LARS Contract Report 120180

NASA-CR-167792

E83-10149

Quarterly Progress Report

Evaluation of SLAR and Thematic Mapper MSS Data for Forest Cover Mapping Using Computer-Aided Analysis Techniques

Contract No. NAS 9-15889

Reporting Period: September 1, 1980 - November 30, 1980

Submitted to: Exploratory Investigations Branch
NASA Lyndon B. Johnson Space Center

Prepared by: Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47906

Technical Monitor: Mr. Norman Hatcher
NASA Mail Code SF5
Exploratory Investigations Branch
Houston, Texas 77058

Principal Investigator: Dr. Roger M. Hoffer
Ecosystems Program Leader
LARS/Purdue University
West Lafayette, Indiana 47906

(E83-10149) EVALUATION OF SLAR AND THEMATIC MAPPER MSS DATA FOR FOREST COVER MAPPING USING COMPUTER-AIDED ANALYSIS TECHNIQUES Quarterly Progress Report, 1 Sep. - 30 Nov. 1980 (Purdue Univ.) 21 p HC A02/MF A01

N83-16824

Unclas
G3/43 00149

TABLE OF CONTENTS

I. ACTIVITIES OF THE PAST QUARTER

A. Analysis of the NS-001 MSS Data

During this quarter, the major thrust of effort involved analysis of the four different spatial resolutions of the NS-001 MSS data. Analysis of the training data for this area indicated that 33 spectral classes would be adequate to characterize the various cover types present. The test area in Flight Line 1 south of Camden (referred to as CAM1S) was the first area on which a detailed analysis was conducted. This area contained 11 different informational classes as described in Table 1.

In order to facilitate the comparison between the different spatial resolution data sets (i.e., 15 m, 30 m, 45 m, and 80 m), a supervised approach was taken in defining training blocks for each of the different cover types. The training fields representing each cover type category were then grouped and this group of training fields were clustered in order to define the individual spectral classes within each cover type category and which would effectively characterize the entire test site. Care was required to ascertain that each of the spectral classes within each of the different cover types was adequately represented in terms of the number of pixels of data present, especially at the Landsat spatial resolution. In addition to briefly describing each of the informational categories or cover classes present in CAM1S, Table 1 also indicates the number of spectral classes representing each of the different cover class groups. Table 2 indicates the number of pixels present in each of the 33 spectral classes for each spatial resolution data set. Table 3 indicates the number of training fields that were defined for each of the cover class groups and also indicates the average number of pixels for each of the individual training fields, as a function of spatial resolution. As one can see from

Table 1. Description of the Cover Classes and Number of Spectral Classes within each Cover Class Defined for the CAMIS Study Area.

Cover Class	Number of Spectral Classes	Description of Cover Class
Tupe	3	Water tupelo; generally restricted to narrow ox-bow lakes and other areas of inundated soils.
Mveg	2	Misc. shrubs and small trees; located on saturated and inundated soils.
Crop	4	Row crops and small grain crops in varying stages of development and maturity.
Past	5	Pasture and old fields; plant cover varies from healthy, improved pasture grasses to senescent forbs and invader species.
Soil	3	Bare soil areas associated with agri-cultural activities; varies in sand, clay, and organic material content as well as moisture content.
Pihd	2	Pine-hardwood mix; generally varies between 35 to 65% hardwood intermixed with pine (determined by visual inspection).
Hdwd	3	Old age bottom-land hardwood; sweet gum is the dominant species, crowns are large, inter-crown gaps are generally deep and result in dark shadowed areas.
Ccut	4	Areas subjected to clearcut forestry practices; ground cover comprised of dry to inundated soils without vegetation, to dense vegetative cover of slash, grasses, shrubs and residual trees. Windrowed slash is common on these areas.
Sghd	3	Second growth hardwood; species composition is highly diverse, crown height and diameter is variable, inter-crown gaps are generally shallow and do not result in dark shadowed areas.
Pine	2	Pine forest areas; the principle species is slash; long-leaf, and loblolly are common; age class varies from recently planted (5-10 years) to mature, closed canopy.
Watr	2	Water; primarily associated with the Wateree River (approximately 70-90 meters in width). Other areas comprising the water class are associated with surface mining and open marsh.

Table 2. The Number of Pixels in each Spectral Class of each Cover Class, by Spatial Resolution (CAM1S).

Cluster Class	Spatial Resolution 15 Meter	30 Meter	45 Meter	80 Meter
Tupe 1	511	139	72	27
Tupe 2	452	104	36	20
Tupe 3	403	99	45	21
Mveg 1	658	158	68	29
Mveg 2	534	136	62	27
Crop 1	598	130	58	28
Crop 2	2887	746	312	152
Crop 3	1003	266	127	65
Crop 4	1227	299	126	54
Past 1	432	112	37	18
Past 2	572	164	70	61
Past 3	1154	296	127	21
Past 4	1233	303	137	68
Past 5	419	104	36	23
Soil 1	765	375	184	83
Soil 2	1919	909	428	187
Soil 3	1366	662	259	114
Pihd 1	246	72	28	16
Pihd 2	1015	242	115	45
Hdwd 1	1159	1319	693	335
Hdwd 2	1846	1701	656	268
Hdwd 3	1043	955	418	189
Ccut 1	771	714	335	157
Ccut 2	1480	1294	582	285
Ccut 3	1414	1445	634	280
Ccut 4	666	732	324	132
Sghd 1	1597	909	428	203
Sghd 2	1979	817	324	139
Sghd 3	757	396	187	93
Pine 1	1244	356	156	85
Pine 2	1946	429	205	72
Watr 1	925	215	*	11
Watr 2	164	39	121	53

*Spectral class was deleted due to an insufficient number of observations to compute the covariance.

Table 3. The Average Number of Pixels per Training Field for each Spatial Resolution, for each Cover Class in CAMIS.

Cover Class	No. of Training Fields	Spatial Resolution 15 Meter	30 Meter	45 Meter	80 Meter
Soil	35	223.0	55.6	25	11.0
Past	51	75.7	19.4	8.0	3.8
Crop	34	168.6	42.5	18.4	8.9
Pine	16	204.4	50.3	23.1	9.8
Pihd	4	318.2	78.5	35.7	15.2
Hdwd	17	926.2	235.1	104.8	46.6
Sghd	16	557.7	140.1	60.9	28.8
Tupe	17	82.0	20.6	9.1	4.1
Ccut	22	772	194.4	85.9	40.7
Mveg	2	596	147.0	65.0	28.0
Watr	10	182.7	42.8	20.3	11.1
Total	224	303.6	76.3	33.7	15.5

Table 3, the number of pixels present in many fields at the 80 meter resolution was very small for some of the cover type classes such as pasture, pine, tupelo and others, indicating the relatively small size of many of the individual fields and forest stands present in this area. However, since a large number of training fields were defined for most of the cover class groups, (except Pihd and Mveg), it was believed that a reasonably good representation of each cover class had been obtained.

The classification results for the training data set are summarized in Table 4 by cover class group and for each of the spatial resolutions. In order to evaluate the significance of possible differences in classification performance as a function of spatial resolution, a technique had to be defined which would adequately take into account the fact that there are different numbers of pixels involved for each of the four spatial resolutions for each of the different cover types. This was accomplished through the use of the harmonic mean, which is a weighted average, where the weight is proportional to the inverse of the relative magnitude of each element included in the average. The harmonic mean is, therefore, a mean value of lower magnitude than the arithmetic mean in every case where the elements are not equal (the harmonic mean equals the arithmetic mean where the elements are equal). The harmonic mean is regarded as more appropriate than the arithmetic mean for estimating a common variance among factor levels (eg., each resolution) sampled at different intensities, since the lowest sampling intensity has the greatest weight in determining the mean.

Table 4. Statistical Evaluation of Classification Performances by Cover Class for each Spatial Resolution (Training Field Pixels, Per-Point GML Classifier, CAMIS).†

Cover Class	Spatial Resolution: 15 Meter	30 Meter	45 Meter	80 Meter	Harmonic Mean
Tupe	96.3[a]	98.9[a]	100.0[a]	100.0[a]	182.49
Mveg	94.7[a]	97.6[a]	99.2[a]	100.0[a]	150.64
Crop	94.8[a]	97.1[a]	98.1[a]	97.3[a]	771.28
Past	93.2[a]	95.6[a]	96.6[a]	97.4[a]	503.43
Soil	94.9[a]	95.7[a]	96.7[a]	96.6[a]	1019.80
Pihd	83.7[a]	89.8[b]	91.6[b]	95.1[b]	163.78
Hdwd	82.5[a]	88.5[b]	91.2[c]	93.3[d]	2092.56
Ccut	79.3[a]	87.0[b]	89.7[c]	92.4[d]	2297.24
Sghd	72.9[a]	85.1[b]	91.3[c]	96.3[d]	1183.66
Pine	72.1[a]	81.1[b]	82.9[b]	95.5[c]	420.12
Watr	79.1[ab]	74.8[a]	79.3[ab]	82.9[b]	232.17

†Dissimilar superscripts within each particular cover class denotes a significant difference at the $\alpha = 0.10$ level of confidence based on the Newman-Keuls' range test conducted on the arcsin transformed proportions. The proportions are the relative rates of omission in classification.

The harmonic mean is computed by:

$$HM = m / \sum_{r=1}^{m} \frac{1}{n_r}$$

where:

HM = harmonic mean

m = the number of elements included in the mean.

n_r = the number of pixels sampled in computing the proportion correctly classified using the r(th) spatial resolution.

In Table 4, as indicated, dissimilar superscripts within each particular cover class denote a significant difference between the various spatial resolutions at the 10% confidence level. Thus, one sees that hardwood, clearcut, second growth hardwood, and in some cases pine, pine-hardwood, and water classes all show statistically significant differences in classification performance between the different spatial resolutions. Agricultural cover types, including crops, pasture, and soil, as well as the mixed vegetation and tupelo forest cover, did not show significant differences between the various spatial resolutions. Thus it appears that it is primarily the forest cover types (with the exception of tupelo) in which the impact of different spatial resolutions causes significant differences in classification performance. These results are perhaps more easily seen in Figure 1, which shows the response surface for each of the individual cover classes for each of the four resolutions tested. As clearly shown on this response surface, for most of the forest cover types, classification performance tends to increase rather dramatically with a

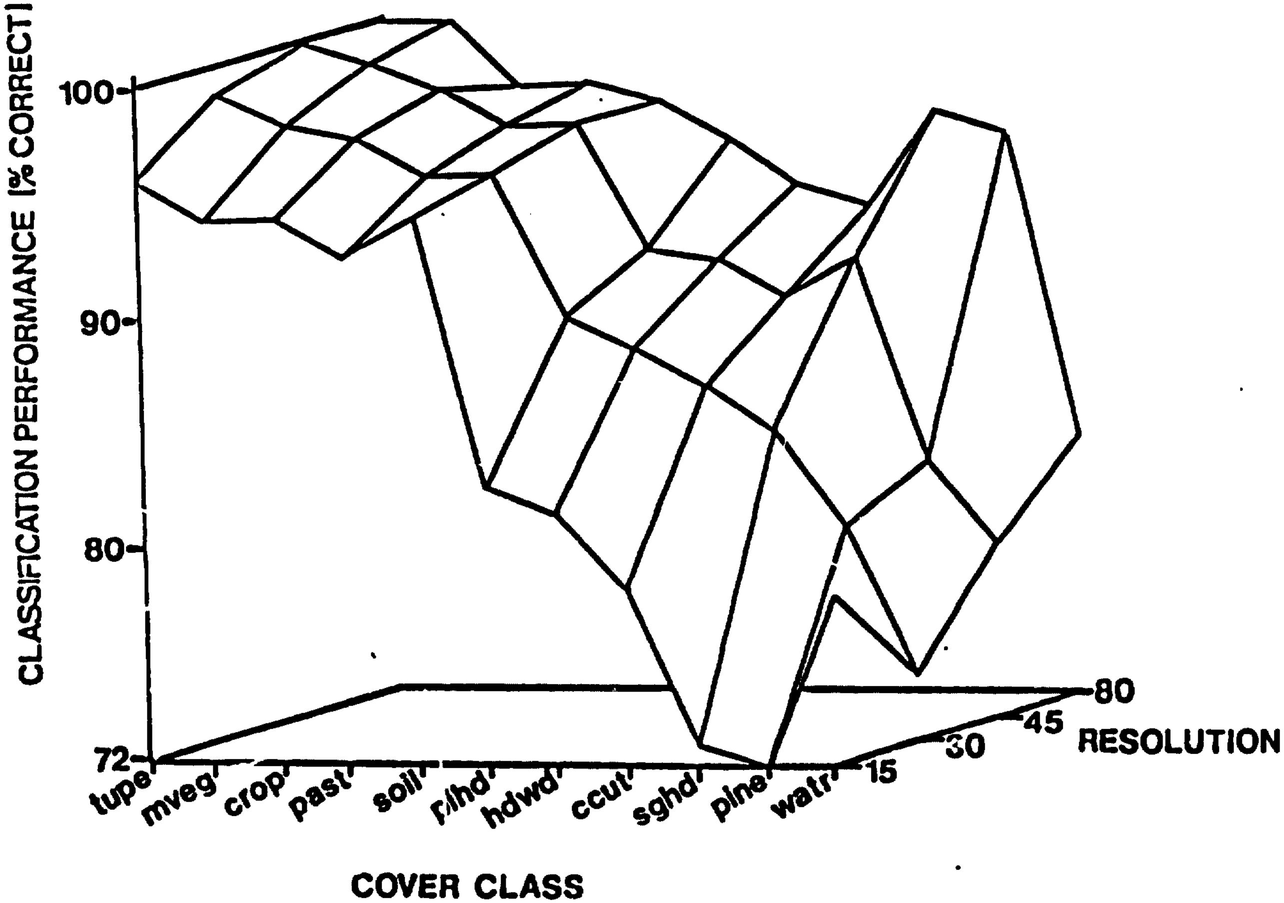


Figure 1. Response Surface in Percent Correct Classification by Cover Class, for each Spatial Resolution (Training Field pixels, Per-point GML classifier, CAMIS).

decreased or larger spatial resolution. On the other hand, tupelo, mixed vegetation, crop, pasture and soil have very high classification performances at all four spatial resolutions. (In considering the high classification performances shown here, one must keep in mind that these results are for the training data only.) These results indicate that although agricultural cover types may not be significantly impacted by the higher spatial resolution of Thematic Mapper data, the classification performance that can be achieved for forest cover types can be significantly affected by the higher spatial resolution of Thematic Mapper type data. Figure 2 indicates that on the basis of overall classification results, there is a distinct increase in classification performance with larger spatial resolution. Thus it would appear that the spatial resolution of the Thematic Mapper scanner system may have a very significant, and possibly detrimental, impact on classification performance achieved when analyzing data obtained over primarily forested areas.

Further evaluation of the characteristics of the data for the different spatial resolutions indicated that the spectral variability from among adjacent pixels was much higher with the higher spatial resolution data sets. Such variation in spectral response level is clearly shown in Figure 3, which depicts the across-track variation is spectral response for each of the spatial resolution data sets. These graphs provide some insights as to why the classification performance at the 15 meter spatial resolution was sometimes much poorer than at the Landsat spatial resolution. It is thought that at the 15 meter spatial resolution, pixels for a given cover type tend to have so much spectral variability that many pixels apparently are spectrally similar to a completely different cover type. However, at the Landsat spatial resolution, the texture in the data

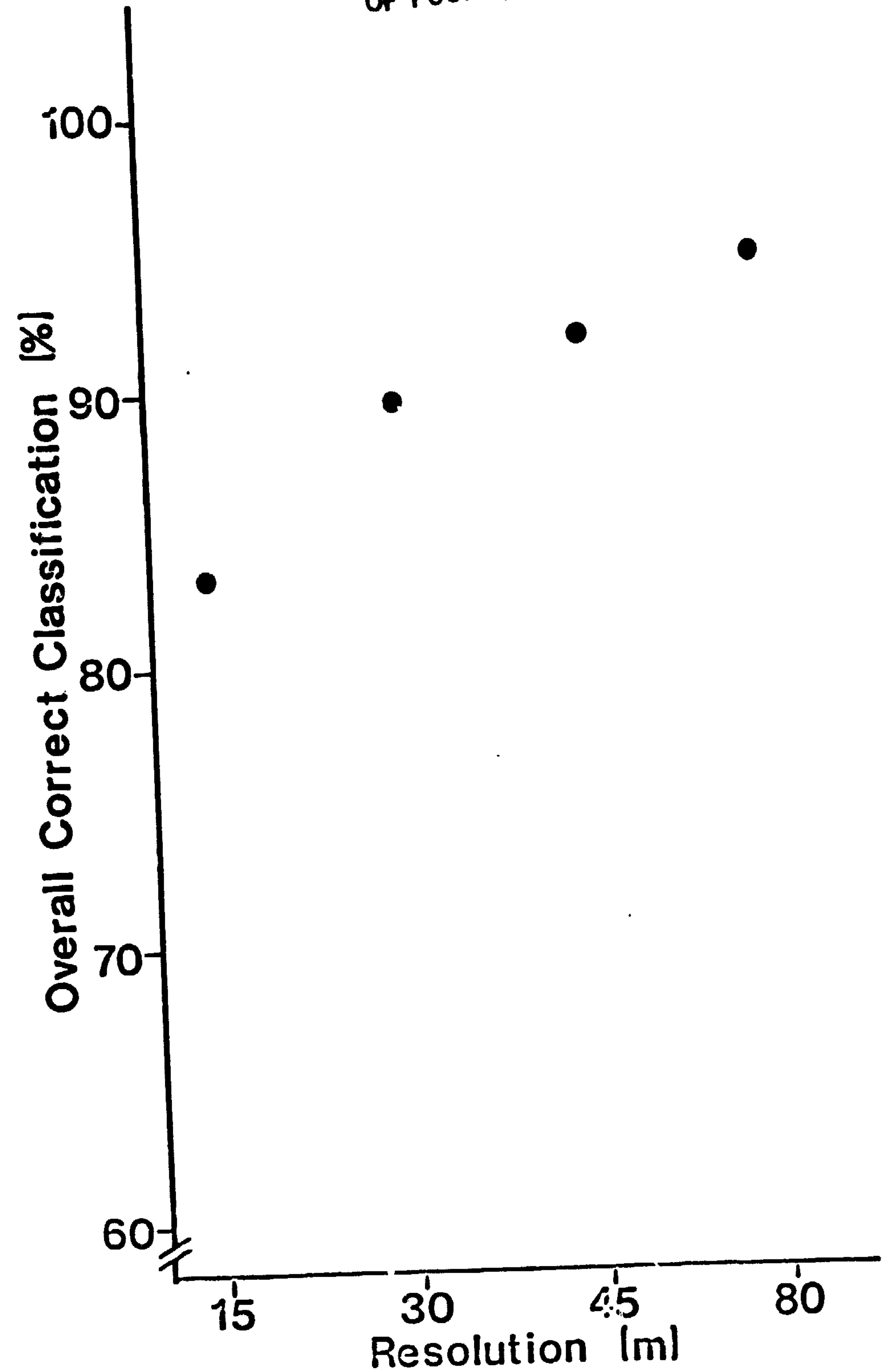
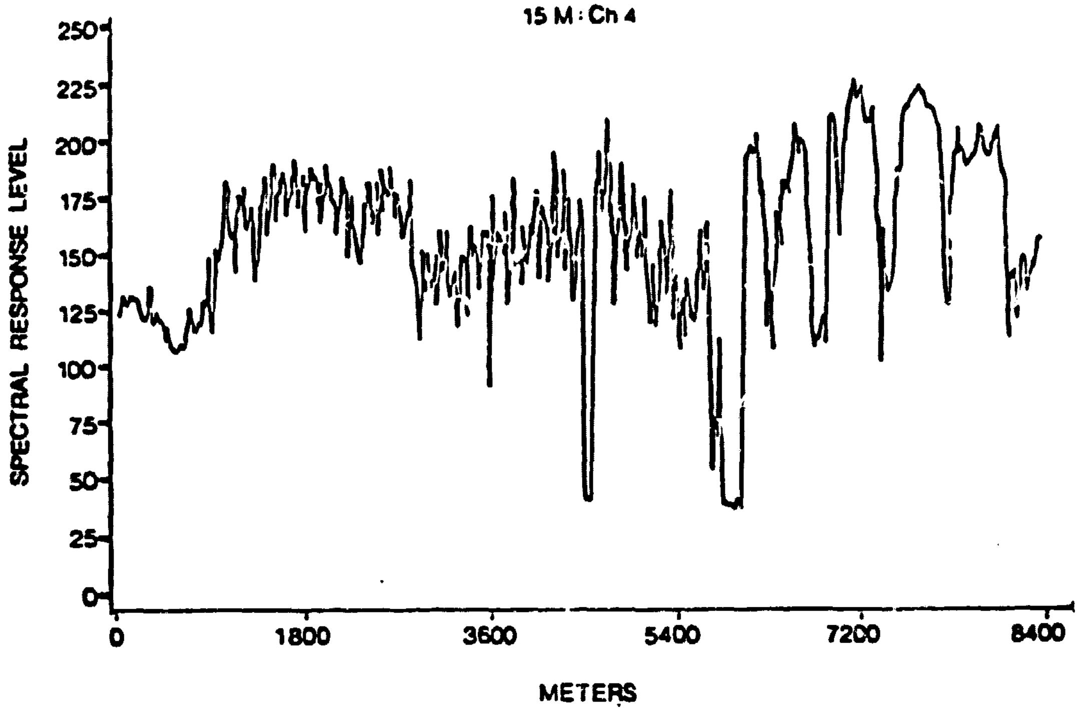


Figure 2. Overall Classification Results for each Spatial Resolution (Training Field pixels, Per-point GML classifier, CAMIS).



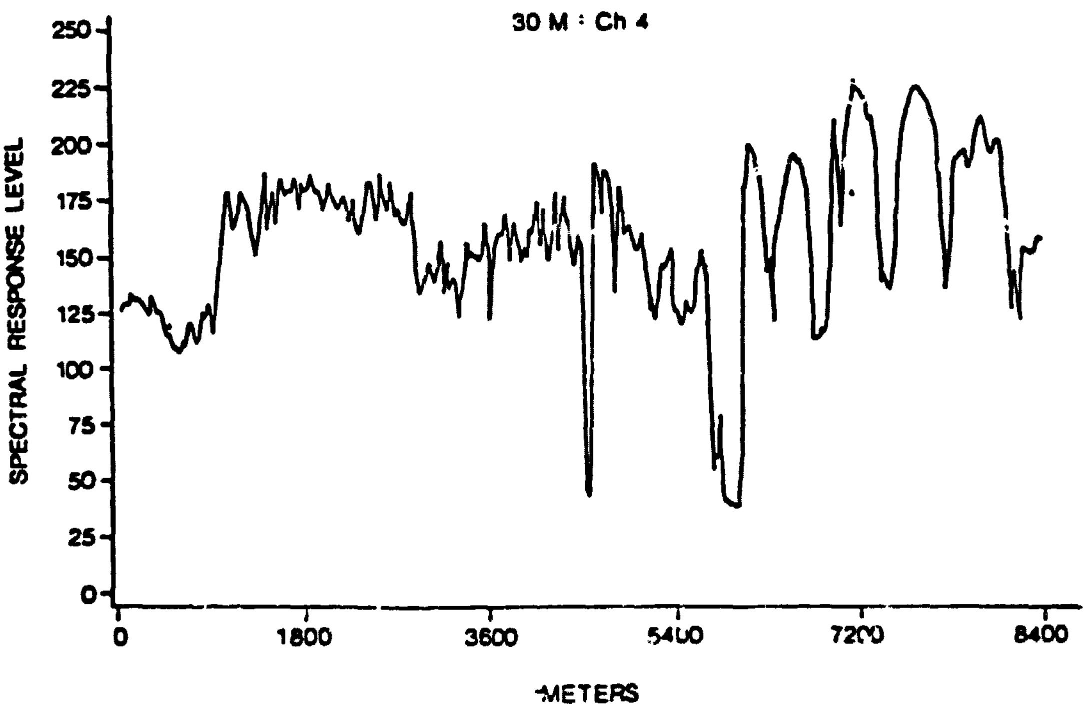
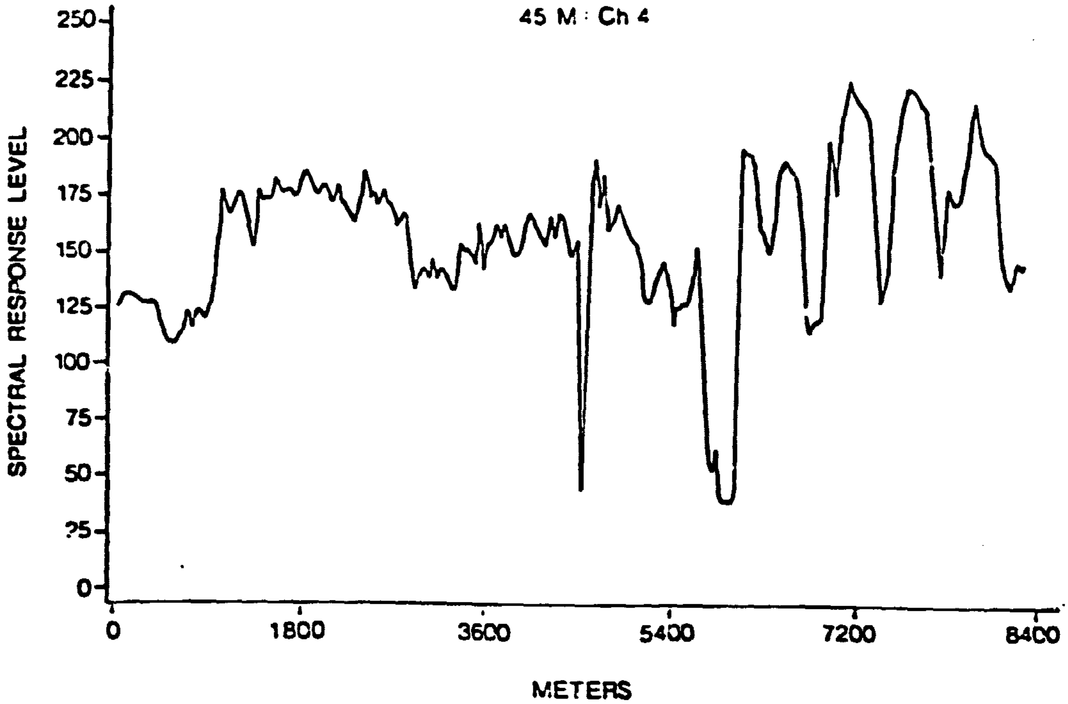


Figure 3. Variation in Spectral Response Level with respect to distance in the across-track dimension for 15, 30, 45, and 60 meter sampling intervals (Flight Line CAM1S).



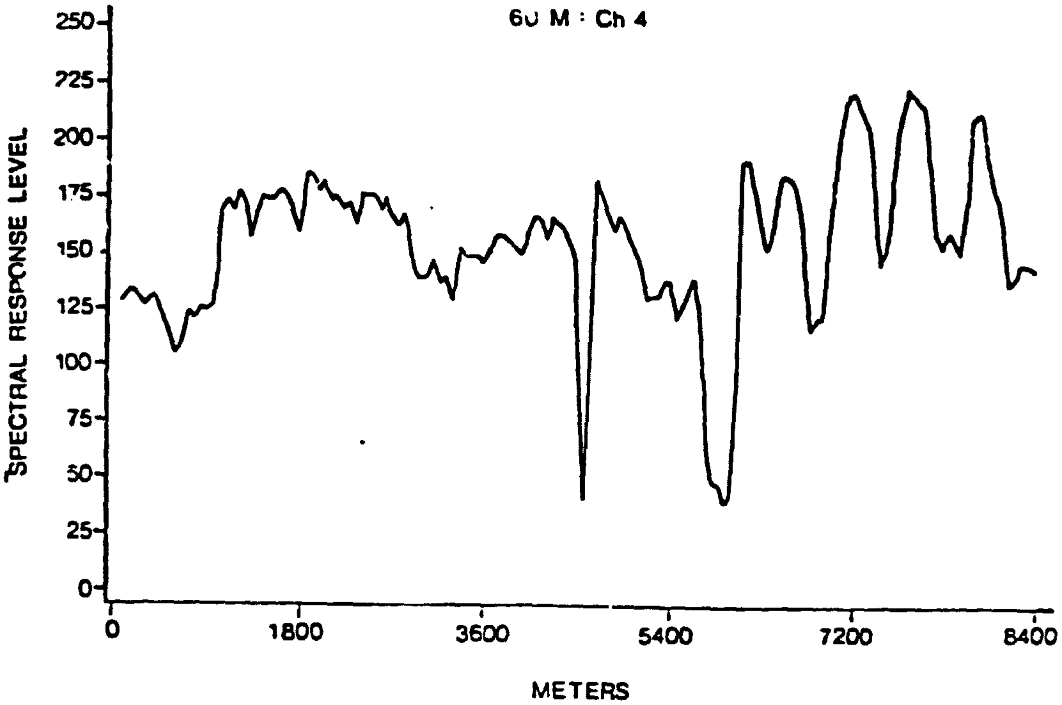


Figure 3. (Continued.)

tends to be averaged out within a particular pixel and the reflectance for that pixel is a representation of the overall spectral response from each individual component within the pixel area. This overall or averaged spectral response is often sufficiently different for different cover types that pattern recognition algorithms can be used to effectively differentiate between the cover types involved. For example, the spectral response of Landsat resolution pixels of hardwood is sufficiently different from pine to allow effective differentiation, whereas at the 15 m spatial resolution, some pixels within the hardwood area may actually fall partially on a shadow area between two tree crowns, possibly resulting in a spectral response similar to that of illuminated pine crowns. In such a case, this pixel within the hardwood forest area probably would be misclassified as pine. Thus, due to the greater spectral variability found among the individual pixels in the higher resolution data, many pixels are misclassified, particularly in the areas of forest cover (where spectral variability is higher than in the agricultural cover types).

In summary, although Thematic Mapper data will undoubtedly be better than the current Landsat data from a mensurational standpoint, these preliminary results showing a decreased classification performance with higher (eg., smaller) spatial resolution data tend to indicate that conventional per-point classification techniques may not be as effective when using higher resolution data. Thus, classification techniques such as "ECHO" (which utilizes the spatial variability in addition to the mean spectral response of an entire forest stand or agricultural field), need to be further tested and refined for potential use with Thematic Mapper data.

Since these results are based upon training data, they are considered to be preliminary, and the trends indicated must be further tested and evaluated using the test data from both CAM1S and CAM2N (the north end of Flight Line 2, near Wateree Reservoir). Evaluation of the test data set for CAM1S and the training and test data set for CAM2N will be pursued during the coming quarter.

A total of 271 training fields have been selected in CAM2N in preparation for the analysis of the data from this area. Because of the differences in land cover characteristics of this area, many of the training fields tend to be much smaller. This is particularly true for the hardwood cover types which tend to follow the drainage system present in the area. Definition of the test data pixels has been completed for CAM1S and is nearly complete for CAM2N.

B. Definition of Radar Data Digitization Procedures

As indicated in previous reports, in order to get the radar data into quantitative format, the imagery obtained must be digitized using a microdensitometer. Thanks to the efforts of our contract monitor, Mr. Norman Hatcher, the radar data for the Flight Line 1 area, obtained on June 30, 1980, has been provided to Lockheed Corporation at the Johnson Space Center for digitization. Both the HH and HV polarization images are to be digitized.

The parameters for digitizing the imagery were calculated using the specifications of the radar system (shown in Appendix A), and an approximate scale of the imagery of 1:376,000. The scale was determined by making several measurements between points on the contact radar image and USGS topographic maps. According to the information obtained from NASA concerning the characteristics of the system, the ground resolution for both the across track and along track resolutions is slightly less than 15 m. This figure was then used as the minimum allowable dimension for a ground resolution element. From the above scale and ground resolution values, it was determined that an aperture setting of 40 μm should be used. The scanning interval and scan line spacing will be equal to the aperture setting. This will prevent any overlap and sidelap of the adjacent pixels, thus providing independence between them. If there is any overlap and/or sidelap of the pixels, the variance between the adjacent pixels would be reduced. This would not allow as effective a comparison among various classification algorithms since the design of some algorithms are more sensitive to a change in variance than others. Table 5 summarizes the parameters for the digitization of the imagery. It is anticipated that the digitization will be completed soon and we can initiate the quantitative analysis of the data. A qualitative interpretation of the imagery is currently being pursued.

Table 5. Flight Characteristics and Parameters for the Digitization of Radar Imagery.

Radar Flight Characteristics

Date of Flight: June 30, 1980

Location: Camden, South Carolina

Flight Line: 4 (Corresponds to Flight Line 1 of the photographic and MSS data)

Mode II used. Near-range look angle = 14.04°
Far-range = 51.34°

Parameters for the Digitization of Radar Imagery

Number of Gray Levels: 256 (0-255)

Aperture setting: 40 μm

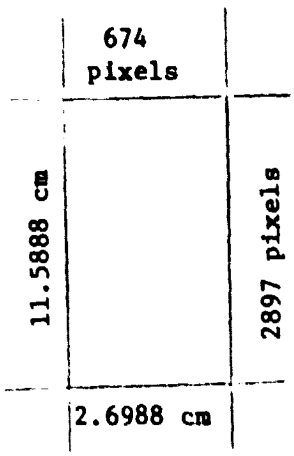
Scanning interval: 40 μm

Scan Line spacing: 40 μm

Number of pixels per cm: 250

Area of pixel on ground: 177.93 m^2

Dimension of test area on image:



Area of test site: 31.2759 cm^2

Total number of pixels: 1,597,448

Digitization should be carried out in West to East scan lines, starting at the North end of the Flight Line.

C. Revision of Phase III Statement-of-Work

Due to some programmatic changes at JSC, it was learned during the past quarter that the funding level for Phase III of this investigation would be reduced to $50K. Because of this situation, a review of the statement-of-work for Phase III was conducted and appropriate modifications were agreed to by the Principal Investigator and Mr. Hatcher. The Modified Statement-of-Work and Revised Budget is in the process of being approved by NASA and Purdue Contract personnel.

D. Participation of Semi-Annual Convention of the American Society of Photogrammetry

A paper had been submitted to the American Society of Photogrammetry for possible presentation at the Semi-Annual Convention of ASP, held in Niagara Falls, New York in October, 1980. A copy of the paper to be presented was included in the last Quarterly Progress Report. The paper was accepted for presentation and was published in the proceedings of the 1980 Semi-Annual Convention of ASP. Mr. Rick Latty made the trip to Niagara Falls and presented the paper, which was very well received.

II. PROBLEMS ENCOUNTERED

No problems of significance were encountered during the past quarter.

III. PERSONNEL STATUS

The following personnel committed the respective percentages of time to the project during the past quarter:

Name	Position	Ave. Monthly Effort (%)
Bartolucci, L.	Prof. Research Analyst	10
Dean, Ellen	Research Associate	50
Frazee, Michael	Research Assistant	50
Hoffer, Roger	Principal Investigator	28
Knowlton, D.	Research Associate	50
Latty, Rick	Research Associate	75
Prather, Brenda	Secretary	53
Stiles, Stephanie	Secretary	05

IV. ANTICIPATED ACCOMPLISHMENTS

The following are the anticipated accomplishments of the forthcoming quarter (December 1, 1980 - February 28, 1981):

1) Completion of the digitization of the SAF data for Flight Line 1, HH and HV polarizations.
2) Reformatting and rectification of the 1980 TMS data.
3) Completion of the analysis of the four different spatial resolutions of the 1979 data.
4) Continuation of the analysis of the spectral characteristics of the 1979 TMS data.

Appendix A -- X-Band Side-Looking Radar Specifications

Transmit Frequency	9600 MHz
Transmit Output Power	50 Kw
Transmit Polarization	Selectable Horizontal or Vertical
Pulse Width (Half Power)	0.90 sec
PRF (Pulse Repetition Frequency)	Variable with Ground Speed (at 400 Knots, PRF is 1352 PPS)
Pulse Spectrum Bandwidth	15 MHz

Antenna Stabilization Limits

- Pitch - Up 4.5°, Down 2.5°
- Azimuth - $\pm 6.75^{\circ}$
- Roll - $\pm 3^{\circ}$

Range

Swath Coverage at 60,000 ft.

- Mode I - 2.5 to 12.5 miles
- Mode II - 10 to 20 miles

Range Resolution (across-track)	< 50 ft.
Azimuth Resolution (along-track)	< 50 ft.
Azimuth Beamwidth	1.45° One Way Half Power
Receivers	One Vertical, One Horizontal (Hard Wired to Recorder)
Recording Mode	Optical and Selected Digital