NASA CONTRACTOR REPORT 166393

Analysis of the Tanana River Basin using Landsat Data

Leslie A. Morrissey
Vincent G. Ambrosia
Charlotte Carson-Henry

CONTRACT NAS2–11101
April 1981

NASA

NASA CONTRACTOR REPORT 166393

Analysis of the Tanana River Basin using Landsat Data

Leslie A. Morrissey
Vincent G. Ambrosia
Charlotte Carson-Henry
Technicolor Government Services

Prepared for
Ames Research Center
under Contract NAS2-11101

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035

N83-17927 #

Analysis of the Tanana River Basin using Landsat Data

Tanana Demonstration Project

Alaska ASVT

Prepared in cooperation with:

Alaska Department of Natural Resources
Division of Research and Development
Division of Forest, Land and Water Management

By:

Leslie A. Morrissey
Vincent G. Ambrosia
Charlotte Carson-Henry
NASA Ames Research Center

April 1981

ABSTRACT

The Tanana River Basin Project, as part of NASAs Applications System Verification and Transfer (ASVT) program, is a cooperative effort with agencies in the state of Alaska to assess the utility of remote sensing technology in meeting critical resource management needs, while providing DNR with baseline resource information. Digital image classification techniques were used to classify land cover/resource information in the Tanana River Basin. Portions of four scenes of Landsat digital data were analyzed using computer systems at Ames Research Center in an unsupervised approach to derive cluster statistics. Identification of the spectral classes was accomplished using the IDIMS display and color-infrared photography. Classification errors, identified during two workshops with DNR personnel, were corrected using stratification procedures. The classification scheme resulted in the following eleven categories: sedimented/shallow water, clear/deep water, coniferous forest, mixed forest, deciduous forest, shrub and grass, bog, alpine tundra, barrens, snow and ice, and cultural features. Color coded maps and acreage summaries of the major land cover cateogories were generated for selected USGS quadrangles (1:250,000) which lie within the drainage basin. The Tanana Project, initiated in October, 1980, was completed within six months.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Scope of the report

The Tanana Project, as part of NASA's technology transfer program, is comprised of two subprojects: generation of a regional land cover map of most of the Tanana River Basin using Landsat digital data and a forestry resource inventory for an area northwest of Fairbanks. This report was prepared to summarize the regional resource inventory.

Chapter 1 provides background information concerning overall project objectives, participants, and agency needs related to legislative mandates. An overview of the technical approach, responsibilities, desired output products, and delineation of the study area are discussed in Chapter 2. Chapter 3 examines in detail digital processing of the Landsat data. This document should assist agency participants in assessing the utility of remote sensing technology in relation to specific state needs. The final chapter is to be completed by state agency participants and will be included in the final report.

Introduction

The Tanana Project is a cooperative effort with agencies within the state to demonstrate and facilitate the operational use of remote sensing technology in meeting state inventory objectives in a timely and cost efficient manner. While the Tanana project was conceived as a demonstration project, the main goal was to provide the Department of Natural Resources with land cover information which will be utilized in the Tanana Basin Area Plan as a basis for initial delineation of potential settlement areas, substantiation of forestry public interest lands, fire planning, habitat verification, and identification of areas requiring more detailed information. Moreover, the project initiated in the fall of 1980, needed to be completed the following spring to be useful in DNR's planning process.

Completion of the analysis phase of the project will serve to meet three main objectives: 1) to provide resource inventory information useful for the development of a general management plan, 2) train agency personnel to understand the advantages and limitations of the technology, and 3) serve as a basis for evaluation of the technology for meeting specific state needs.

Recent and impending legislative mandates in the state require transfer of millions of acres of land to federal, state, local agencies and Native groups. One of those mandates, the Alaska Statehood Act, tranferred 104 million acres to state jurisdiction, while the Land Policy Act (1978) mandates regional management plans for state lands selected under the Statehood Act. To meet these critical needs the Division of Research and Development is responsible for developing a comprehensive regional plan for the Tanana River Basin. Land transfer requirements of these acts dictate statewide land inventory and use - classification, now largely non-existent in Alaska.

Considering the size of the state, complicated and changing land ownership patterns, inaccessiblity, and high cost of land - based inventories, coupled with numerous legislative mandates, the availability of color infrared (CIR) aerial photography and Landsat data provides an ideal data base to meet these various inventory requirements. In addition, analysis of Landsat digital data will provide state agencies with information readily compatible with geo-based information systems.

Background

Landsat digital data, used in conjunction with high altitude color infrared photography, can provide agencies with accurate resource information in a timely manner which becomes increasingly cost effective for large project areas such as the Tanana River Basin. Additionally, digital image classification provides rapid, consistent classification decisions where each picture element is assigned to a resource class. The use and application of the final output products is an important factor when considering image classification techniques. The digital format provides an efficient means for summarizing land cover statistics within a specified areal unit. Digital classification results provide greater flexibility in displaying data in various formats and map scales and results can be aggregated to simulate various minumum mapping units.

Photo-interpretation of land cover information can provide polygonal information for extensive regions given the availability of aerial photography or satellite imagery. The resultant polygonal format of the data base, however, is not particularly useful. It is difficult and laborious to convert polygonal data into a digital format which would facilitate integration with other data sets. In contrast, the digital format of Landsat data provides a capability for easy data extraction and manipulation.

A digital inventory base, used in conjunction with other data, offers a multitude of unique opportunities for specialized study needs. The successful merging of Landsat resource information with other data sets can dramatically extend the usefulness of the digital data base and begin to provide innovative solutions to many problems confronting resource managers and planners. In addition, the initiation of specific applications utilizing the generalized land cover inventory will provide a basis for final assessment of the remote sensing technology by participating agencies. The Tanana Project has been designed to test the usefulness and feasibility of utilizing remote sensing technology in a wide variety of situations and applications. The following list summarizes possible applications where the landcover inventory might prove useful.

Land use/cover inventory
Wetland delineation
River Basin study support
Floodplain mapping
Water quality monitoring
Change detection applications
Land use suitability mapping
Baseline data for GIS integration
Habitat evaluation

Integration of the land cover maps produced from Landsat digital data with other ancillary data sets in the Alaska Land and Resources System (ALARS) geo-based information system would further enhance the usefulness of the land cover information. Many facets of land capability can be keyed to land cover information. For example, areas with impeded drainage, saturated soils and permafrost can often be infered from a black spruce bog vegetative cover. In addition, the land cover map used in conjunction with landform, slope, aspect and other data can be reinterpreted to represent desired land capability categories. Follow-on subprojects, identified by each of the participating state agencies, will be supported with technical assistance from NASA in utilization of the digital data base for specific applications during the third year of the ASVT.

Project Responsiblities

One of the major factors determining the success of a technology transfer project is the level of commitment and contribution of participating personnel. This demonstration project requires active agency participation in all phases of the project to be able to adequately assess the utility of the technology for meeting state needs. The major responsiblities of NASA and participating state agencies are summarized below.

NASA's Role

Information and Training
Image Analysis Processing
User needs survey

Overall project support
Provide satellite and aircraft imagery
Conduct training workshops

Participating State Agencies

Ground data collection
State coordinating functions
Identification of applications needs
Active participation
Commitment of personnel
Technology evaluation

The cooperative nature of the demonstration project requires active participation of agency personnel. At the completion of the project, the burden of evaluation (the most important phase of the project) will reside with the agencies. The final assessment of the technology, hence the success of the project, is solely determined by agency personnel.

Agency Participation

The following divisions within DNR are actively participating in this demonstration:

Alaska Department of Natural Resources
Division of Research and Development
Division of Forest, Land and Water Mangement

The following agencies are providing project support of demonstration activities, such as field data collections, training course attendance, and evaluation of output products:

Alaska Department of Fish and Game

Alaska Department of Natural Resources

- Division of Technical Services
- Division of Research and Development
- Division of Forest, Land and Water Management
- Division of Minerals and Energy Management

Alaska Department of Community and Regional Affairs

Alaska Department of Environmental Conservation

U.S. Department of Agriculture

- Soil Conservation Service

CHAPTER II

PROJECT CONSIDERATIONS AND ORGANIZATION

Technical Overview

An unsupervised classification of Landsat digital data covering the Tanana River Basin has been developed to provide a regional land cover inventory base.[1] Four Landsat scenes falling within the basin were acquired. Portions of two scenes within the same orbital path were mosaicked to create one central scene, in addition to the scenes located east and west. The spectral information for each scene was analyzed using clustering algorithms to yield spectral statistics. The unsupervised cluster statistics were then used to classify (assign) each picture element to the cluster it most nearly resembled using a maximum likelihood classifier. Classification results were evaluated by DNR personnel to identify the resource characteristics of each spectral class. This step was completed using the interactive display at the EROS Field Office in Anchorage. Each spectral class was examined using high altitude CIR photography. Classification errors or classes which represent more than one resource category, were identified for later stratification. A number of enhancement techniques were utilized to resolved these errors. A series of color coded maps of the eleven land cover categories with acreage summaries by cover type were produced at 1:250,000 scale.

1 In a unsupervised approach, the MSS data is paritioned into an arbitrary number of spectrally distinct groups of clusters.

Project Roles

The successful completion of the digital analysis project depended, to a major extent, on the participation of DNR personnel for spectral class identification. In addition to the project objectives stated in Chapter 1, specific support provided by DNR included:

Map of the drainage basin boundaries

CIR photography for areas within the eastern and western Landsat scenes

IDIMS systems cost (second session)

Landsat color composite prints

Assessment of final products

Evaluation and documentation

Two evaluation workshops were held at the IDIMS EROS Field Office during the months of January and February. P. McNees and J. Wehrman provided expertise in the interpretation of spectral classes for the three data sets. It must be noted that the overall classification accuracy is proportional to the amount of time spent in the spectral class evaluation phase. Due to the limited coverage of aerial photography and ground information, the classification scheme used in the analysis was modified during the workshop to maintain a high degree of reliability. Utilization of the final classification by DNR will provide additional time for further evaluation of results, and potentially, a refinement of the eleven land cover categories by data set.

Study Area

Selection of the study region was determined by DNR and the availability of Landsat data. The main goal of the project was to provide DNR with land cover information which would be utilized in their Tanana Basin Area Plan. In addition, a Landsat search of available imagery revealed a number of possible scenes. Four scenes were chosen because they encompassed most of the lands within the Tanana Basin under state jurisdication.

The study region, located within the Tanana River Basin, extends north to Livengood, south to the Alaska Range, west to the confluence of the Yukon and Tanana River, and just east to George Lake (Figure 1). Along the southern boundary, the elevation of the Alaska Range varies from 6,000 to 9,000 feet culminating with Mount Denali (20,269 feet). North of the Alaska Range, the foothills reach elevations of 2,000 feet. The lowlands south of the Tanana river are characterized by low, gentle terrain. North of the Tanana river the uplands consist of forested mountains up to 6,000 feet in elevation.

The test site is traversed by major highways, the Alaskan railroad, and the Alaska pipeline. Fairbanks is the largest city with numerous other smaller settlements and a number of military installations. Land ownership patterns have undergone radical change due to numerous state and federal legislative mandates regarding land selection, disposal and classification. As a result, a need for cooperative land management exists. In addition, the checker - board ownership patterns have increased

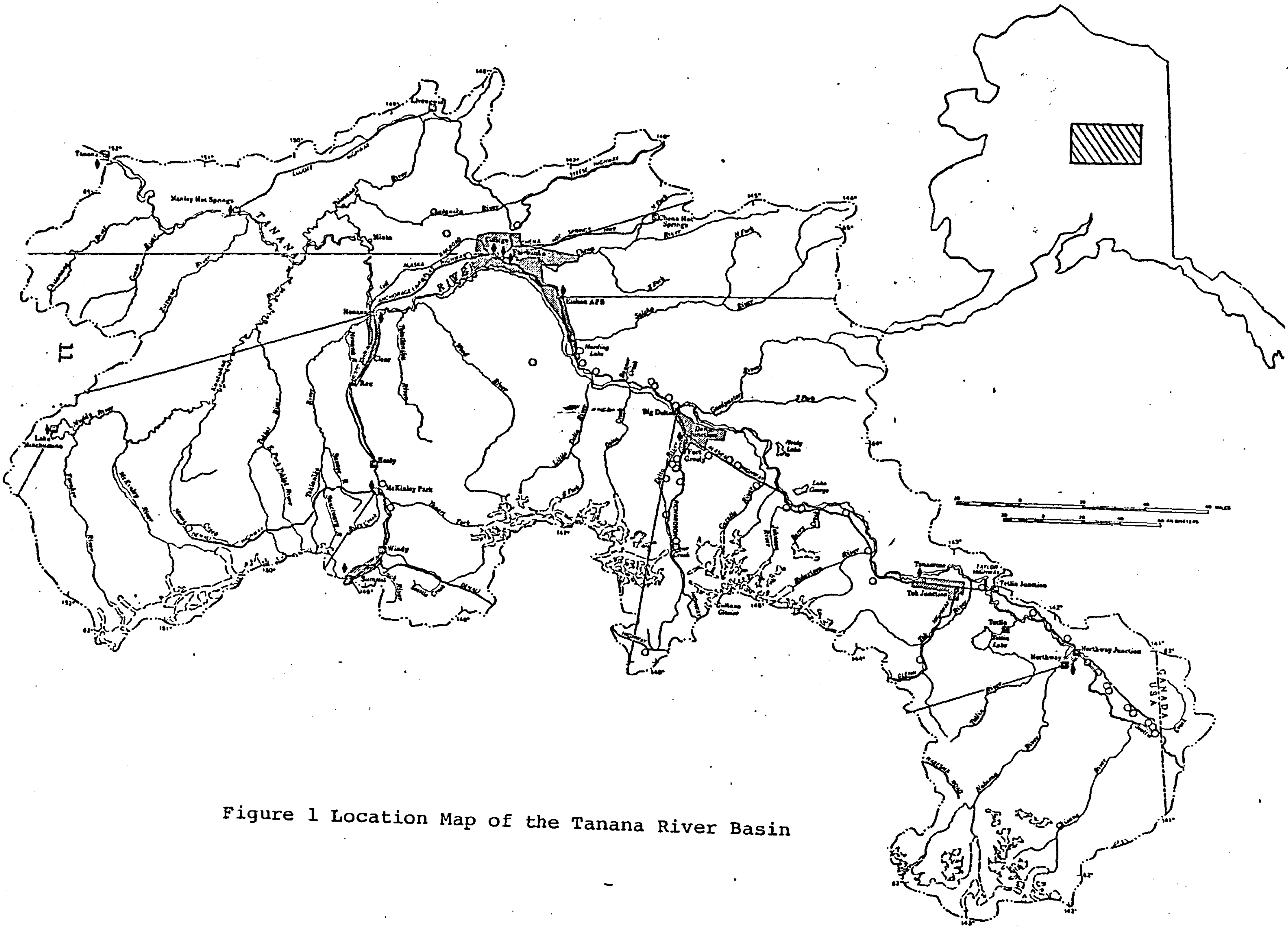


Figure 1 Location Map of the Tanana River Basin

accessibility problems. The principal political subdivision within the basin is the Fairbanks North Star Borough. However, federal, state and local agencies have jurisdiction of lands within the basin.

Classification Scheme

The classification scheme used to characterize the Tanana Basin was devised to satisfy the needs of a number of state agencies in an effort to demonstrate the use of Landsat satellite data, while providing a reliable land cover inventory for the region. The classification scheme was generalized to maintain a high degree of accuracy and consistency in all three data sets, although greater detail exists in each individual scene. For instance, DNR was interested in differentiating shrublands above and below the treeline. Hoewever, the spectral reflectance of shrubs at any elevation appear identical. The resulting classification scheme, therefore, evolved to reflect the characteristics of the satellite, expertise of the individual photo-interpreters, and limitations of time. The land cover categories were classified according to dominant lifeform, site moisture and other associated landform elements. The classification scheme developed for the Tanana Basin, based on the analysis of Landsat digital data, resulted in the following eleven land cover/use categories:

Cultural

Commercial, industrial, transportation, and residential land uses comprise this class including the two major urban areas of Fairbanks and Big Delta. Growing crops, fallow fields and pastures are also included in this class.

Grass and Shrub

Encompasses those communities with a major vegetative component of shrubs, grasses and sedges. This class occurs as tundra above the treeline, as the primary successional community on timber cuts, and along river floodplains. It may also include up to 30% tree cover.

Coniferous Forest

This class includes any areas which have a dominant crown cover of coniferous trees (white and black spruce). It includes open and closed coniferous forests and may have up to approximately one-third deciduous crown cover. In addition, black spruce bogs with a forested crown cover of greater than 30% are included in the class.

Mixed Coniferous/Decidous Forest

This class occurs whenever the deciduous or coniferous vegetative components cover at least one-third of the overall vegetative cover.

Deciduous Forest

Encompasses open and closed forests with deciduous trees as the major vegetative component (greater than 30%). This class can include up to one-third coverage of conifers; usually the conifers present were below the deciduous crown canopy. Main species found include birch and aspen with an understory of various forbs, grasses and shrubs.

Barrens

All bare rock or soil surfaces with less than one-third vegetative cover. Includes bare rock outcrops on the major mountain ranges, floodplains, sandbars, gravel pits, rubble on glaciers and some north-facing slopes in shadow.

Alpine Tundra

Basically, mat and cushion, low-lying vegetation commonly found above the treeline (berries, prostrate shrubs, lichens).

Snow and Ice

This class includes all perennial snow and ice fields in the major mountain ranges surrounding the Tanana Basin.

Shallow and Sedimented Water

Shallow lake shelves, turbid lakes, deltaic plumes, and rivers and lakes with high sediment loads comprise this class. Emergent and aquatic vegetation may also be present.

Deep and Clear Water

Includes many of the hundreds of inland lakes and stream courses in the Tanana Basin.

Bog

That community comprised of a black spruce or shrub overstory with an understory of mosses, berries and small shrubs. Bogs are basically found in two distinct areas: depressional basins and on north-facing mountainous slopes. This cover type is characterized by impeded drainage, saturated soils, and, in many cases, is underlain with permafrost.

The original classification scheme proposed by DNR was modified to reflect the spectral and ground truth information. The information classes were generalized, when necessary, to maintain a high level of accuracy. Spectral classes which were not consistently identified during the evaluation workshops by agency participants were grouped into a more generalized information class. For example, spectral confusion between Black and White Spruce necessitated their grouping. Black Spruce occupy depressional basins with impeded drainage, while White Spruce are found on slopes with integrated drainage. These two species could be visually differentiated with a map overlaid on the color-coded classification if the interrelationship of vegetation and terrain is known. To reiterate, the Landsat classification will provide more reliable and useful information when used in conjunction with ancillary data.

Output Products

The digital format of the final classification is condusive to a variety of display methods. Line printer maps, electrostatic printer/plotter maps (in which the spectral classes are represented by symbols), and photographic products generated from a film recorder are common types of output products. A fourth product is a tabular summary of picture elements by land cover category. During early planning discussions with DNR, NASA agreed to provide the following products for selected 1:250,000 U.S.G.S. quadrangles within the Tanana River Basin:

Color-coded maps (two sets) of the classified data

Acreage summaries by quadrangle

8" X 10" Polaroid prints of selected areas

In addition to providing DNR with immediate products, NASA will reproduce color coded prints of each of the three data sets (at a scale of 1:500,00) for evaluation by other participating state agency personnel.

Training Workshops

In order to assist state agency participants in developing an understanding of the advatages and limitations of remote sensing technology, two training workshops sponsored by NASA, were held in Anchorage. The first workshop centered on the principles of photo-interpretation using color-infrared high altitude photography. The workshop was presented at the EROS Field Office (Anchorage) in June, 1980. Participants included five personnel from the Department of Natural Resources, three from the Alaska Department of Fish and Game, two from the Department of Community and Regional Affairs, and one from the Department of Environmental Conservation. Participants were given a set of aerial photographs for the Fairbanks region, a course notebook and instruction in ground data collection procedures.

During a second workshop, state agency participants were given hands-on training on the Interactive Display Image Manipulation System (IDIMS), held at the EROS Field Office the last week in October, 1980. Participants to the introductory digital workshop included four from the Department of Natural Resources, one from the Department of Environmental Conservation, one from the Department of Community and Regional Affairs and two from the Soil Conservation Service. All participant had previous photo-interpretation training; hence made an easy transition to Landsat digital data. The use of CIR photography to increase the usefulness and reliability of the Landsat derived information was emphasized in the digital workshop.

CHAPTER III

TECHNICAL ANALYSIS METHODOLOGY

Digital analysis refers to the steps involved in the generation of a classified Landsat scene. Simply stated, in the digital process the reflectance characteristics of each picture element are correlated with resource categories. This is possible because the reflectance values are recorded in a digital (numerical) format easily manipulated with computers. This chapter describes the various steps in the digital analysis of the Tanana Basin (Figure 2).

Digital analysis of the three data sets utilized in the Tanana Basin encompassed five major steps: clustering, classification, evaluation, stratification and generation of output products. In an unsupervised clustering, the multispectral (MSS) data is clustered into a number of spectrally distinct groups to develop statistics (means, variances). These unsupervised cluster statistics are then utilized in the classification of each Landsat scene, a process in which each picture element (pixel) is assigned to the statistical cluster it most nearly resembles. Evaluation of the classification involved identifying the spectral classes by comparing the classes (viewed on the IDIMS Comtal display) with color infrared photography. Classification errors, identified during the evaluation workshops, are the result of spectral confusion between two distinct land cover categories. Stratification techniques involve renaming "problem"

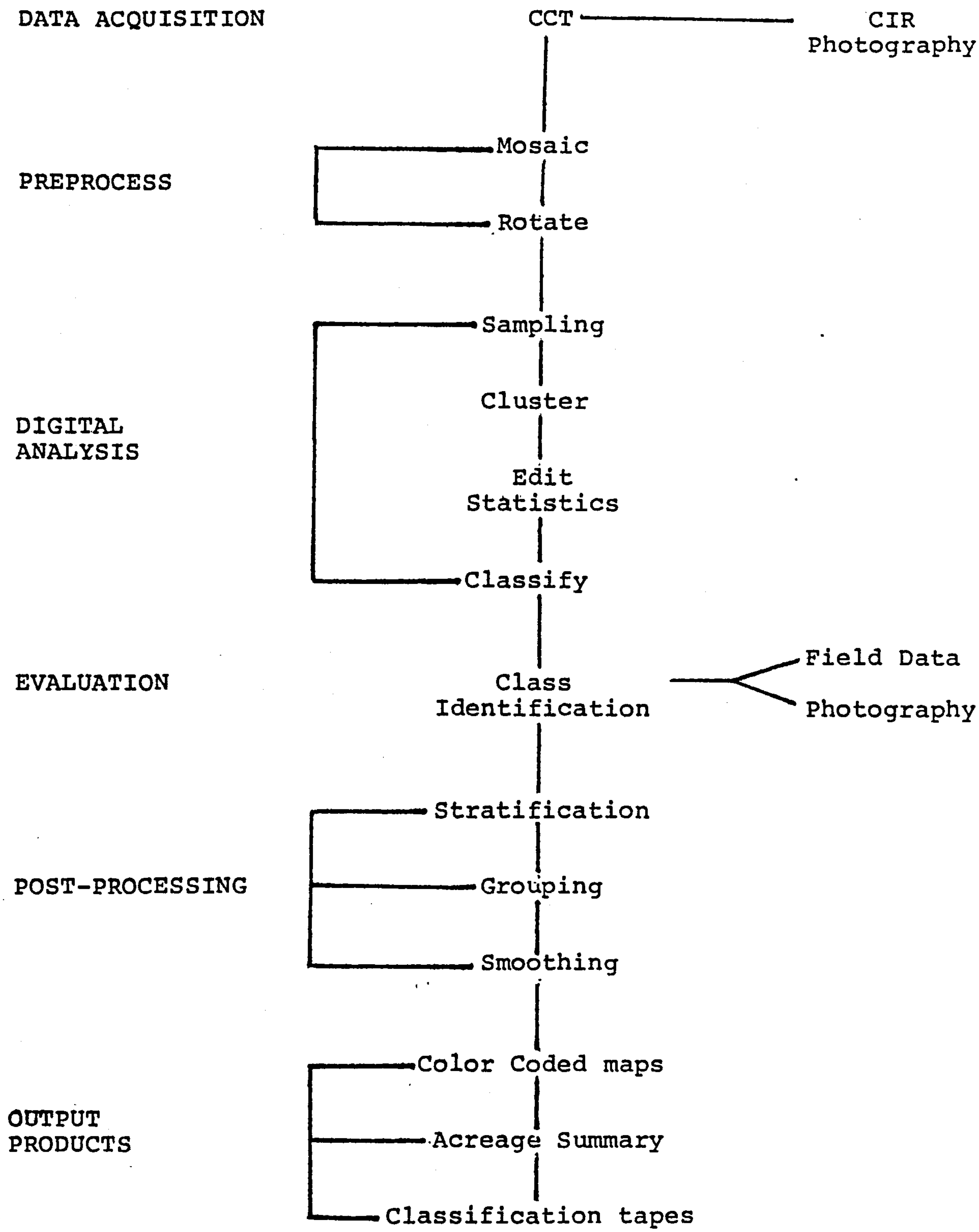


Figure 2 Major Analysis Steps

classes within specified areas. Upon completion of the stratification step, output products, color coded maps and acreage summaries, were produced.

Imagery Acquisition

The initial step in the digital anlysis of Landsat data involves the acquisition of imagery and ground data information. Classification of the Tanana River Basin required the following data inputs:

Landsat digital data (in the form of computer compatible tapes)

False-color composite photographic prints

Field data

Color-infrared U-2 photography

Landsat scenes were selected on the basis of cloud cover, season of the year, and orbital path. Coverage of most of the Tanana Basin lands under state jurisdiction (Figure 3) was obtained with portions of the following Landsat scenes:

2583 - 20190	August 27, 1976
30918 - 20501	September 8, 1980
30555 - 20433	September 11, 1979
30555 - 20435	September 11, 1979

Two of these scenes (September 11, 1979) are within the same orbital path, and the other two scenes are from orbital paths to the east and west. The 1979 Landsat scenes were mosaicked to form a central scene prior to digital analysis. This central mosaicked scene, in addition to the other two scenes, comprise the three data sets utilized in the analysis. Computer compatible

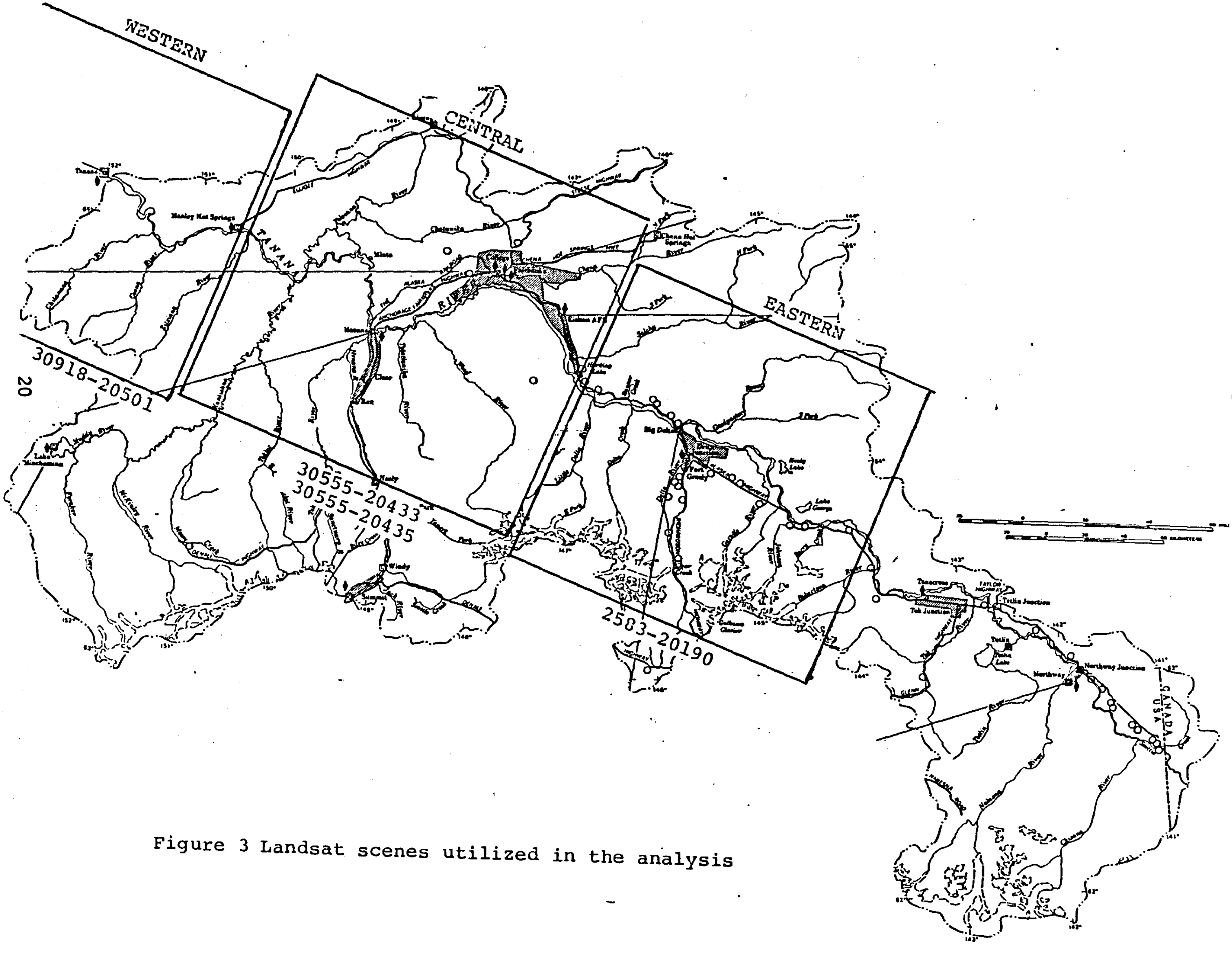


Figure 3 Landsat scenes utilized in the analysis

tapes (CCTs) were purchased from EROS (Earth Resources Observation Systems) Data Center in Sioux Falls, South Dakota. Color-infrared photography (CIR) provided a necessary link between ground visits and the Landsat digital data with the development of photo-interpretation keys which allow the analyst to extrapolate information gathered during ground visits over larger areas. The CIR photography is primarily useful for collection of ground data and spectral class verification.

A search of available CIR photography was made at the U-2 Data Facility located at Ames Research Center. Flight lines of all CIR photography falling within the study region were plotted on 1:1,000,000 scale map (Figure 4). Photography with excessive cloud cover were eliminated from further consideration. Each of the participating agencies was given selected photographic coverage (Appendix A) of areas they had designated high priority NASA also maintained a duplicate set of photographs for evaluation of the classifications.

Computer Systems

Analysis of the Tanana scenes utilized several computer systems available at Ames Research Center. The EDITOR (ERTS Data Interpreter and TENEX Operation Recorder) software package developed by the Center for Advance Computation (CAC) provided the mainstay for interactive image analysis. The EDITOR package is available on the TENEX PDP-10 at the computer facility of Bolt, Berenak, and Newman (BBN) in Boston, and is accessed through telephone lines via the ARPA (Advanced Research Projects

Figure 4 CIR Photographic Coverage

Agency) Network. Development of cluster statistics for the eastern scene, generation of output products, and evaluation of the spectral classes was completed using IDIMS (Interactive Digital Image Manipulation System) software and Comtal display implemented on the HP 3000 computer at Ames. Large scale clustering and classification required the use of the ILLIAC IV, also on the ARPA Network, while preprocessing, reformatting and acreage summaries were completed using the IBM 360/67. The CDC 7600 was used for bulk processing steps: classification of the western scene and post-processing tasks. Access to several computers allowed use of the most efficient system for each specific processing task.

Field Data Collection

The accurate classification of land use and land cover information with Landsat multispectral data requires the careful collection of ground data throughout the study region. In an unsupervised clustering approach ground data provides the basis for identification and evaluation of spectral classes. In addition, ground excursions coupled with the recognition of spectral response on the aerial photographs supplements photo-interpretation training for state agency personnel.

Personnel from a number of state agencies participated in the ground data collection effort in the summer of 1980. Each agency was responsible for a specific area within the drainage basin. NASA provided CIR photography, coordination, and instruction in the acquisition of training site information in support of Landsat analysis. The procedure for establishing training sites is shown in Appendix B.

To familiarize agency personnel in the recognition of major vegetative species, a photo key was completed containing ground photographs of major tree species along with written descriptions. In addition, data collection forms (Appendix C) and a preliminary classification scheme (based on Vierecks) were also included in the ground data collection package. Prior to the onset of the field season, and as part of a photo-interpretation workshop sponsored by NASA, personnel participated in field orientation excursions in Anchorage and Fairbanks. One day was spent in the field to familiarize personnel with major vegetative communities and most importantly, to standardize the completion of the data forms.

During the first week in July, a TGS analyst accompanied project team members into the field. Ground truth forms were completed for each training site and ground photographs were taken to provide a permanent record of each site. In addition, an aerial overflight provided reconnaissance information for inaccessible areas south of the Tanana River. Upon completion of the field season, ground site information was compiled for use in the digital analysis described in the next section.

Preprocessing

Preprocessing of the data prior to classification included removal of bad data lines, geometric correction and rotation of the data to north. Preprocessing of the three scenes varied and each will be described individually.

Preprocessing of the eastern scene using the IBM 360 involved the removal of skew (caused by the earth's rotation beneath the satellite), and rotation of the picture elements to a north orientation for ease of use. Skew parameters for this type of geometric correction were computed using a program available on the IBM 360. The corrected data was then reformatted for use on the Illiac and BBN TENEX systems.

The Landsat data sets for the central and western scenes were processed at the EROS Data Center. Beginning in 1980, standard processing of Landsat MSS data at EROS includes radiometric correction, haze removal and geometric correction to a 57 square meter pixel size. The western scene was missing one-fourth of the data due to line start problems with the scanner. Despite this problem, the scene was utilized because the missing data fell outside the study region. In addition to the line start problem, 10 lines of data in the upper portion of the scene were only intermittently recorded.

The central data set is a composite of portions of two Landsat scenes. The north half of the southern-most scene and south half of the norther-most scene were mosaicked using the VICAR software package available on the IBM 360. Greyscales of the areas of overlap of the two scenes were utilized to determine the degree of overlap of the two data sets. Upon completion of mosaicking the data sets, the composite central scene was then reformatted for use on other computer systems.

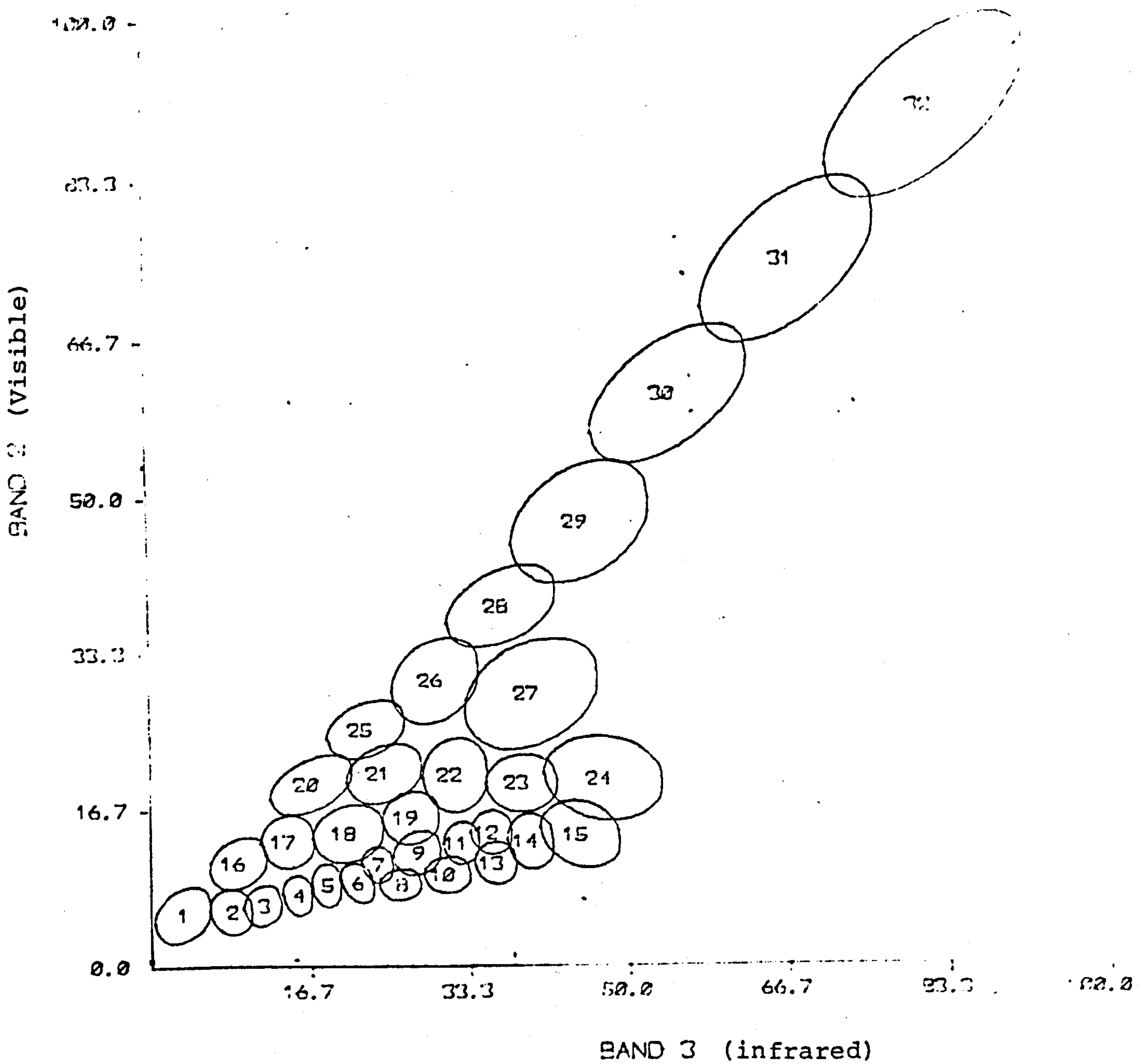
Clustering and Classification

Analysis of the central and western scene was completed using the EDITOR software package and the ILLIAC IV. A data reduction program in EDITOR enables clustering of an entire Landsat scene on the Illiac IV. The program maintains a count of the number of pixels which have the same reflectance values in the four bands. Though the program will condense an entire scene, only 25% of the data set was sampled (every second line and column) for each Landsat scene. This large sample increased the probability of sampling non-extensive land cover types. This "weighted file" created with the condensing program was then transferred over the ARPA Network to the Illiac IV where it was clustered into a number of spectrally distinct clusters as specified by the analyst. Each scene was clustered twice, specifying 30 and 40 clusters. Each clustering provided a set of statistics, including a mean, variance, and covariance matrix of the four-band data for each cluster. Preliminary examination of the statistics, including separability (distance between clusters), determines whether the clusters were statistically acceptable. Clusters with few pixels or high variances were deleted while clusters with low separability and low variances, indicating spectral overlap, were pooled. Statistics from each clustering were merged and edited. The central scene utilized 31 clusters (Appendix E) and 36 clusters for the eastern scene (Figure 5). The statistics for each scene, as well as the MSS data, were transferred to the Illiac IV to perform a maximum likelihood classification, assigning each pixel to the cluster it most nearly resembled. The assumption is made in the unsupervised approach that spectral classes - individualy or collectively correspond to desired resource categories.

The development of cluster statistics for the western scene was accomplished using the IDIMS software package. Random boxes were located throughout the scene using the RANDSAMP program. The RANDSAMP created a composite image comprised of all the sample boxes, with each box being 15 x 15 pixels, and the total sampled pixels reflecting a 5% sampling of the entire data set. This composite RANDSAMP output was then input to the ISOCLS clustering algorithm on IDIMS. A maximum of 30 clusters were requested; the minimum number of pixels per cluster was specified as 50, with clusters to be split if the standard deviation exceeded 2.5 while containing over 100 pixels. The ISOCLS recommended pooling of two sets of clusters, but all thirty were maintained separately, in preference for post-classification grouping. The statistics file generated by ISOCLS was reformatted for use on the CDC 7600. The statistics and MSS data, was submitted to the maximum likelihood classifier on the CDC 7600.

Figure 5

Statistical Plot of the Eastern Scene



A two-dimenstional representation of spectral space is shown. Each number represents the location of the cluster mean, while the circle portrays the variance.

Due to extremely low variances (below 0.1) two of the clusters were not used by the classifier. Of the remaining 28 spectral classes, five classes represented background information on the scene. A two-dimensional plot of the cluster statistics for the western scene is shown in Appendix D.

Identification of Spectral Classes

Spectral class identification was accomplished using the IDIMS Comtal display by enlarging areas within the three scenes and comparing these same areas to available CIR photography. Individual classes were identified and evaluated to ensure consistent identification throughout each scene. Two digital workshops were held at the EROS Field Office in Anchorage. P. McNees, J. Wehrman, and N. Mahalow participated in the evaluation process. A number of classification errors were located and stratification strategies devised. Generally, spectral confusion was found to occur between physiographic provinces. For example, on the central scene one spectral class represented deciduous forests north of the Tanana River and shrub tundra on the foothills of the Alaska Range. Once identified, problems could be corrected with a number of stratification techniques.

The reliability of the identification of spectral classes reflect three major components: 1) the expertise of the photo-interpreter, 2) the correlation of spectral classes to resource categories, and 3) the availability of CIR photography. A map of CIR photography flight lines (Figure 4) indicate areas in the drainage basin which have photographic coverage and, therefore, a higher degree of reliability for the digital data. In contrast, areas where there was not any aerial photography were not checked during the workshops and may, therefore, have a lower reliability. The spectral classes were grouped into one of the eleven desired land cover categories specified by DNR. Individual spectral classes, however, represent a greater level of detail than is found in the final grouping. For instance, the individual spectral classes, collectively identified as deciduous forest, represent varying crown densities, species composition and aspect. The constraints of field information, time, and budget necessitated a more generalized grouping of land cover categories which could be consistently identified on the CIR photography.

Post-Processing

Post-processing refers to digital processing steps which are performed after clssification and may include geometric correction, stratification for refinement of the classification, grouping of the spectral classes into desired resource categories and smoothing to simulate a minimum mapping unit. Geometric correction of the data was performed prior to classification and was, therefore, is not part of the post-processing for the Tanana scenes.

During the evaluation workshop, spectral confusion of classification errors were noted and strategies devised to correct the problem areas. There are two basic sources of classfication error: 1) distinct land cover categories with similar spectral characterists, and 2) poor identification of spectral classes. Stratification techniques are available in which the misclassified pixels are identified, and reassigned within specific polygons. The strategy works well when spectral confusion can be separated along physiographic boundaries. For instance, in the Tanana River Basin, most stratification was accomplished within physiographic provinces: the Tanana lowlands, the Alaska Range and foothills, and the uplands north of the Tanana River. More important, the correct identification of spectral classes is required prior to stratification and is dependent on two major factors: availability of CIR photography for checking classes and the experience and ability of the photo-interpreter. This phase of the project is exceedingly time consuming and of utmost importance in the generation of accurate and reliable land cover information.

Stratification of the three scenes was completed on IDIMS using the ZIP function. An area of misclassified pixels were located and digitally isolated within polygons delineated using the Comtal track ball to locate vertice points. Pixels within the analyst specified polygons were then reassigned to a new spectral class number. The new spectral class was then assigned to the land cover category it actually represented on the ground, therby correcting the initial spectral confusion.

Upon completion of stratification, the spectral data is then grouped into one of the eleven desired land cover categories and then smoothed to a nine acre minimum mapping unit. The grouping of the spectral classes was completed on the IBM 360/67 and the IDIMS system. Grouping is a process whereby spectral classes are generalized to desired resource categories. This generalization of the data maintains reliability of class assignments, however, it effectively eliminates more detailed information known about each spectral class. For instance, the spectral classes which make up each grouped or generalized land cover category represent variations in crown density, slope, aspect etc,. However, because these individual classes could not consistently be identified throughout the scene (due to lack of adequate photographic coverage), were grouped into a more reliable and consistently identified category. Smoothing of the classified data, completed on the CDC 7600, examined each pixel in comparison to its surrounding eight pixels and reassigned the center pixel based on the dominant land cover of surrounding pixels. Smoothing enhances the classified data visually by removing the random, isolated occurrence of single pixels and thus generalizing land cover patterns. Smoothing is advised when output products are requested at a fairly small scale (1:250,000).

Output Products

Final output products for the Tanana project include color coded maps, tabular acreage summaries by quadrangle, and a computer tape of the final data sets. Negative and positive transparencies were generated using the DICOMED film recorder. To fulfill DNR requirements, each of the three scenes were reproduced at a scale of 1:250,000. For general distribution, these same data sets were enlarged to 1:500,000. Additional color coded prints were generated on the DUNN film recorder for specific areas within the study area. An example of a DUNN print of a subsection of the eastern scene near Delta Junction is shown in Figure 6. Notice the extent of agricultural fields in this data set (acquired in 1976). Tabulations by land cover category, acreage and percent of total, were generated for each quadrangle encompassed within the data (Table 1). In many cases, the acreage summary includes only a portion of the quad. Computer tapes of the final classification will also be provided to DNR for integration into the ALARS geo-based information system.

Cost Approximation

Analysis of Landsat digital data provides reliable resource information over large areas in a cost effective manner. Computer charges by task, labor and materials have been compiled for the land cover classification of the Tanana Basin (see Table 2). This section should provide baseline information for those who are considering the use of Landsat data for generating resource information.

A number of computer systems utilized in the analysis are subsidized by Ames Research Center and not charged by task. For instance, the Illiac IV has no equivalent commercial system or software. Therefore, the charge for a specific processing

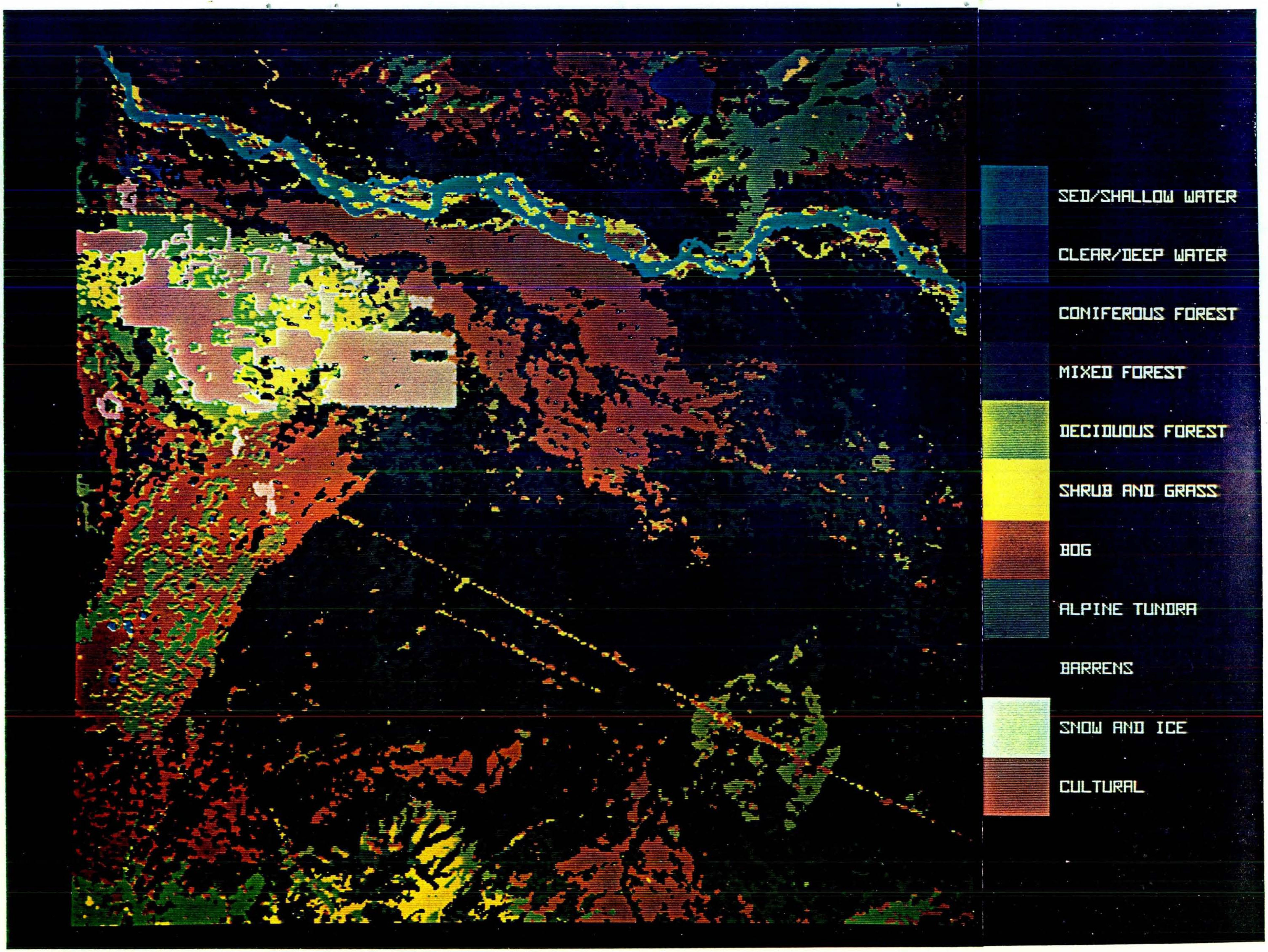


Figure 6 Subsection of the Eastern Scene near Delta Junction

U.S.G.S Quadrangles (1:250,000)

Land Cover	Livengood		Tanana		Big Delta		Fairbanks		Kantishna River		Mt. Hayes	
	Acres	Percent	Acres	Percent	Acres	Percent	Acres	Percent	Acres	Percent	Acres	Percent
Sedimented/ Shallow Water	46819	1.2	63750	2.2	26902	.7	48186	1.0	14789	.6	17824	.8
Clear Water	33561	.8	13030	.5	7845	.2	20299	.4	15450	.7	13696	.6
Coniferous Forest	890185	23.0	700347	23.7	1268031	32.8	1093668	23.5	502018	20.7	491733	21.3
Mixed Forest	142580	3.8	36508	1.2	443945	11.5	176437	4.0	65710	2.7	155196	6.8
Deciduous Forest	411770	11.0	450313	15.2	622660	16.1	276045	6.0	212684	8.8	210391	9.1
Shrub and Grass	890398	23.0	537435	18.2	323206	8.4	1307474	28.0	502103	20.7	473384	20.5
Bog	1395593	37.2	1073896	36.3	948726	24.4	1463795	31.6	977911	40.4	236860	10.3
Alpine Tundra	1392	.03	1462	.04	115757	3.0	140305	3.0	26891	1.1	91647	4.0
Barrens	29200	.6	46017	1.6	98519	2.5	97064	2.0	90300	3.7	503327	21.8
Snow and Ice	2864	.07	28064	.9	6025	.1	8249	.2	14475	.6	107349	4.7
Cultural	10411	.3	6760	.2	10650	.3	13508	.3	0	0	2747	.1
TOTAL	3854773	100%	2958382	100%	3872266	100%	4645030	100%	2422331	100%	2304154	100%
Portion of Scenes	CENTRAL		CENTRAL & WESTERN		EASTERN		EASTERN & CENTRAL		WESTERN & CENTRAL		EASTERN	

TABLE 1a ACREAGE SUMMARIES BY U.S.G.S QUADRANGLES (1:250,000)

LAND COVER	CENTRAL SCENE		EASTERN SCENE		WESTERN SCENE	
	ACRES	PERCENT	ACRES	PERCENT	ACRES	PERCENT
Sedimented/ Shallow water	72177	.8	48320	.8	2395	.1
Clear Water	46124	.5	22645	.4	52431	2.4
Coniferous Forest	2166014	24.3	972891	15.5	18647	.8
Mixed Forest	294510	3.4	609049	9.7	399723	17.6
Deciduous Forest	706141	7.9	929972	14.8	2673	.01
Shrub and Grass	2341368	26.4	1072467	17.0	416053	18.4
Bog	2844590	32.0	1404436	22.4	297944	13.1
Alpine Tundra	120841	1.4	317649	5.0	1031876	45.6
Barrens	248622	2.8	762026	12.1	29220	1.3
Snow and Ice	37285	.4	131974	2.1	16212	.7
Cultural	13508	.1	13414	.2	0	0
TOTAL	8891183	100%	6284343	100%	2267174	100%

TABLE 1b ACREAGE SUMMARIES BY INDIVIDUAL SCENE

task completed on the Illiac was equivalent to a commercially available system. The IDIMS system is also supported through NASA overhead. Charges for IDIMS were computed by connect time ($100/hour) at a commercially available rate. Charges to the CDC 7600, which is also subsidized by NASA, were multiplied by a factor of three.

Many of the costs associated with a demonstration project would not necessarily be duplicated in an operational mode. They do represent the charges incurred by NASA in conducting the project. For example, they include the necessity of maintaining an in-state coordinator and travel costs associated with conducting the project at an analysis center outside the state of Alaska. Cost figures were not included for agency training, agency personnel ground data collection, or use of the Illiac IV. Based on this model, the cost of classifying portions of three Landsat scenes, 79,000 square kilometers (30,620 square miles), was approximately $1.96 per square kilometer ($5.06 per square mile). The unsupervised classification technique is useful for mapping land cover information over large areas where multispectral Landsat data are available and other source materials are not. Although a rigorous evaluation of the results has not been com pleted, a comparison of the color coded maps and aerial photographs indicates the classification is a reliable representation of land cover information.

Table 2

Approximate cost of land cover classification of the Tanana Basin (portions of three scenes)

Category	Cost
Data Acquisition (CCTs, aerial photos)	7,000
Preprocess	1,900
Cluster and Classify	1,850
Evaluation	5,200
Post-Processing	7,800
Output Products	6,000
Staff Support (person/years)	50,000
Photo Products	5,000
Travel	10,000
IPA (in-state coordinator)	30,000
Project Coordination	20,000
NASA Overhead	10,000
	$ 154,750

CHAPTER IV

AGENCY EVALUATION

Completion and generation of final products for the Tanana unsupervised classifications await evaluation by the participating agencies. As the most important phase of the project, evaluation of the final products is the sole realm of the agencies. Factors which should be considered in the assessment process can include:

1. Did the project address a real resource management need?

2. Ease of incorporating results (color coded maps) or procedures into existing agency operations.

3. Relate usefulness of the data products produced to data needs of your agency.

4. Discuss and compare alternate ways that the data products might have been produced.

5. Discuss and compare the data results with other available data sources (Spetzman e.g.).

6. Describe differences and discrepancies between the project results and the planned results (products).

7. Describe your experience in team activities (workshops) with participants from other agencies.

8. Assess the final products (color coded maps and acreage summaries) in terms of:

    Information Content
    Format
    Geographic Coverage
    Scale
    Resolution
    Accuracy

9. What is the long term potential for operational application in your agency?

APPENDIX A

CIR PHOTOGRAPHY FOR TANANA

YEAR/FLIGHT #	SCALE	FRAMES	QUADS	AGENCY
78112	1:60,000	5684-5692	Livengood	Forestry
78112	1:60,000	5664-5672	Livengood	"
78112	1:60,000	5609-5618	Livengood/Tanana	"
79113	1:60,000	3821-3833	Livengood	"
78112	1:60,000	5580-5590	Livengood/Tanana	"
79113	1:60,000	3887-3899	Livengood	"
78112	1:60,000	5513-5525	Livengood/Tanana	"
79113	1:60,000	3808-3820	Livengood	"
78112	1:60,000	5492-5507	Fairbanks	"
79113	1:60,000	3900-3912	Livengood	"
79113	1:60,000	3726-3754	Fairbanks/Kantishna	Land
79109	1:60,000	3415-3421	Fairbanks	"
79119	1:60,000	4954-4959	Kantishna	LRP
	1:60,000	3422		
74104	1:120,000	9795-9797	Healy	Land
77088	1:60,000	1769-1775	Fairbanks	"
74100	1:120,000	9538-9544	Fairbanks	LRP
79113	1:60,000	3942-3964	Fairbanks	"
78112	1:60,000	5784-5791	Fairbanks	DEC
79113	1:60,000	3973-3975	Fairbanks	"
78125	1:60,000	8365-8380	Big Delta	"
78125	1:60,000	8351-8364	Mt. Hayes	F & G
79113	1:60,000	3976-3981	Big Delta/Fairbanks	DEC
79113	1:60,000	3982-3993	Big Delta/Fairbanks	"
77088	1:60,000	1940-1948	Mt. Hayes	F & G

APPENDIX B

PROCEDURE FOR ESTABLISHING TRAINING SITES

PREFIELD STEPS:

1. Obtain stereo model prints (sets of 3 per location) of aerial photography to be used in the field with mylar overlay.

2. Scotch tape phot and mylar to sleeve. On mylar overlay indicate north arrow, frame number and fiducial marks.

3. Locate a number of homogenious training sites for each landcover type (or spectral signature) on the center frome of the stereo prints on aerial photography. Sites should be distributed over entire test region.

   a. Locate sites in relation to prominent surface features which can be found in the field (roadwayus, intersections, creeks, power lines, railroads tracks, lakes, houses, etc.).

   b. Training sites should be at least 10 acres in size, preferable 40 acres. Avoid linearsites; at least 5 acres in width (40 acre field is roughle ¼" x ¼: 1:62,500 photo).

   c. Check for road access to sites. Will four-wheel drive be required?

   d. Concentrate training sites for maximum use of field time.

4. Outline each homogeneous training site on mylar over aerial photography. Assign a field number along side of boundary.

5. Transfer training site boundaries to topographic maps.

6. Delineate training site locations on maps and determine best route. (With an auto approximately 10 sites can be visited per day.)

FIELD WORK:

1. For each honogeneous training site outlined & referenced on an aerial photograph, fill out a ground truth data form from information readily available on topo maps and aerial photographs.

2. If on field examination, area selected does not meet criteria in step 2b, reject area and select another.

3. Horizontal and vertical ground photos will provide a permanent record of each site.

APPENDIX C

TRAINING SITE DATA SHEET

GENERAL

Observer ____________________ Date ____________________

Quadrangle ____________ Scale __________ Time ___________

Locality Description Township______ Range _______

Section _____ _______ ____

Latitude/Longitude_________

Roll _______

Photography Flight Line & Frame No. _________ BLM Code _____

Photo Date _______ Film Type _______ Test Site No. _______

Road Access to Site____________Aircraft Access to Site_____

Other _____________

PHYSIOGRAPHY (See attached legend)

Elevation___________Slope__________Aspect N S E W ____

Position on Slope (Toe, Mid, Upper, Ridge)__________________

Macrorelief________________ Landform ___________________

VEGETATION (See attached legend)

Level I Forest__ Woodland___ Tundra__ Grassland__ Shrubland__Aquatic__ Wetland__

Level II ______________ Level III _______________

Community Type (Dominant species in order of prominance, phenology)

Overstory _____________ 1. ______ 2. ______ 3. ______

Intermediate _____________ 1. ______ 2. ______ 3. ______

Ground Layer _____________ 1. ______ 2. ______ 3. ______

Cumulative Vegetative Ground Cover (enter number of crown cover classes after each species listed above)

1.	95 - 100%	4.	25 - 49%
2.	75 - 94%	5.	10 - 25%
3.	50 - 74%	6.	- 10%

Percentage Cover

____Coniferous Trees
____Deciduous Trees
____Tall Shrubs (2 m)
____Medium ht Shrubs (.5-2 m)
____Dwarf Shrubs (10-50 cm)
____Prostrate Shrubs (10 cm)
____Grass & Sedges
____Forbs
____Mosses & Lichens
____Bare soil
____Rock
____Water

LAND USE ______________________ (See attached legend)

Soils: (optional) Sand ___Silt___Clay___Gravel___Loam___Stones____

Bedrock ___

Moisture: Dry _______ Moist ________ Saturated _______

COMMENTS

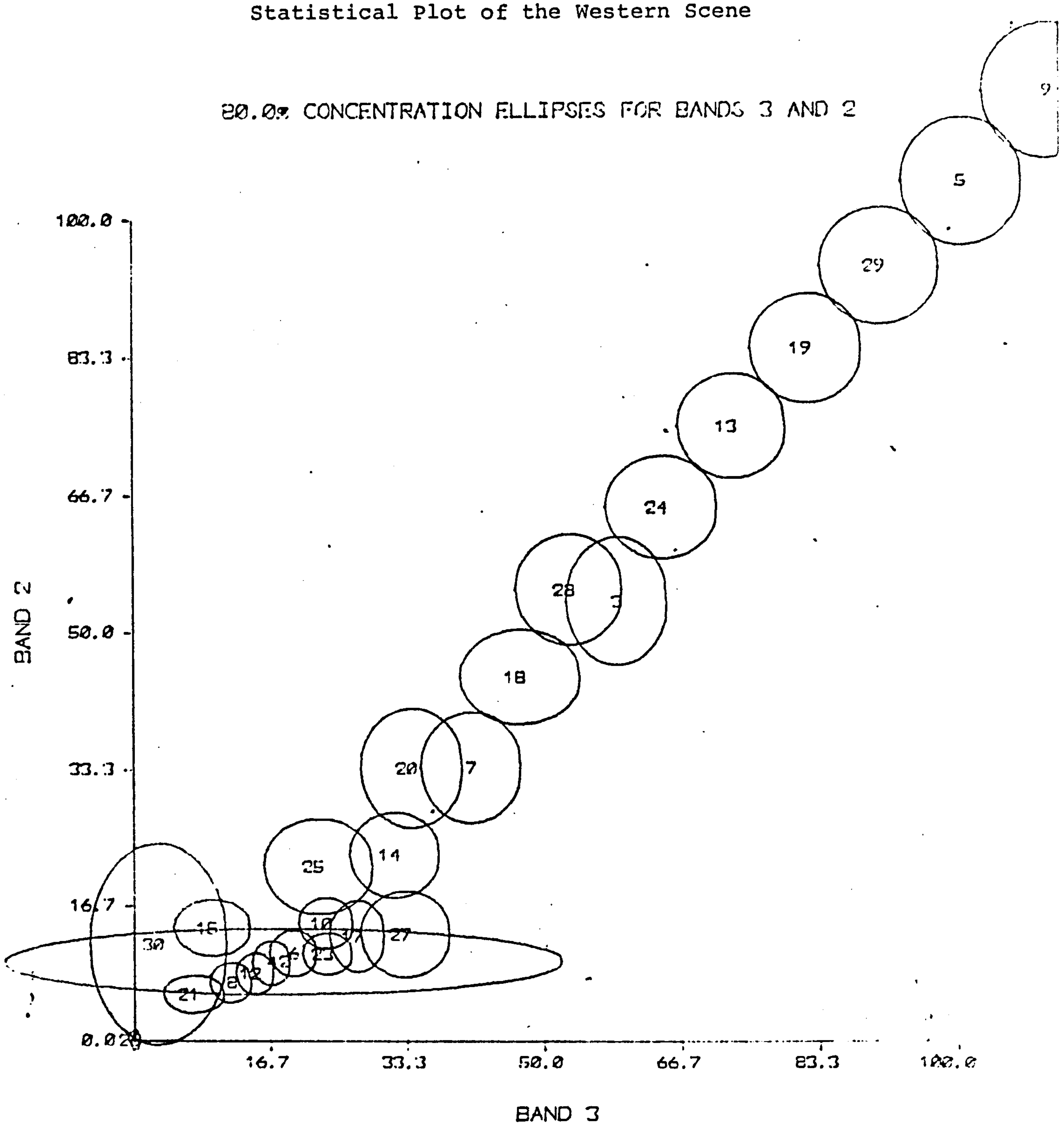
____

____

____

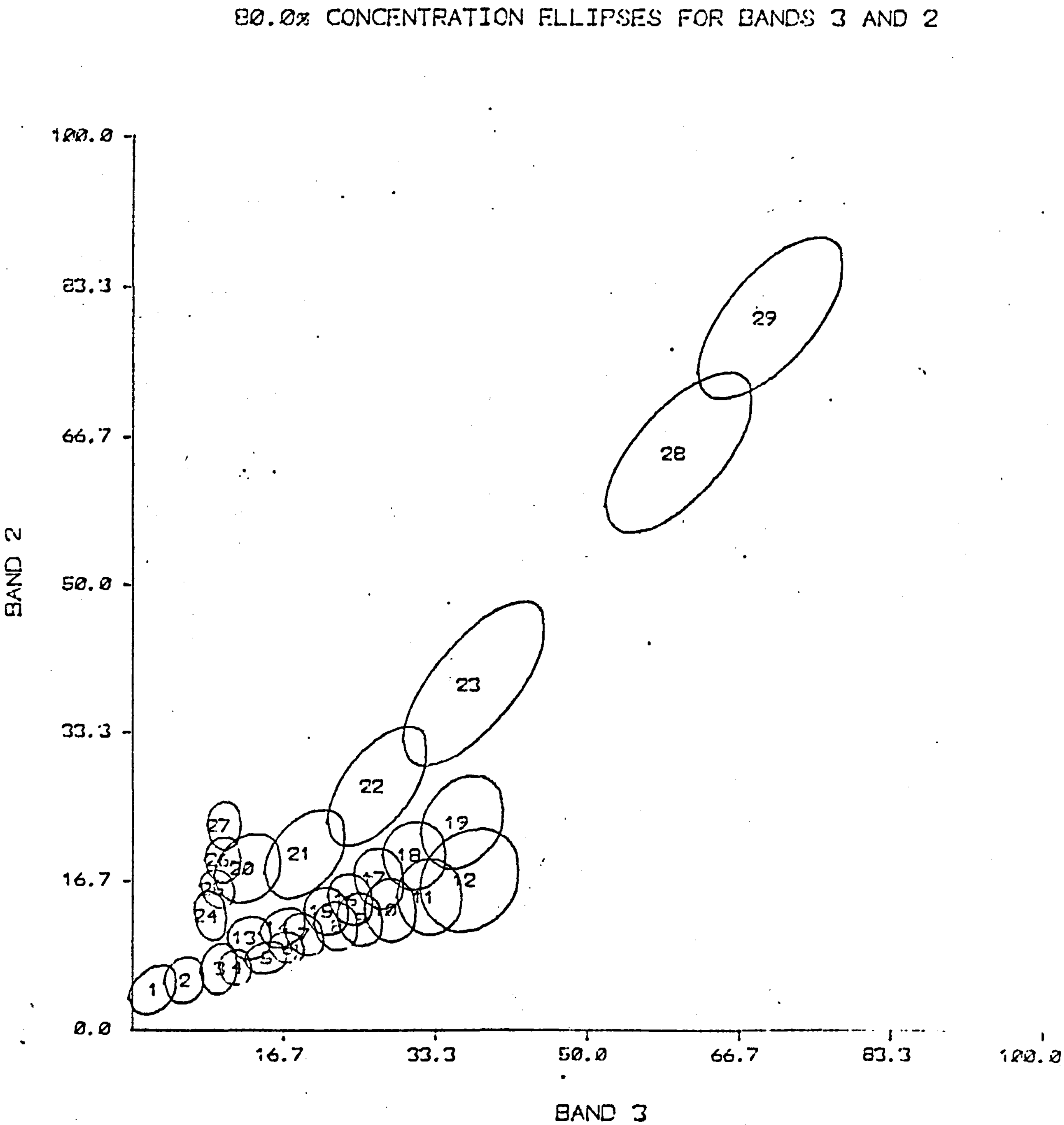
APPENDIX D

Statistical Plot of the Western Scene



APPENDIX E

Statistical Plot of the Central Scene



1. Report No. NASA CR 166393	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Analysis of the Tanana River Basin using Landsat Data		5. Report Date April 1981 6. Performing Organization Code
7. Author(s) Leslie A. Morrissey, Vincent G. Ambrosia and Charlotte Carson-Henry		8. Performing Organization Report No.
9. Performing Organization Name and Address Technicolor Government Services NASA Ames Research Center Moffett Field, California 94035		10. Work Unit No. T 5037 11. Contract or Grant No. NAS2-11101
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C.		13. Type of Report and Period Covered Contractor Report 14. Sponsoring Agency Code 658-80-01

15. Supplementary Notes

Point of Contact: Susan Norman MS 242-4
NASA Ames Research Center
Moffett Field, CA 94035 FTS 448-5912 COM 965-5912

16. Abstract

The Tanana River Basin project, as part of NASAs Applications System Verification and Transfer (ASVT) program, was a cooperative effort with agencies in the state of Alaska to assess the utility of remote sensing technology in meeting critical resource management needs while providing the Alaska Department of Natural Resources (ADNR) with baseline resource information for state land selection and disposal programs.

Digital image classification techniques were used to classify land cover information in the Tanana River Basin. Portions of four scenes of Landsat digital data were analyzed using computer systems at Ames Research Center. An unsupervised clustering approach was utilized to derive cluster statistics for classification of the MSS data. Classification errors, identified during two workshops with ADNR personnel using color infrared photography, were corrected using stratification procedures. Color coded maps and acreage summaries of the major land cover classes were generated for selected USGS quadrangles which lie within the basin. This project, initiated in October 1980, was completed within six months.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Landsat Land cover Alaska	Unclassified - unlimited Subject Category 43

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	43	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

End of Document