National Aeronautics and Space Administration

EARTH RESOURCES LABORATORY

AN INITIAL ANALYSIS OF LANDSAT 4 THEMATIC MAPPER DATA FOR THE CLASSIFICATION OF AGRICULTURAL, FORESTED WETLAND, AND URBAN LAND COVERS

REPORT 215

NOVEMBER 1982

NATIONAL SPACE TECHNOLOGY LABORATORIES

AN INITIAL ANALYSIS OF LANDSAT 4 THEMATIC MAPPER DATA FOR THE CLASSIFICATION OF AGRICULTURAL, FORESTED WETLAND, AND URBAN LAND COVERS

NOVEMBER 1982

BY

DALE A. QUATTROCHI
JAMES E. ANDERSON
DAVID P. BRANNON
CHARLES L. HILL

EARTH RESOURCES LABORATORY
NATIONAL SPACE TECHNOLOGY LABORATORIES
NSTL STATION, MS 39529

Original photography may be purchased
from EROS Data Center
Sioux Falls, SD 57198

REPORT NO. 215

ABSTRACT

An initial analysis of Landsat 4 Thematic Mapper (TM) data for the delineation and classification of agricultural, forested wetland, and urban land covers has been conducted using a scene of data collected on August 22, 1982. A study area in Poinsett County, Arkansas has been used to evaluate a classification of agricultural lands derived from multitemporal Landsat Multispectral Scanner (MSS) data in comparison with a classification of TM data for the same area. Data over Reelfoot Lake in northwestern Tennessee have been utilized to evaluate the TM for delineating forested wetland species. A classification of the study area has been assessed for accuracy in discriminating five forested wetland categories. Finally, the TM data have been used to identify urban features within a small city. A computer-generated classification of Union City, Tennessee has been analyzed for accuracy in delineating urban land covers. An evaluation of digitally enhanced TM data using principal components analysis to facilitate photointerpretation of urban features has also been performed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF MAPS

LIST OF TABLES

PREFACE

This report describes the results of initial work by several investigators at NASA/NSTL/ERL on evaluating Landsat 4 Thematic Mapper (TM) data for the delineation and classification of agricultural, forested wetland, and urban areas. Although there are three component sections to the report, one for each of the specific land cover analyses, a common bond links these disparate investigations. All TM data have been extracted from scene ID#40037-16031, which is one of the first seven channel data sets made available to TM investigators after the launch of Landsat 4 in late July 1982. The TM scene used here was collected on August 22, 1982 and encompasses portions of four states -- Arkansas, Missouri, Kentucky, and Tennessee.

To retain the cohesive identity of the individual investigations, each is presented as a separate unit within the report. The first section compares a classification previously derived from Landsat Multispectral Scanner (MSS) data for agricultural land covers in Poinsett County, Arkansas, with a TM classification of the same area. An assessment is then made of the amount of improvement afforded by the TM over what had been derived from MSS data.

The second segment of the report describes the results that have been obtained from an evaluation of TM data for delineating forested wetland species. Data over Reelfoot Lake in northwestern Tennessee have been examined for accuracy in separating five types of forested wetland categories.

In the last section, the ability of the TM to discriminate urban features is addressed using Union City, Tennessee, as a study area. A computer implemented classification is evaluated for accuracy in delineating land covers within the city. An analysis of TM data for photointerpretation of urban features is also discussed in this section of the report.

SECTION I: EVALUATION OF MSS AND TM DATA FOR CLASSIFICATION OF AN AGRICULTURAL STUDY AREA IN NORTHEASTERN ARKANSAS

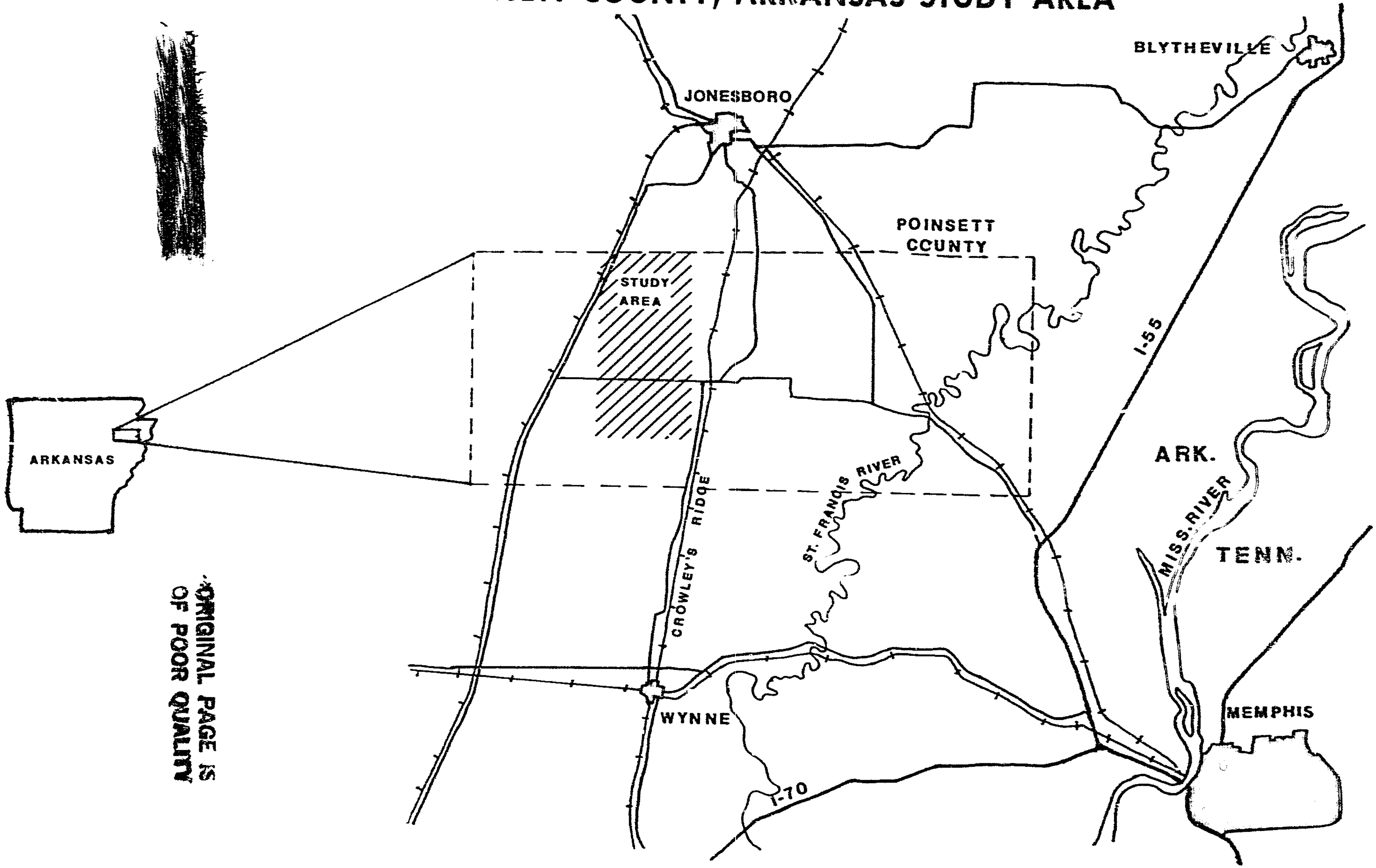
The investigation had as its purpose: (1) to examine the potential of Landsat 4 TM data for providing agricultural information; and (2) to compare the general capabilities of TM and Landsat 2 and 3 MSS data for crop mensuration and mapping. Poinsett County, Arkansas, located in the northeast part of the state, has been used as a study area for agricultural analysis. The following discussion provides a description of the study area, results from the MSS and TM analysis, and an assessment of the comparative performance between the two sensor systems for delineating agricultural land covers within the study area.

DESCRIPTION OF THE STUDY AREA

Poinsett County lies partially within the floodplains of the Mississippi and St. Francis Rivers and the well-drained uplands west of Crowley's Ridge (Map 1). The land area of the county is approximately 486,208 acres (196,845 ha) with 84% of this acreage maintained in farmland. Elevations above mean sea level in the county range from about 400 feet on Crowley's Ridge to about 140 feet near the St. Francis River at the southern boundary of the county. County topography can be divided into three distinct sections: the level bottomlands in the east; the gently steep Crowley's Ridge; and the moderately sloping upland plains to the west.

Bottomlands east of Crowley's Ridge comprise approximately 45% of the county's land area. With the exception of a few wooded tracts in the St. Francis River Floodway, the entire area of bottomland soils is cultivated. Most of the uplands are suitable for crop and pasture land, except for the

MAP OF POINSETT COUNTY, ARKANSAS STUDY AREA



Map 1

heavily dissected lands along Crowley's Ridge. Upland soils, comprising approximately 55% of the county, extend westward from Crowley's Ridge.

The primary crops produced in Poinsett County include cotton, rice, soybeans, wheat, and to a lesser degree, grain sorghum, and hay. During the past few years improvements in crop varieties, flood control measures, and drainageways have led to a rapid expansion of croplands into the wet bottomlands, and a subsequent reduction in woodland acreage.

MSS DATA PROCESSING PROCEDURE AND CLASSIFICATION ANALYSIS

Landsat MSS data for the Poinsett County area were selected over Path 25, Row 35 of the Landsat Worldwide Reference System corresponding to critical phases of the growing season. Data sets were obtained for (1) pre-planting conditions and winter wheat mapping (February 26, 1981, ID#2222715584); (2) midseason vigor (July 20, 1981, ID#2237115563); and (3) senescence (September 30, 1981, ID#2244315554). The dates were used to give the broadest possible range of spectral values for the target land cover types.

Multi-date analysis requires the registration of the Landsat data sets before processing begins. This procedure entails selecting a number of points held in common by the data sets to be registered. Registration of Landsat scenes using this method is accurate to within one pixel (57 x 79 meters). Generally, the highest quality, cloud-free data set is used as base for overlaying the remaining data sets. For the Poinsett area, the February scene has been used as the base; channels 2 and 4 from each data set have then been registered to create a single six-channel data set used for analysis.

Data processing for the study utilized the Earth Resources Laboratory Applications Software (ELAS) (Graham, et al., 1981). ELAS is a Fortran-based operating subsystem designed to analyze digital imagery data. The ELAS module Search (SRCH) which employs an unsupervised approach to develop spectral signatures was used to derive reflectance statistics from the MSS data. SRCH moves a 3x3 pixel "window" through the data and develops signatures for those windows of data which meet specified criteria for homogeneity. Once spectral statistics have been derived, they are classified using the gaussian maximum likelihood technique.

For the 75,053 acres (30,386 ha) within the study area, which is comprised of portions of the Powers Slough and Otwell 7.5 minute U.S.G.S. topographic quadrangle maps, SRCH developed 30 spectral signatures. These spectral responses were classified, and the resulting classes were identified by specific land cover types. Class identification or naming was accomplished by selecting field sites, visiting these areas to record and photograph the characteristic land covers, and then the field data along with ancillary information were used to name the classes. The 30 spectral classes developed for the study were related to specific land cover types using ground truth, aerial photography, or crop production records. During this investigation, the naming and verification of the spectral classification was made with a series of ground truth polygons delineated from aerial photography and USGS 7.5 minute quadrangle maps.

Polygons used in the ground truth or class naming and verification exercises were randomly selected from a 3,200 acre (1,280 ha) area within the study area; these 3,200 acres (1,280 ha) encompassed 20 quarter-sections of land. Half of the polygons were utilized for class naming, and half were used for accuracy assessments. Aerial photography employed as ancillary information helped to define field boundaries within each quarter section. Only interior field pixels were digitized and used for data analysis, thereby reducing spectral confusion within the polygons. Class naming polygons totaled 52 with verification polygons totaling 56. Crop statistics compiled by the Agricultural Stabilization and Conservation Service were used to name the agricultural polygons for the 1981 growing season. Aerial photography and field observations were utilized to identify the non-agriculture polygons.

MAPPING RESULTS AND ACCURACY

The 52 class naming polygons were identified and used to group the 30 MSS spectral classes into six land cover categories:

1. Hardwood
2. Fallow
3. Rice
4. Soybeans
5. Winter Wheat/Soybeans
6. Water

Prior to computing areal acreage assessments for each class, the data were geographically referenced to a Universal Transverse Mercator (UTM) grid with a precision of 50 meters root mean square (rms) error. Geographic referencing is accomplished by identifying control points on appropriate U.S.G.S. 7.5 minute quadrangle maps for the study area. The appropriate row and element number and longitude/latitude or UTM coordinates are then correlated to effect rectification.

The geographically referenced data and corresponding acreages for each class within the study area are illustrated in Figure 1. A U.S.G.S. 7.5 minute map base has been manually overlaid onto the map to facilitate geographic orientation.

The overall accuracy of the land cover map in Figure 1 was determined using the Accuracy of Classification Table (ACTB) module in ELAS. ACTB used the 56 verification polygons, coded by land cover type, to construct an accuracy table. Comparisons were then made between the polygons in the accuracy table and the corresponding pixels in the classification.

Table 1 groups the accuracy assessments according to land cover type; overall accuracy of the land cover classification based on MSS data was 80.91% correct. The highest percentages of classification accuracy were recorded for the hardwood, rice, and soybean categories with percentages of 100, 98.74, and 89.69, respectively. Low accuracies were recorded for double-cropped areas of winter wheat and soybeans (18.63%) and water (22.81%). Seventy-three percent of the double-cropped areas of winter wheat and soybeans were classified as soybeans. The relatively smaller fields of winter wheat in the study area did not contribute a significantly large number of interior field pixels for signature development. Reduced availability of interior field pixels will result in errors of ommission for small land cover features, especially when applying the sliding window approach (Stoner, et. al. 1981). This problem can be alleviated to a certain degree by utilizing a pixel-by-pixel analysis which will gather a sufficient number of small field pixels during signature development. Also. impoundments in this area are normally smaller than 10 acres and are usually bordered by hardwoods. The 3x3 pixel sliding window in SRCH, therefore, did not find a significant number of water pixels for signature development.

ASSIFICATION DERIVED FROM MSS DATA SETS
QUIRED 2/26/81 7/20/81 9/30/81

POWERS SLOUGH/OTWELL QUADRANGLES

AN AUTOMATED CLASSIFICATION OF LANDSAT II
MSS DIGITAL DATA DERIVED FROM SRCH
GENERATED STATISTICS

LANDCOVER CATEGORY		ACREAGE
HARDWOOD	■	5,001
FALLOW	▨	117
RICE	■	28,740
SOYBEANS	▨	31,510
WINTER WHEAT/SOYBEANS	▨	575
WATER	■	717

FIGURE 1
CLASSIFICATION OF POINSETT COUNTY STUDY
AREA DERIVED FORM MULTI-TEMPORAL MSS DATA

TABLE 1: ACCURACY OF CLASSIFICATION FOR MSS DATA - POINSETT COUNTY, ARKANSAS STUDY AREA

LAND COVER CATEGORIES

MAPPING ACCURACY

GROUND VERIFICATION CLASSES	HARDWOOD 1	RICE 2	SOYBEANS 3	WINTER WHEAT/ SOYBEANS 4	WATER 5	NO. OF PIXELS
1	307	0	0	0	0	307
	100.00	0.00	0.00	0.00	0.00	
2	0	2191	28	0	0	2219
	0.00	98.74	1.26	0.00	0.00	
3	0	149	1296	0	0	1445
	0.00	10.31	89.69	0.00	0.00	
4	0	41	339	87	0	467
	0.00	8.78	72.59	18.63	0.00	
5	44	0	0	0	13	57
	77.19	0.00	0.00	0.00	22.81	

PERCENT CORRECT OVERALL 80.91

CLASS	%CORRECT
1 Hardwood	100.00
2 Rice	98.74
3 Soybeans	89.69
4 Winter Wheat/Soybeans	18.63
5 Water	22.81

Although a fallow category has been defined in Figure 1 no ground verification sites were selected for this class. Field conditions within the study area did not permit delineation of both ground truth and ground verification sites for the fallow category. Thus, the class was identified, but no accuracy statistics were produced for inclusion in Table 1.

TM DATA PROCESSING PROCEDURE AND CLASSIFICATION ANALYSIS

TM data from the August 22, 1982 data set used in the investigation were subjected to the SRCH automated signature development algorithm for derivation of spectral signatures. It is significant to note that at least one year had elapsed between MSS and TM data acquisition for the study area. Figure 2 illustrates the relative frequency of occurrence (i.e., spectral response range) for the TM data utilized in the Poinsett County study area. Channels 2, 4, and 5 were used to develop signatures from the data; the other four channels were not input to SRCH. Analysis of spectral scattergrams for selected ground verification plots using all seven channels of data indicated that separation of agricultural cover types with channels 2, 4, and 5 produced statistically better results overall than those produced using all seven channels of information.

The unsupervised statistical development technique resulted in 43 spectral signatures. All ground truth polygons evaluated during the MSS analysis were revisited in August 1982 and were used to: (1) establish the relationship of spectral signatures with specific land covers; and (2) develop estimates of accuracy. Because of crop rotation practices, land uses in the study area had changed in the intervening time period between MSS and TM data acquisition. The distribution of land cover types for TM analysis, therefore, was not identical to that found in the MSS evaluation.

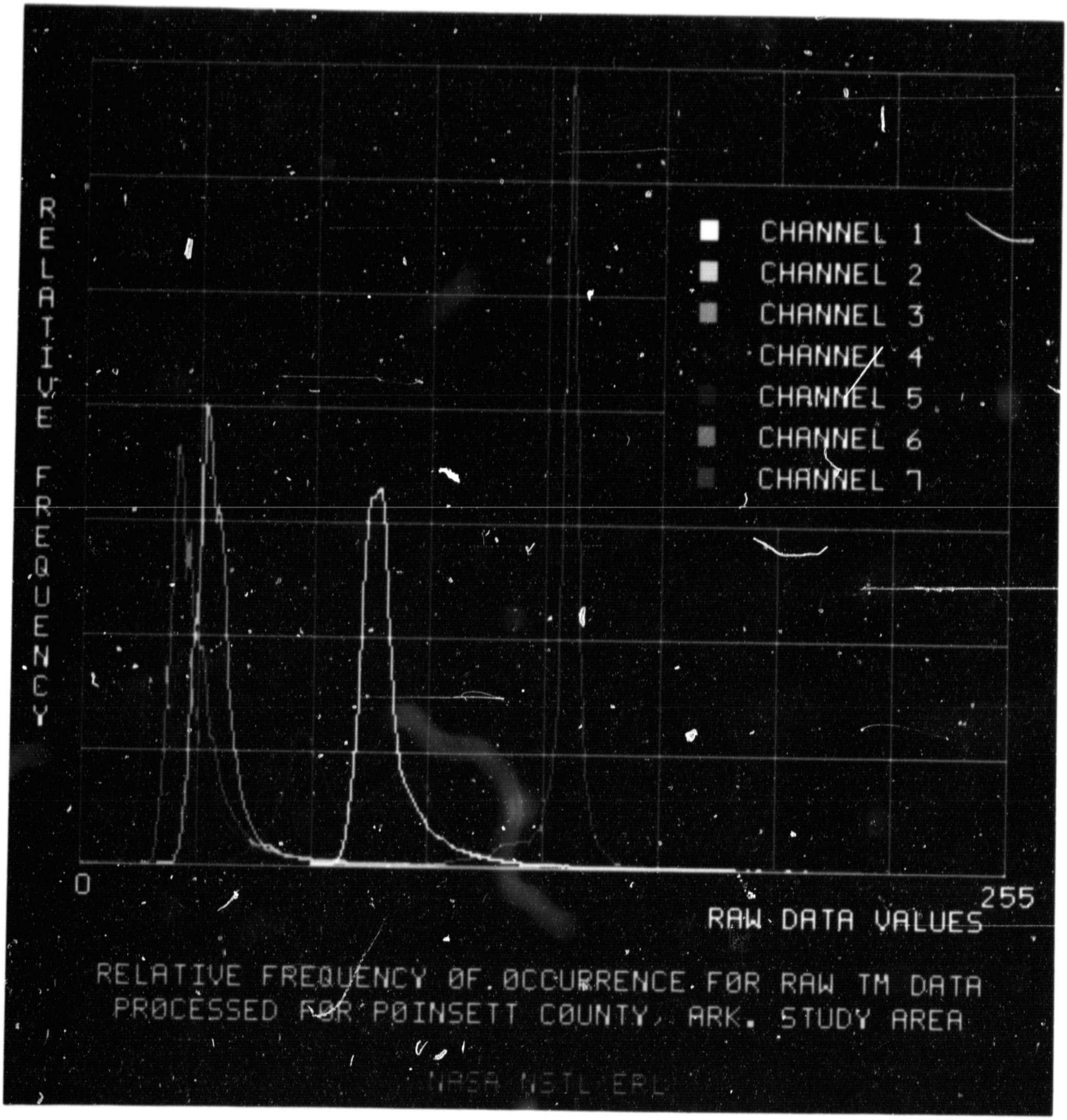

RELATIVE FREQUENCY
CHANNEL 1
CHANNEL 2
CHANNEL 3
CHANNEL 4
CHANNEL 5
CHANNEL 6
CHANNEL 7
0
255
RAW DATA VALUES
RELATIVE FREQUENCY OF OCCURRENCE FOR RAW TM DATA
PROCESSED FOR POINSETT COUNTY, ARK. STUDY AREA
NASA NSTL ERL

Figure 2

Consequently, the 43 spectral signatures developed from the data were identified as belonging to one of five land cover categories (Figure 3):

(1) Soybeans

(2) Rice

(3) Fallow/bare soil

(4) Hardwood

(5) Water

As in the MSS analysis, ACTB was used to establish accuracy relationships between the TM classification and the ground verification polygons selected for the study area. The ACTB accuracy results are presented in Table 2.

COMPARISON OF MSS AND TM CLASSIFICATION RESULTS

The accuracies listed in Tables 1 and 2 have been used to compare the performance of the MSS versus the TM within the Poinsett County study area. Results from a Newman-Keuls test of comparative accuracies, based on the arcsine $\sqrt{p}$ transformation at the .05 significance level, illustrate that: (1) the TM did significantly better than the MSS in an accuracy assessment of the water class; (2) the MSS and TM performed equally well (i.e., no statistically significant difference in accuracies) in the soybeans, rice, and hardwood categories; and (3) most importantly, the TM performed significantly better than the MSS in an overall comparison of accuracies.[1] These results are made even more interesting when it is considered that only one date of TM data has been used for analysis, as opposed to the three dates of data (February, July, and September) utilized in the MSS evaluation.

---

[1] Unlike in the MSS evaluation, an accuracy assessment for a fallow/bare soil category was derived for the TM data because land uses within the study area allowed delineation of both ground identification and verification sites.

ORIGINAL PAGE IS
OF POOR QUALITY

Figure 3

TABLE 2: ACCURACY OF CLASSIFICATION FOR TM DATA - POINSETT COUNTY, ARKANSAS STUDY AREA

LAND COVER CATEGORIES

MAPPING ACCURACY

GROUND VERIFICATION CLASSES	HARDWOOD 1	FALLOW 2	RICE 3	SOYBEANS 4	WATER 5	UNCLASSIFIED	NO. PIXELS
1	194	0	0	0	0	0	194
	100.00	0.00	0.00	0.00	0.00	0.00	
2	0	149	0	1	0	11	161
	0.00	92.55	0.00	0.62	0.00	6.83	
3	9	0	4783	18	0	1	4811
	0.19	0.00	99.42	0.37	0.00	.02	
4	45	0	212	4672	0	1	4930
	0.91	0.00	4.30	94.77	0.00	.02	
5	0	0	0	0	83	0	83
	0.00	0.00	0.00	0.00	100.00	0.00	

Percent Correct Overall 97.06

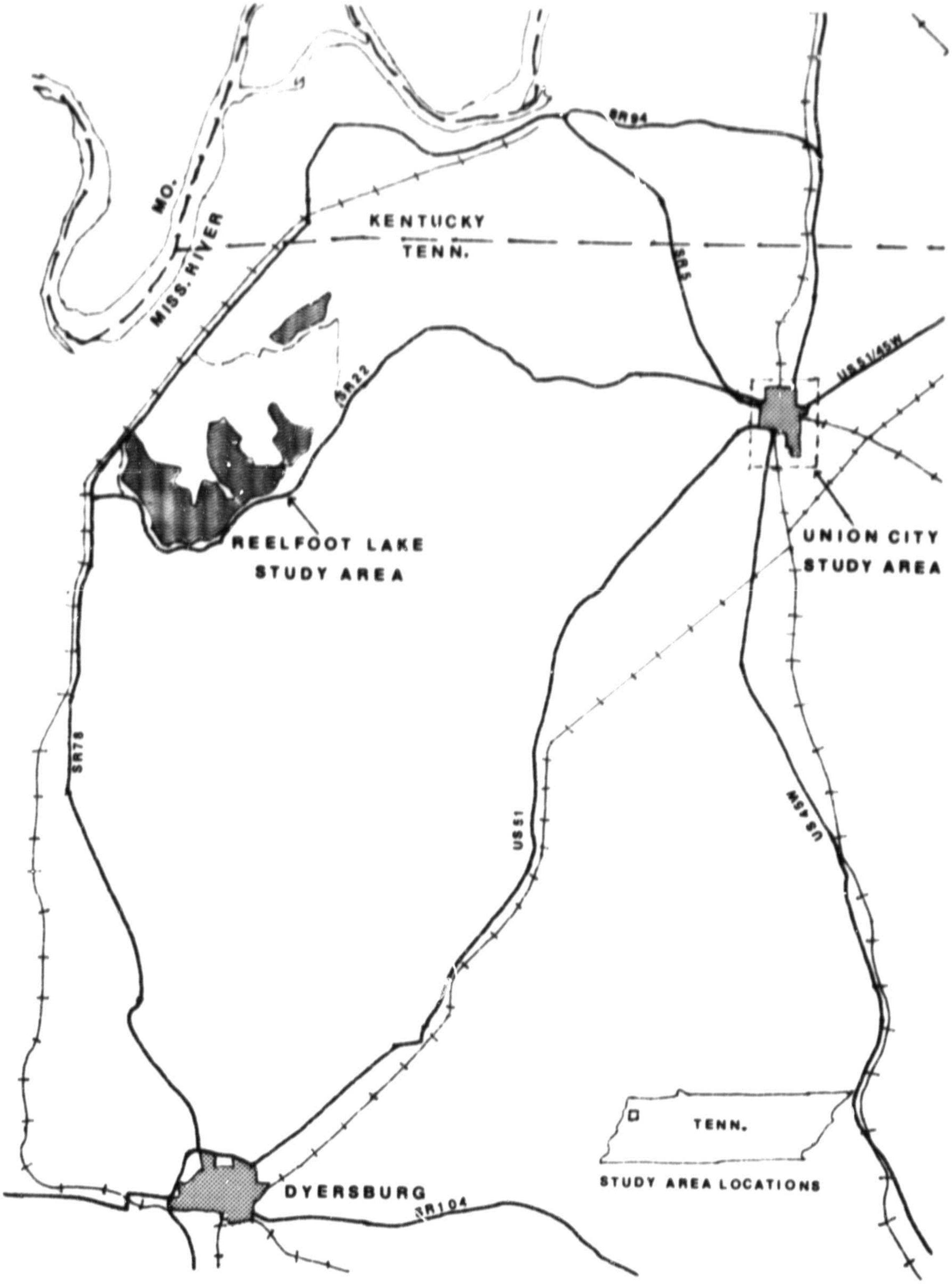
CLASS	%CORRECT
1. Hardwood	100.00
2. Fallow	92.55
3. Rice	99.42
4. Soybeans	94.77
5. Water	100.00

SECTION II: EVALUATION OF TM DATA FOR CLASSIFICATION OF FORESTED WETLANDS IN THE REELFOOT LAKE AREA OF NORTHWESTERN TENNESSEE

Section II discusses the analysis of Landsat 4 TM data collected on August 22, 1982 for the discrimination and inventory of forested wetlands in the vicinity of Reelfoot Lake, Tennessee. Because of its unique aquatic ecology, the Reelfoot Lake area presents an unusual opportunity to test the capabilities of the TM for delineating specific biomes within the interior wetland regime.

DESCRIPTION OF THE STUDY AREA

Reelfoot Lake, located in extreme northwestern Tennessee, is a tectonic feature created by the New Madrid Earthquake of 1811-12 (Carter et al., 1977). The lake is situated within the Mississippi River floodplain of western Tennessee and is bordered on the east by a series of hills known as the Loess Bluffs (Map 2). The Bluffs rise in elevations ranging from 100-200 feet (30.5-61 m) above the Mississippi Alluvial Valley; these hills are heavily dissected by streams which have cut narrow, steep valleys into the unconsolidated loessial deposits (i.e., areas created by aeolian or wind-blown deposition). Reelfoot Lake contains approximately 8,117 acres (3,245 ha) of wetlands which are protected as part of the Reelfoot National Wildlife Refuge, and they are also protected by the state as a fish and game preserve. The lake's wetlands are comprised mainly of cypress mixed with other hardwoods, exposed mudflats vegetated with aquatic grasses and shrubs, and shallow water with emergent and floating aquatic plants. Except for the forested Loess Bluffs to the east, Reelfoot Lake is surrounded by agricultural land with soybeans the predominant crop.



MAP 2
LOCATION MAP OF REELFOOT LAKE, TN
STUDY AREA

TM DATA PROCESSING PROCEDURE AND CLASSIFICATION ANALYSIS

Figure 4 represents a false color composite of the Reelfoot Lake area used in the initial stage of data processing. The image in the figure was created by assigning the blue, green, and red guns of a display device to channels 2, 3, and 4 respectively, of the TM data used in study. The blue colors in the image, therefore, are governed by responses in channel 2 (.52 to .60μm), the greens by channel 3 (.63 to .69μm), and the reds by channel 4 (.76 to .90μm). The collective spectral reflectances from these channels produce a pseudo-false color image which bears a strong likeness to a color infrared aerial photograph. Forested wetlands exhibited in Figure 4 are mostly red in color, while the bright pinks, cyans, and grey-whites represent aquatic grasses and plant communities. Dark blues represent open water which has been excluded from the data analysis procedure.

To focus attention on areas where wetland heterogeneity was most prominent, a polygon delineated by the white line in Figure 4, was drawn around the lake. Only data which fell within this polygon were actually processed for the study. Figure 5 shows the relative frequency of occurrence for the raw TM data used in the study area. Preliminary evaluation of the data for the Reelfoot Lake area, based on the analysis of spectral scattergrams for selected ground verification plots, indicated that TM channels 1 (.45 to .52μm) and 6 (10.40 to 12.50μm) did not contribute a significant amount of information which could be used to spectrally separate wetland classes. Consequently, these channels were eliminated from consideration and data analysis was restricted to channels 2, 3, 4, 5, and 7.

ORIGINAL PAGE
COLOR PHOTOGRAPH

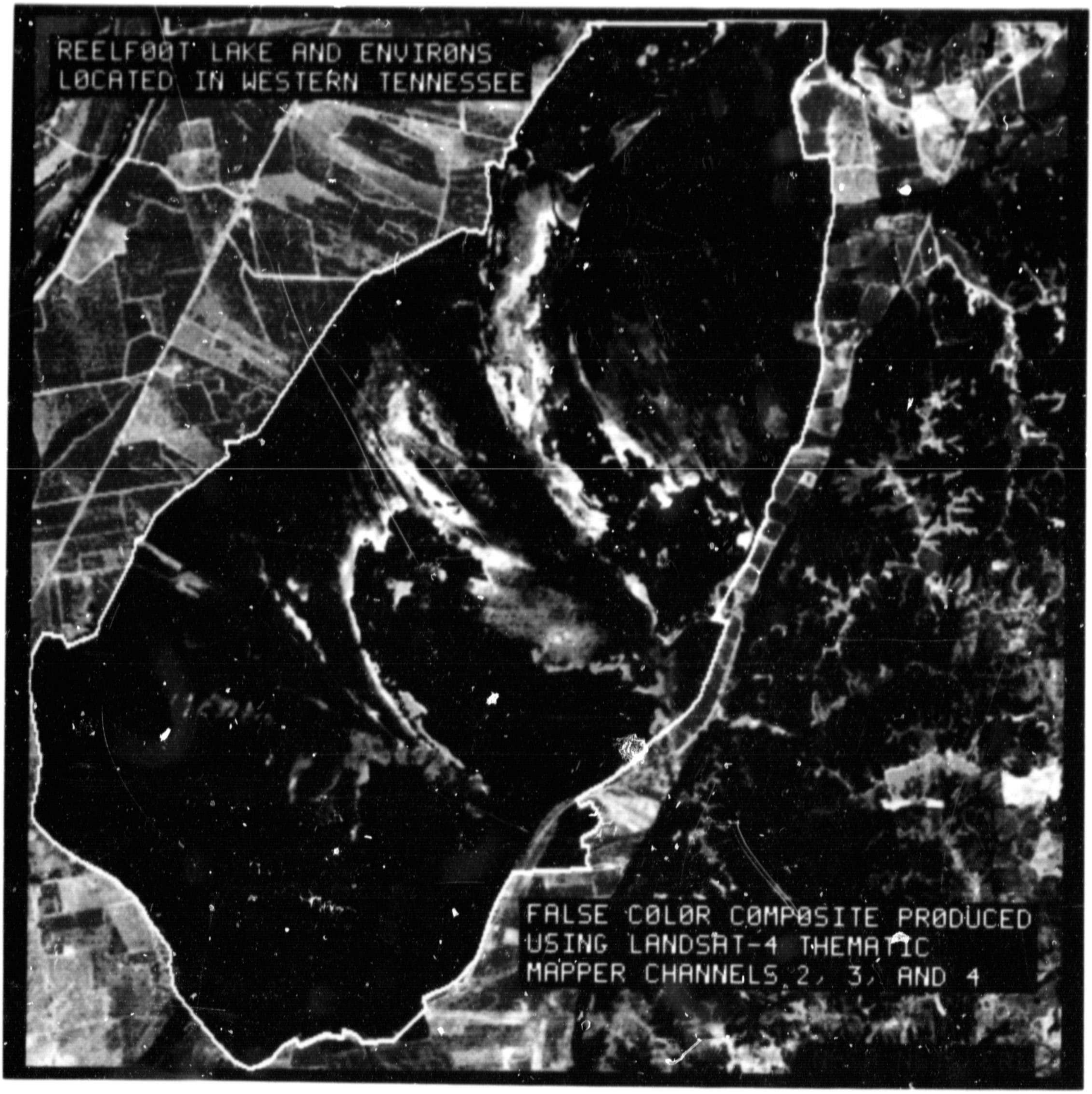
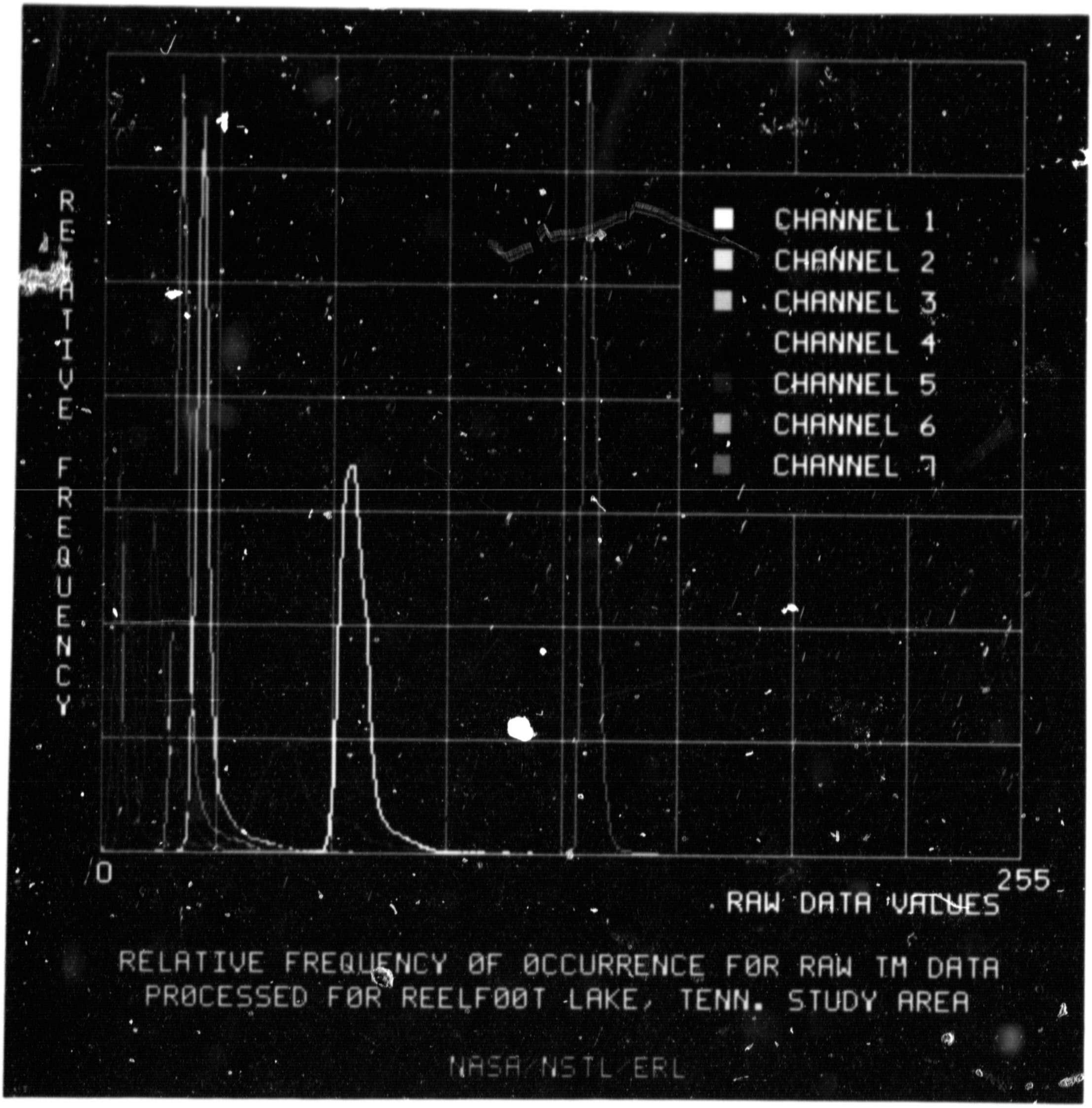


Figure 4



RELATIVE FREQUENCY OF OCCURRENCE FOR RAW TM DATA PROCESSED FOR REELFOOT LAKE, TENN. STUDY AREA

NASA/NSTL/ERL

Figure 5

The SRCH unsupervised spectral signature development approach and the Within Class Cluster (WCCL) pixel-by-pixel unsupervised signature development algorithm in ELAS were used to derive spectral signatures for the polygon-defined area. Sixty spectral statistics were produced by these algorithms, and they were then classified using a gaussian maximum likelihood technique. The 60 classes were ultimately grouped into five primary cover types: cypress; mixed hardwoods; willow and cypress; brush, grasses, and emergents; and floating aquatic vegetation. Ancillary information such as aerial photography taken over the study area and field observation data, was used to group the original 60 classes into appropriate categories. For display purposes, the five primary cover types outlined above were further refined into the nine specific classes illustrated in Figure 6. This nine-category classification was not used in assessing the accuracy of the TM data because ground truth sites could not adequately be defined within some of the classes; i.e., the nine classes in Figure 4 were interpreted from the data, but sites with areal extents large enough for ground truthing could not be defined.

Independent of the class naming procedure, ground verification samples were established for use in determining associated accuracy estimates for each of the five primary cover types. These verification samples were identified using ground observation information, U. S. Fish and Wildlife Service timber stand maps of the Reelfoot Lake area, and recently acquired aerial photographs. The ACTB module in ELAS was used to derive accuracy estimates, based on evaluating the classified TM data (Figure 6) against the known cover types within the ground verification samples. Table 3 summarizes the results generated from the accuracy assessments of the data.

From a close inspection of the accuracies listed in Table 3, several interesting conclusions can be drawn about the performance of the classsification. First because of the high overall percent correct classification,

ORIGINAL PAGE
COLOR PHOTOGRAPH

Figure 6

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE 3: SUMMARY OF "PERCENT CORRECT" ACCURACY VALUES ASSOCIATED WITH THE FIVE LAND COVER TYPES OF INTEREST - REELFOOT LAKE STUDY AREA

			LANDSAT-4 TM CLASSIFICATION LAND COVER CLASS				
		CYPRESS	MIXED HARDWOOD	WILLOW CYPRESS	BRUSH, GRASS, EMERGENTS	FLOATING AQUATICS	UNCLASSI-FIED
GROUND VERIFICATION CLASS	CYPRESS	98.75	0.83	0.0	0.0	0.42	0.00
	MIXED HARDWOOD	5.75	94.25	0.00	0.0	0.0	0.00
	WILLOW CYPRESS	7.07	0.0	91.92	0.0	0.0	1.01
	BRUSH, GRASS, EMERGENTS	2.38	1.79	0.0	93.45	2.38	0.00
	FLOATING AQUATICS	0.0	0.0	0.0	0.0	97.20	2.80
	OVERALL PERCENT CORRECT						95.36

it appears that the signatures developed using TM channels 2, 3, 4, 5, and 7 precisely define the five wetland land covers. Also, where confusion (misclassification) exists, it can logically be explained. For example, the willow/cypress class is confused only with land covers classified as cypress; this sort of misclassificaton is plausible because of the extreme heterogenity of the land covers within the study area. Again, classification of some pixels as "unclassified" is expected, particularly in such a heterogeneous area as Reelfoot Lake.

The results derived from this early investigation into the use of TM data for discrimination of forested wetlands are encouraging. This is particularly true considering that similar studies in the past utilizing MSS data have not obtained a product of such detail and accuracy. Moreover, MSS data collected in late August have been found to be less than optimum for the discrimination of land cover types analogous to those delineated here.

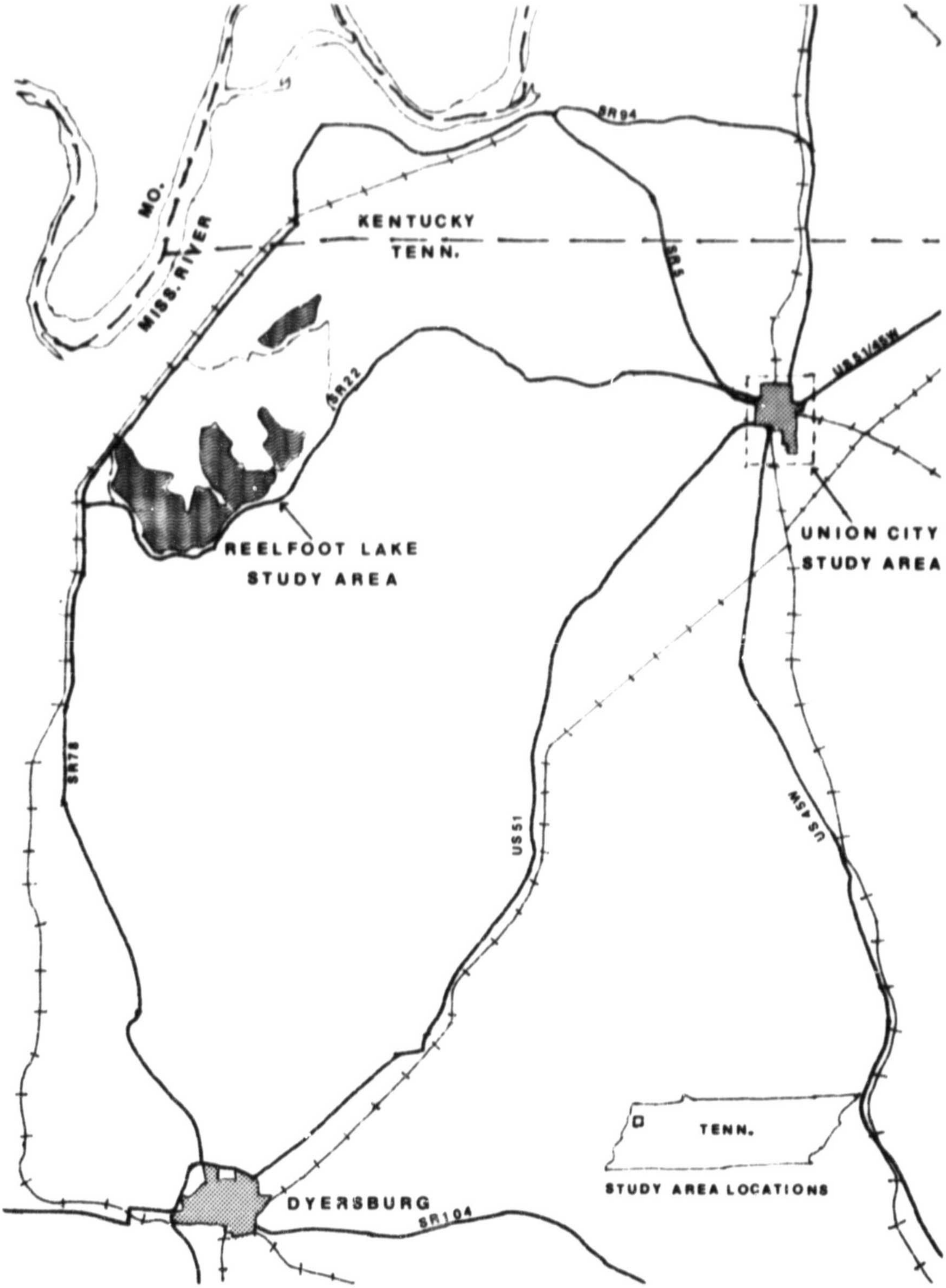
SECTION III: EVALUATION OF TM DATA FOR CLASSIFICATION OF A SMALL URBAN AREA IN NORTHWESTERN TENNESSEE

The increased spatial and spectral resolution of the Landsat 4 TM over the MSS used as the primary imaging system on previous Landsat platforms greatly enhances the utility of TM data for discriminating and classifying urban features. With its 30 meter resolution and the added information available from the new sensor in the mid-IR (1.55 to 2.35um) bandwidth range, the TM enables cultural features to be perceived from an entirely new perspective. Additionally, the inclusion of channel 6 (120 meter resolution) extends the range of data collection into the thermal IR region (10.40 to 12.50µm).

Given the new opportunities presented by these data, an initial evaluation of the TM for urban analysis has been conducted using a small city in western Tennessee as a test site. TM data were analyzed for both digital and visual characteristics. Digital processing techniques were used to produce a computer implemented land cover classification of the study area. Also, an analysis of the TM data for photointerpretation of urban features was performed.

DESCRIPTION OF THE STUDY AREA

Union City, Tennessee, located in extreme northwestern Tennessee, was chosen as a study area (Map 3). The city was selected because of its close proximity to the Reelfoot Lake area (Section II) where another TM investigation was in progress. Moreover, the small size of the town expedited the collection of ground truth information used to evaluate the land cover classification produced from the data.



MAP 3
LOCATION MAP OF UNION CITY, TN
STUDY AREA

Union City is the county seat of Obion County, and it is the major commercial center of northwestern Tennessee. The city has a population of approximately 14,650 according to preliminary estimates from the 1980 census. Union City has about eight square miles ($20.7km^2$) within its city limits. The land cover surrounding the town is predominantly agricultural, with soybeans, corn, wheat, and small grains the principal crops. Although the city is in a rural area, it does support a number of diverse light industries.

DATA PROCESSING PROCEDURE AND CLASSIFICATION ANALYSIS

Information utilized in the study was extracted from TM imagery gathered over Union City on August 22, 1982 (ID#40037-16031). A pixel-by-pixel automated signature development algorithm called Point Cluster (PTCL) in ELAS was used to derive spectral signatures within a polygon defined from the data. This polygon was selected to encompass the majority of the Union City urban area.

Figure 7 illustrates the relative frequency of occurrence for the raw TM data processed for the study area. An analysis of the seven-channel TM data, predicated on the evaluation of spectral scattergrams for selected ground truth plots, indicated that channel 1 (.45 to .52μm) did not contribute a significant amount of information which could be used to separate urban signatures. Consequently, channels 2 through 7 were used for data processing. Thirty-nine spectral signatures were derived from the unsupervised signature development procedure utilizing the six channels of data as input. These spectral responses were then classified via the gaussian maximum likelihood technique. Collectively, the signatures comprised six categories for classification: (1) roads and inert materials; (2) commercial

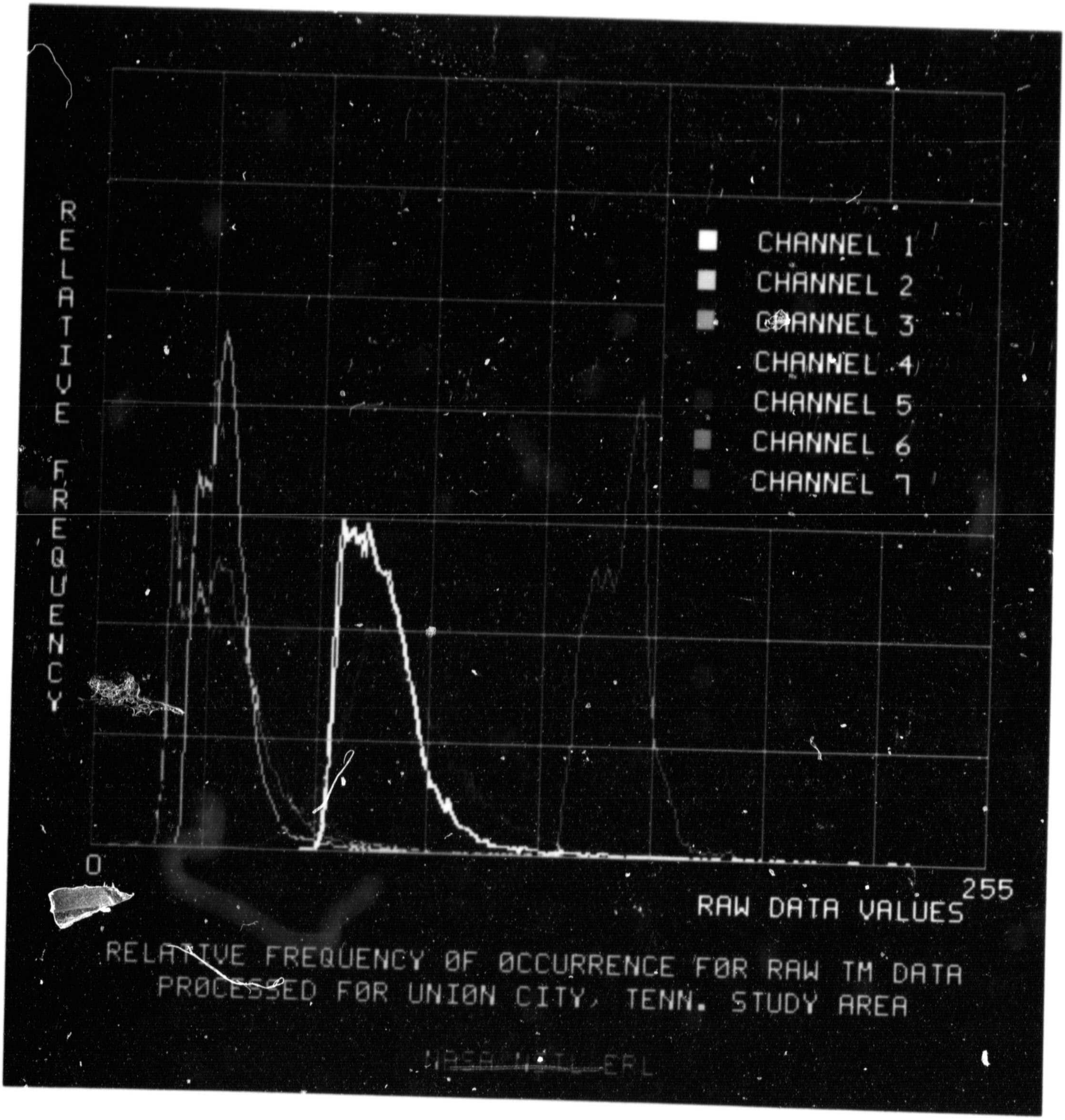

RELATIVE FREQUENCY
■ CHANNEL 1
■ CHANNEL 2
■ CHANNEL 3
CHANNEL 4
CHANNEL 5
■ CHANNEL 6
■ CHANNEL 7
0
255
RAW DATA VALUES
RELATIVE FREQUENCY OF OCCURRENCE FOR RAW TM DATA
PROCESSED FOR UNION CITY, TENN. STUDY AREA
NASA NSTL ERL

Figure 7

and industrial development; (3) residential development; (4) agriculture and bare soil; (5) transitional or grassland areas; and (6) forested areas.

Histograms depicting specific land cover signatures were produced for the classification. These histograms were used as a tool to help group each of the original 39 spectral classes into six appropriate classification categories. Of particular interest were the histograms for the commercial/industrial (Figure 8) and the residential classes (Figure 9). Three distinct types of peaks in channel 5 for the commercial/industrial histogram are indicative of three unique reflectances: red values exemplify buildings and associated features which have relatively low reflectance in channel 5 (e.g. buildings with black or dark roofs or asphalt-covered surfaces); yellow values are associated with surfaces of medium reflectance, such as buildings with medium-grey roofs (e.g., tar roofs that have been sprinkled with a light coating of gravel); and blues are indicative of high reflectance surfaces (e.g., buildings with very light-colored roofs). The residential histogram shows how high reflectance areas, such as trailer parks or newer residential areas (in white), can be separated from older, more established neighborhoods (in green) which have a greater amount of tree cover.

In addition to the histograms, ground truth data were used in naming specific classes. This information was obtained via field work and from medium altitude aerial photography taken over the Union City area. Ultimately, two different perspectives of the classified TM data for the Union City area were produced. One product represents a classification of the polygon initially used to delineate the Union City urban area for signature development purposes (Figure 10). The other photo product represents an extension of the polygon classification to Union City and its environs (Figure 11).

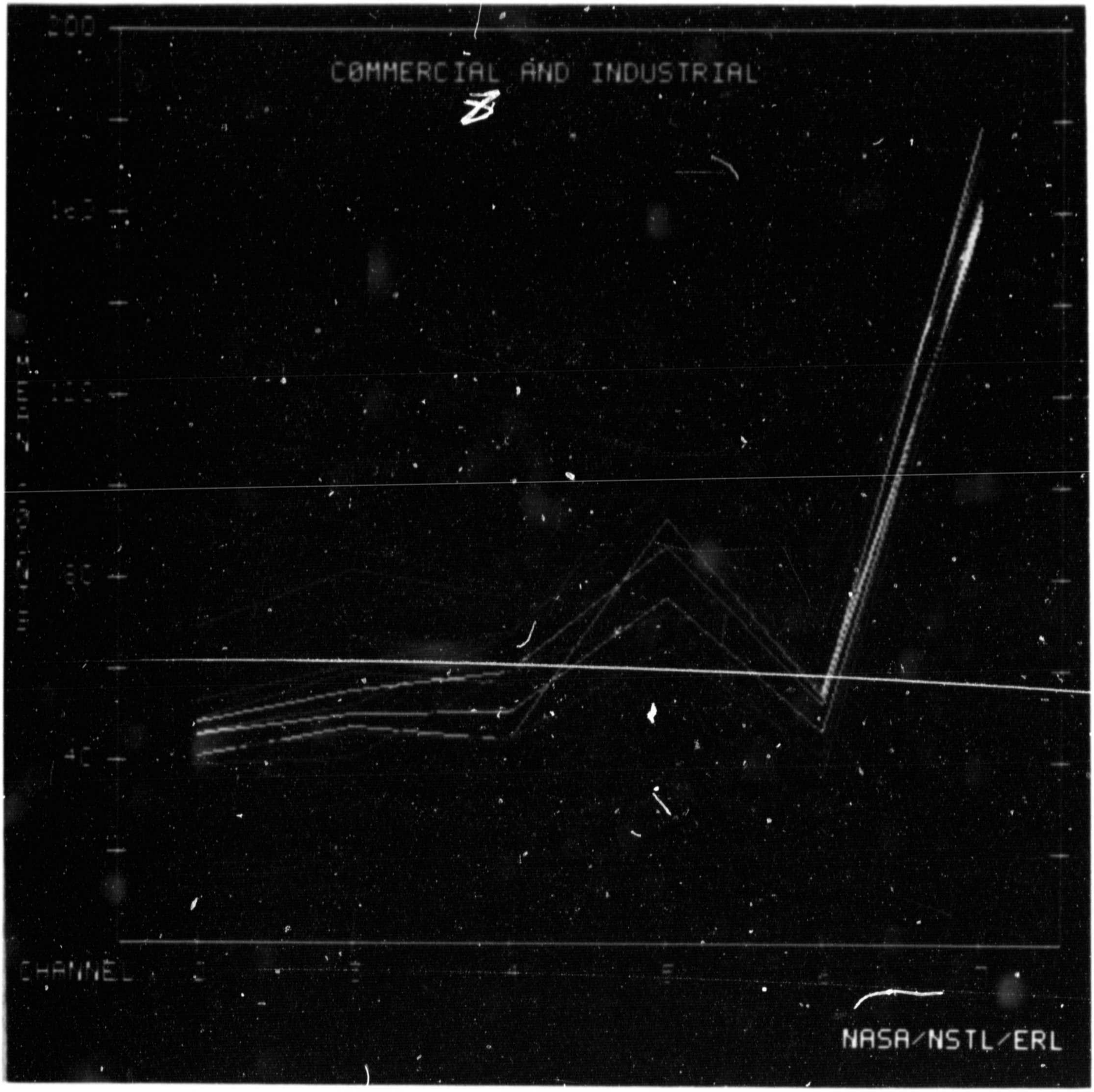


FIGURE 8
HISTOGRAM OF SPECTRAL SIGNATURE FOR
THE COMMERCIAL/INDUSTRIAL CLASS

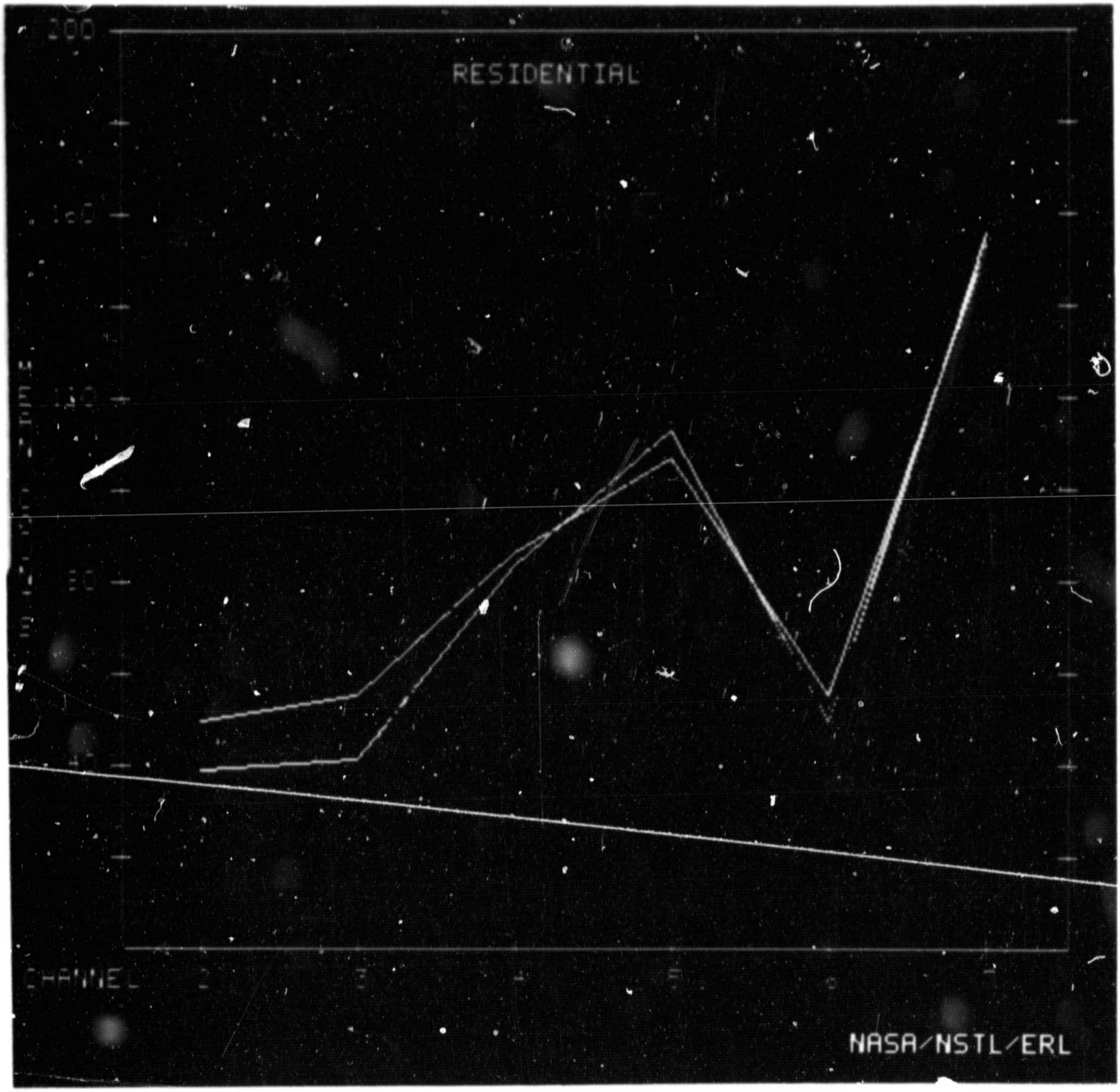


FIGURE 9
HISTOGRAM OF SPECTRAL SIGNATURES FOR
THE RESIDENTIAL CLASS


CLASSIFICATION OF UNION CITY, TN FROM TM DATA
COMMERCIAL/INDUSTRIAL
RESIDENTIAL
ROADS AND INERT MATERIALS
AGRICULTURE/BARE SOIL
FOREST
TRANSITIONAL/GRASSLAND
NASA/NSTL/ERL

Figure 10


COMPUTER IMPLEMENTED LAND COVER CLASSIFICATION
OF UNION CITY, TENNESSEE AND ENVIRONS FROM
LANDSAT 4 THEMATIC MAPPER DATA
NASA NSTL ERL

Figure 11

Classification accuracies were computed for the classified data set based on the evaluation of 45 ground verification polygons with known land covers selected from the data. Because of the difficulty in defining precise ground truth polygons for the transitional/grassland and roads/inert materials classes, polygons for these categories were combined with the agriculture/bare soil, commercial/industrial, or residential categories. Table 4 gives the accuracies for the resulting four major land cover types used to estimate classification accuracy.

The accuracy estimations in the table illustrate how much within class confusion occurred for each of the four ground truth categories. Table 4 shows that the agriculture/bare soil class had an accuracy of approximately 90%, while the commercial/industrial class produced an accuracy of about 96% correct classification. Residential areas had an accuracy of 83.50%. The confusion with agriculture was predominantly due to the overlap of spectral signatures within the transition class. Newer residential areas have fewer trees than older, more established neighborhoods. These newer subdivisions appeared spectrally similar to grassland or scrubland areas and were confused with the transition land component of the agriculture class. Forested areas produced the best classification accuracy, with a percentage correct of almost 99%. The overall map accuracy for the four major classes as compared to the ground verification data totaled 89.90% correct.

ANALYSIS OF PHOTOINTERPRETATIVE PRODUCTS

The three black and white products for the Union City area (Figures 12-14) illustrate how the improved spatial characteristics of the TM


1ST PRINCIPAL COMPONENT FROM THEMATIC MAPPER DATA
UNION CITY, TENNESSEE
NASA/NSTL/ERL
GOODYEAR
BROWN SHOE CO
CBD
DIXIE
SPIABL
DIV

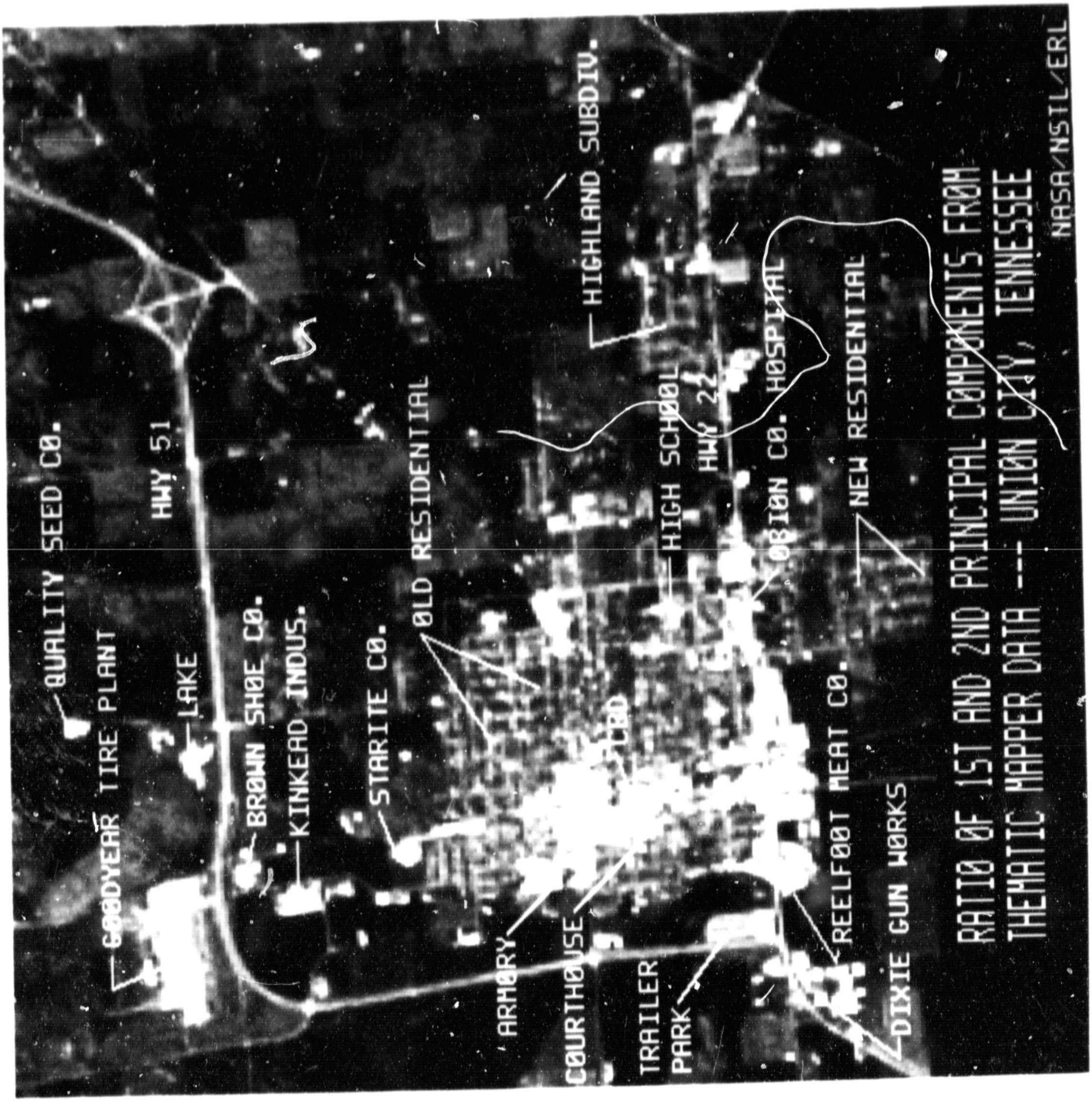
Figure 12

ORIGINAL PAGE
COLOR PHOTOGRAPH

Figure 13

ORIGINAL PAGE
COLOR PHOTOGRAPH

[redacted]



RATIO OF 1ST AND 2ND PRINCIPAL COMPONENTS FROM THEMATIC MAPPER DATA --- UNION CITY, TENNESSEE

NASA/NSTL/ERL

Figure 14

greatly augment the photointerpretative qualities of the data. To enhance the spectral contrast within the urban area, Principal Components Analysis (PCA) was applied to the original seven channels of TM data. PCA is a statistical procedure predicated on the variance and covariance of a data set. Variance is a measure of scatter or spread within one variable in the data set, while covariance is a measure of the scatter between two variables. PCA applies a linear transformation to the raw data and creates a new set of information which maximizes the variance across the data. The variance within the original data set is then redistributed in decreasing orders of magnitude within the new transformed data set. Each channel of transformed data, therefore, contains information that is uncorrelated with data in the other channels.

The first principal component (PC1) incorporates or explains the most variance and consequently, has the most information content (Figure 12). PC1 is commonly referred to as "brightness" because it emphasizes surface reflectance albedo; thus, highly reflective materials such as those in the Central Business District (CBD) and in other commercial areas appear as bright white tones. Note in Figure 12 that individual buildings and other urban components can be discerned. Moreover, different responses from the roofs of individual buildings are evident as exhibited by the reflectances from the Starite Company and Kinkead Industries located north of the Union City CBD. The Starite Company has a roof that is high in reflectance (e.g., a roof comprised of very light-colored gravel) while the Kinkead Industries building has relatively low response from its black, tar-coated roof.

PC2 contains less variability than PC1, but the information in these data is still significant. This second principal component is occasionally called "greenness" since it emphasizes vegetative characteristics, or in the case of urban areas, the absence of vegetative response (Figure 13). Street patterns can be detected in this photo product because they have a low greenness reflectance in the second component. As in PC1, individual structures can also be discriminated from the PC2 product.

To maximize the information content contained in PC1 and PC2, these transformed data channels were ratioed (Figure 14). The ratioed product increased the contrast between highly reflective buildings and more subdued urban components, such as streets or residential areas. Although both PC1 and PC2 enhance cultural features within the city, the ratioed data allowed easier photointerpretation of detail, particularly in the identification of specific structures.

From analysis of the classified and photointerpretative data products, it is evident that the TM greatly enhances the detection of urban features over what could previously be derived from Landsat MSS data. It is enticing to see that with TM data, discrete urban land covers can be discriminated using digital techniques while maintaining accuracies as high as those established in Table 4. More importantly, a town the size of Union City could barely be discriminated through photointerpretation, let alone be digitally classified as an urban area, from MSS data. Yet with the TM, it is possible to classify components of the city and visually locate and discriminate individual buildings. It definitely appears, therefore, that the TM will significantly advance interpretation of the urban milieu in comparison to what formerly could be obtained through evaluation of MSS data.

SUMMARY AND CONCLUSIONS

The investigations described in this report have explored the capabilities of TM data for discriminating land covers within three particular cultural and ecological realms. Although the work presented here has been initial in scope, the results indicate how useful the Landsat 4 TM will be in providing researchers with a new capability to monitor the Earth's environment and landscape. One scene of TM data collected on August 22, 1982 over portions of Arkansas, Missouri, Tennessee, and Kentucky, has set the background for the studies that have been outlined here.

The agricultural investigation in Poinsett County illustrated that TM data can successfully be used to discriminate a variety of crop cover types within the study area. Moreover, in comparison with a classification derived from a multitemporal MSS data set of the same area the single date TM classification produced results that were significantly better than those developed from the MSS.

For the Reelfoot Lake area, TM data processed using unsupervised signature development techniques have produced a detailed classification of forested wetlands with excellent accuracy. In fact, the nine-category classification generated for the study suggests that it is possible to extract so much detail within the wetlands that delineation of areas of sufficient size for use in ground truth verification is difficult.

It also appears that the TM is very well suited to deriving information on specific urban land cover classes. Section III of the report indicates that even in a small city of approximately 15,000 people, TM data can successfully be used to spectrally distinguish commercial and industrial sites from residential areas. Furthermore, the Principal Components Analysis of the data show that it is possible to distinguish individual buildings and roof responses with the TM.

The results from the three investigations that have been discussed here reinforce the hopes first expressed when the TM began telemetering data to the ground -- that from a purely empirical viewpoint, these new data were a revelation to the remote sensing community. Now that digitally processed and enhanced TM data have been initially evaluated in this report, we can begin to see what great potential the TM has for myriad types of scientific applications and endeavors.

REFERENCES CITED

Carter, Virginia, Alan Voss, Donald Malone, and William Godsey. 1977. Wetlands Classification and Mapping in Western Tennessee. U. S. Geological Survey, Reston, Virginia.

Graham, Marcellus H., Bobby G. Junkin, Ronnie W. Pearson, Benjamin R. Seyfarth, and Maria T. Kalcic. 1981. ELAS, Earth Resources Laboratory Applications Software. ERL Report #183.

Stoner, E. R., G. A. May, and M. T. Kalcic. 1981. Evaluation of Multiband, Multitemporal, and Transformed Landsat MSS Data for Land Cover Area Estimation. ERL Report #182.