensing*, GE-27(5):592–605.

———, 1989b. Seasat Over-Land Scatterometer Data, Part II: Selection of Extended Area Land-Target Sites or the Calibration of Spaceborne Scatterometers. *IEEE Transactions on Geoscience and Remote Sensing*, GE-27(6):779–788.

Lieth, H., and M. J. A. Werger, 1989. *Tropical Rain Forest Ecosystems: Biogeographical and Ecological Studies*, Vol. 14b of *Ecosystems of the World* (D.W. Goodall, editor), Elsevier, Amsterdam.

Long, D. G., C. -Y. Chi, and F. K. Li, 1988. The Design of an Onboard Digital Doppler Processor for a Spaceborne Scatterometer, *IEEE Transactions on Geoscience and Remote Sensing*, GE-26: 869–878.

Long, D. G., and P. J. Hardin, 1993. Vegetation Studies of the Amazon Basin Using Enhanced Resolution Seasat Scatterometer Data, *IEEE Transactions on Geoscience and Remote Sensing* (in review).

Long, D. G., P. J. Hardin, and P. T. Whiting, 1993. Resolution Enhancement of Spaceborne Scatterometer Data, *IEEE Transactions on Geoscience and Remote Sensing*, (in press).

Long, D. G., P. T. Whiting, and P. J. Hardin, 1992. High Resolution Land/Ice Imaging Using Seasat Scatterometer Measurements, *IGARSS '92 International Geoscience and Remote Sensing Symposium*, 1:440–442.

Naderi, F. M, M. H. Freilich, and D. G. Long, 1992. Spaceborne Radar Measurement of Wind Velocity over the Ocean: An Overview of the NSCAT Scatterometer System. *Proceedings of the IEEE*, 79(6):850–866.

Sobti, A., R. K. Moore, and F. T. Ulaby, 1975. *Backscatter Response at 13.9 GHZ for Major Terrain Types as Seen from Orbit*, Technical Report 243-4, NASA, Lyndon B. Johnson Space Center, Houston, Texas, August 15.

TREES, 1991. *Strategy Proposal 1991-1992.* TREES Technical Series A, No. 1, A Joint Project of the Commission of the European Communities and the European Space Agency, SPI.90.31, Joint Research Center, Ispra, Italy.

Ulaby, F. T., R. K. Moore, and A. K. Fung, 1982. *Microwave Remote Sensing: Active and Passive, Vol. I.* Addison-Wesley.

———, 1986. *Microwave Remote Sensing: Active and Passive, Vol. III.* Artech House.

UNESCO, 1980. *Vegetation Map of South America.* 1:5,000,000, two sheets, published by the United Nations Educational, Scientific and Cultural Organization, Paris.

———, 1981. *Vegetation Map of South America Explanatory Notes.* Published by the United Nations Educational, Scientific and Cultural Organization, Paris.

(Received 7 August 1992; revised and accepted 28 April 1993)

Appendix

Vegetation Definitions

The definitions below were adapted primarily from Lieth and Werger (1989), UNESCO (1981), and Bourliere (1983).

Caatinga. Semi-deciduous thorn scrub with succulents. It is found in northeastern Brazil in areas where the climate is subhumid to arid. It is bordered on the east by the northern variety of *campo cerrado* and on the east by the coastal Atlantic forest and related degraded formations. Climax *caatinga* is ligneous, drought-deciduous, thickly branched, narrow leaved, and sometimes thorned. The physiognomy of the *caatinga* is variable. In some areas it may resemble shrubland steppe, and in other regions may be a dense scrubland with canopies reaching 10 metres. Cactus may also be found.

Campo Cerrado (N). Medium-tall grassland with several species of broad-leaved evergreen trees and shrubs. This formation is found in central Brazil, and forms a zone between the Amazonian Forests to the east and north and the *caatinga* along its eastern border. Physiognomy of *campo cerrado* (N) varies from a shrub savanna to woodland savanna of light to dense canopy densities. Gallery forests and areas of grassland are also common. Shrubs and trees average 5 metres in height, have thick bark, twisted branches, and partially deciduous leaves which may be hairy and leathery. Anthropogenic change and fire in the *campo cerrado* (N) make the landscape extremely variable.

Campo Cerrado (S). Medium-tall grassland with several species of broad-leaved evergreen trees and shrubs. This formation is located in southern Brazil and north-central Paraguay. It is distinguished from *campo cerrado* (N) by its shorter dry season and a definite cool season. In general, it possesses a denser understory than the northern *campo cerrado*. Primarily grassland with sedges, shorter palms are sometimes common. Edaphic factors are particularly important to explaining the various facies of this formation.

Campo sujo. A prairie grassland with scattered shrubs. Within the study area of this paper, they are located in northwestern Brazil.

Campo limpo. Tall prairie grassland nearly devoid of shrubs and trees. Within the study area, these areas are mixed with *campo sujo* in northern Brazil.

Chaco. Drought deciduous lowland (and submontane) woodland. The *chaco* within the study area is found in Bolivia, along the plain between the Paraguay and Parana rivers in Paraguay, and along the Rio Paraguay in southern Brazil. This woodland is typically multi-canopied and deciduous, with cactus located in drier areas. There are many climatic and edaphic variations on a local scale.

Pantanal. A phytogeographical complex covering over 100,000 km² along the Paraguay River and its tributaries in Brazil, Bolivia, and Paraguay. The *pantanal* is primarily a humid savanna of hygrophytic grassland which is 80 percent flooded during the rainy season. In areas of higher local relative relief, this grassland is replaced by shrubs, woodland, and forest.

Perry J. Hardin
Perry Hardin is an Assistant Professor in the Geography Department at Brigham Young University and directs the BYU Laboratory for Geographic Information Analysis. Holding a Ph.D. in Geography (Remote Sensing) from the University of Utah, he is an investigator on several projects involving the integration of remote sensing and GIS for resource management. His research interests include biogeography, image processing theory, climatology, and habitat modeling.

David G. Long
David Long is an Assistant Professor in the Electrical and Computer Engineering Department at Brigham Young University. He has a Ph.D. in Electrical Engineering from the University of Southern California. He is the principle investigator on several NASA-sponsored interdisciplinary research projects in remote sensing and has numerous publications in signal processing and radar scatterometry. His research interests include microwave remote sensing, radar, signal processing, and mesoscale atmospheric dynamics.