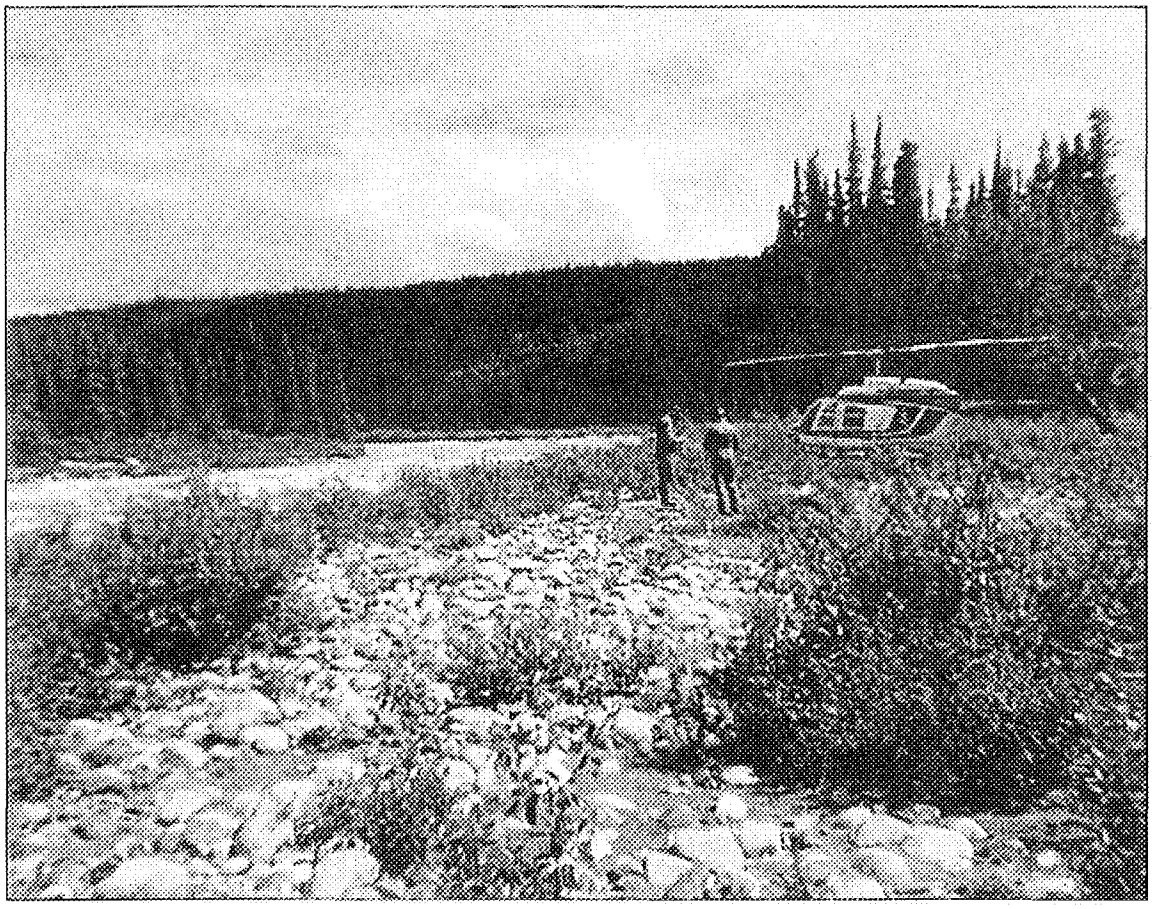
U.S. Department of the Interior
Bureau of Land Management

BLM-Alaska Technical Report 45
BLM/AK/ST-02/018+6500+931
September 2002

U.S. Department of the Air Force

Ducks Unlimited, Inc.

Tanana Flats Earth Cover Classification



Mission Statement

The Bureau of Land Management (BLM) sustains the health, diversity and productivity of the public lands for the use and enjoyment of present and future generations.

Partners

The Department of the Interior, Bureau of Land Management, and Ducks Unlimited, Inc. completed this project under a cooperative agreement.

Cover

The cover photo depicts the remoteness of the area and the need to use helicopters for data collection.

Technical Reports

Technical Reports issued by the Bureau of Land Management-Alaska present the results of research, studies, investigations, literature searches, testing, or similar endeavors on a variety of scientific and technical subjects. The results presented are final, or are a summation and analysis of data at an intermediate point in a long-term research project, and have received objective review by peers in the author's field.

The reports are available while supplies last from BLM External Affairs, 222 West 7th Avenue, #13, Anchorage, Alaska 99513 (907) 271-3318 and from the Juneau Minerals Information Center, 100 Savikko Road, Mayflower Island, Douglas, AK 99824, (907) 364-1553. Copies are also available for inspection at the Alaska Resource Library and Information Service (Anchorage), the United States Department of the Interior Resources Library in Washington, D. C., various libraries of the University of Alaska, the BLM National Business Center Library (Denver) and other selected locations.

A complete bibliography of all BLM-Alaska scientific reports can be found on the Internet at:
http://www.ak.blm.gov/affairs/sci_rpts.html

Related publications are also listed at:
http://juneau.ak.blm.gov.

Tanana Flats
Earth Cover Classification

Technical Report 45
July 2002

U. S. Department of the Interior
Bureau of Land Management
Alaska State Office
222 W. 7th Ave., #13
Anchorage, AK 99513

U.S. Department of the Air Force
611CES/CEVP
10471 20th Street
STE 320
Elmendorf AFB
Anchorage, AK 99506-2200

Ducks Unlimited, Inc.
3074 Gold Canal Drive
Rancho Cordova,CA 95670

Acknowledgements

This project was funded under a cooperative agreement between the United States Department of the Interior, Bureau of Land Management (BLM), Pacific Meridian Resources, Utah State University (USU) and Ducks Unlimited, Inc. (DU). Other cooperators include United States Air Force, National Biological Service, United States Fish and Wildlife Service, Alaska Department of Fish and Game and Alaska Department of Natural Resources. The project was administered by John Payne (BLM) and Dick Kempka (DU).

Special thanks are extended to John Payne, Dick Kempka, Jim Sisk (BLM, Fairbanks), Jim Herriges (BLM, Fairbanks), Craig Altop (BLM, Utah) and Colin Homer (USU) for field assistance, map preparation, field site selection and organizing logistical details. Thanks to Air Logistics for safe air service. Thanks also to Bob McAlpin from Alaska Fire Service for help in coordinating helicopter contracting; and Mark Pearson (GeoNorth, Inc.) for programming the Ducks Unlimited Field Form (DUFF) database.

In addition, we would like to acknowledge the efforts of those who have technically reviewed this report. John Payne and Ed Bovy offered important suggestions.

Table of Contents

Acknowledgements iii

Table of Contents v

List of Figures vii

List of Tables ix

Abstract xi

Introduction 1

- Project Objective 2
- Project Area 2
- Data Acquisition 2

Methods 5

- Image Preprocessing 5
- Classification Scheme 5
- Field Verification 6
- Field Data Analysis 8
- Classification 9
  - Generation of New Bands 9
  - Seeding Process 9
  - Band Selection 9
  - Modified Supervised/Unsupervised Classification 11
- Accuracy Assessment 12

Results 15

- Field Verification 15
- Classification 15
- Accuracy Assessment 20
- Final Products 23

Conclusions 25

Literature Cited 27

Appendices 29

- Appendix A. The decision tree for the Tanana Flats project 29
- Appendix B. Error matrices for the Tanana Flats classification 31
- Appendix C. Contact information 33

List of Figures

Figure 1. The Military Operational Area mitigation effectiveness study - habitat phase project area. 3

Figure 2. Data field collection card for the Tanana Flats project..... 7

Figure 3. Helicopters are the most effective way to complete the field work for this project 6

Figure 4. Standardizing sampling techniques ensures consistent results..... 8

Figure 5. The classification process for the Tanana Flats project..... 10

Figure 6. The digital elevation model coverage of the Tanana Flats project area. 17

Figure 7. The 1000 meter contour for the Tanana Flats project 18

Figure 8. The 100 meter contour intervals from 1000 meters to 1500 Meters. 19

Figure 9. The final classification for the Tanana Flats project..... 21

List of Tables

Table 1. The classification scheme 6

Table 2. The number of field sites collected and the number of accuracy assessment sites withheld 16

Table 3. The acreage statistics for each landcover category within the study area 20

Table 4. The number of photo interpreted sites used in the accuracy assessment 22

Abstract

The Department of the Interior has established the Research and Monitoring Committee to identify research, monitoring, and inventories that must be accomplished to establish short and long-term mitigation for Military Operational Areas (MOAs) in Alaska. The establishment of base-level landcover information is key to any terrestrial research project. Therefore, the United States Air Force has contracted with the Bureau of Land Management (BLM) to map approximately six million acres within designated MOAs in Alaska, including important waterfowl habitats. This data will ultimately be used to assess satellite imagery for mitigating the potential impacts of military overflights on the Delta caribou herd. The cooperators in this project include BLM-Alaska, Ducks Unlimited, Inc., and Pacific Meridian Resources.

One Landsat Thematic Mapper (TM) satellite scene (Path 68/Row 15, shifted 23% south, acquired July 14, 1993) was used to classify the project area into landcover categories. Fifty-five, 1:63,360 scale, quadrangle color infrared plots of the Landsat TM data were produced for the placement of field sample sites and for navigational purposes with the helicopters. A custom field data collection card was used to record field information. After initial on-the-ground sampling, a helicopter was utilized to gain access to field sites throughout the project area. Global positioning system technology was used to record locations of sites sampled in the field. Data was collected on 315 field sites during a five-day span in late August 1995. A portion (30%) of these field sites were set aside for accuracy assessment.

A modified supervised/unsupervised technique was used to classify the satellite imagery. Digital elevation models were used to stratify the project area into the alpine tundra class and to alleviate most of the shadow effects from mountains. The results of this classification indicate low shrub is the dominant landcover in this region accounting for 20.2% of the project area. The percentages for the rest of the landcover types are: other (clouds, cloud shadows, etc.) 15.4%, open needleleaf forest 14.9%, closed mixed forest 6.3%, closed deciduous forest 5.7%, snow 5.0%, closed needleleaf forest 4.9%, tall shrub 4.2%, dwarf shrub 3.8%, graminoid 3.7%, alpine tundra shrub 2.8%, alpine tundra graminoid 2.5%, sparse vegetation 2.0%, bryoid/lichen 1.8%, open deciduous forest 1.6%, agriculture 1.3%, ice 1.2%, clear water 1.0%, turbid water 0.9%, wet herbaceous 0.4%, urban 0.3%, and aquatic bed 0.1%. The overall accuracy for the classification was 82%. This geographic information system database will be used to model the impacts of low-altitude jet aircraft flyovers on wildlife within the MOAs of Alaska's interior.

Introduction

The Institute of Arctic Biology (Institute) at the University of Alaska, Fairbanks, has begun research to determine the effects of low-altitude jet aircraft on the Delta caribou herd. The study predicts that in response to overflights, female caribou with calves may be moving into lower quality habitats, in terms of both foraging and predator avoidance. To plot caribou and other wildlife species movement in relation to habitat, and response to noise stimulation, the Institute will require high quality, high resolution earth and digital elevation mapping products.

Meanwhile, the United States Air Force (Air Force) is preparing an environmental impact statement (EIS) for its Military Operational Areas (MOAs) in Alaska. The Department of the Interior (DOI) and the State of Alaska have agreed to coordinate the studies and research to determine the effectiveness of mitigation for all agencies. Cooperators (United States Air Force, National Biological Service, United States Fish and Wildlife Service, Bureau of Land Management, Alaska Department of Fish and Game, and Alaska Department of Natural Resources) will prepare combined proposals from all agencies so overlap and duplication of information is minimized. In other words, the Air Force can go forward with studies knowing that all agencies have agreed with the proposals.

To coordinate these efforts, the Research and Monitoring Committee has been established by DOI having representatives from the United States Air Force, National Biological Service, United States Fish and Wildlife Service, Bureau of Land Management, Alaska Department of Fish and Game and Alaska Department of Natural Resources. This committee will make recommendations to the Air Force on types, duration, and content of biological studies and research that will be required to adequately address potential impacts from low-flying military aircraft.

Effective mitigation efforts must be defined in terms of geographic placement and levels of disturbance. Thus, studies and research must be designed to result in geographically specific recommendations on the efficiency of mitigating wildlife habitat. In developing modern models that address such things as overflights and their potential effects on wildlife, Alaska is unique in the availability of this type of information. Either information is generally lacking, or is of such age and scale that it is not useful. In addition, most existing data has limited functionality and resides in numerous disjunct formats that meet no national mapping standards or recognized classification structure. Future studies and research for recreation, wildlife, and subsistence issues must be tied directly to this type of ground based information to have landscape and regional meaning.

Most ground-based inventories of vegetation in Alaska have been limited by accessibility to the area, or logistically restricted to a single large or several smaller watersheds. Aerial photography is available for much of Alaska, but is highly variable in scale, and unless it is color infrared, it has limited uses for determining wildlife habitat types. In the last two decades, space-borne remote sensors (Landsat, *Systeme Pour l'Observation de la Terre*, Synthetic

Aperture Radar, and others) have emerged as the best methodologies for developing regional earth cover databases. Access to these large databases allows biologists and managers to define and map crucial areas for wildlife, perform analysis of related habitats, plot movement patterns for large ungulates, generate risk assessments and provide a base to which wildlife and sociological data can be related.

The establishment of base-level landcover information is key to any terrestrial research project. Several project proposals, received by the DOI committee during the fiscal year 1995, had satellite landcover information as the base for species-specific studies. This project, Military Operational Area mitigation effectiveness study - habitat phase, was selected, in part, because it offered an opportunity to establish a landcover database using consistent methodology and national standards. In addition, the products of this project can be utilized with the Institute of Arctic Biology's Delta caribou herd program as well as future research, mitigation and management efforts.

Project Objective

The objective of this project is to develop a baseline landcover inventory using Landsat Thematic Mapper (TM) satellite imagery in Geographic Information System (GIS) format. The landcover data will then be used with Digital Elevation Models (DEM) to assist in determining mitigation for potential impacts of military overflights on the Delta caribou herd. More specifically, this project will purchase, classify, field verify, and produce high quality, high resolution digital and hard copy resource base maps that can be used in the continuing analysis of the effectiveness of mitigation within the MOA's.

Project Area

The project area includes the area southeast of Fairbanks and extends past the Alaska range. It includes the Eielson, Birch, Falcon, and parts of the Yukon1, Buffalo, and Fox MOAs (Figure 1). The project area is located in the interior physiographic province and includes a large portion of the Delta caribou herd traditional use area. The area is underlaid by discontinuous permafrost. The predominant vegetation in this region is deciduous birch and willow in the river valleys and aspen at higher elevations.

Data Acquisition

Commercially available Landsat TM, with a spatial resolution of 30 meters, is available for most of the MOAs. Each Landsat TM scene covers approximately eight million acres. For this project, one terrain corrected Landsat TM satellite image (Path 68/Row 15, shifted 23% south, acquired July 14, 1993) was used for the landcover classification. This is a terrain-corrected image purchased from the Earth Observation Satellite Corporation (EOSAT). In addition, data were collected at 315 field sites from August 21 through August 25, 1995. The ancillary data includes: National Aeronautics and Space Administration (NASA) 1:60,000 aerial photographs and 45 edgematched 1:63,360 scale DEM that cover the majority of the project area, purchased from the United States Geological Survey (USGS). The DEMs will provide information for the analysis of slope, aspect, view sheds, noise sheds, and drainage basin validations to support potential effects on wildlife research and studies.

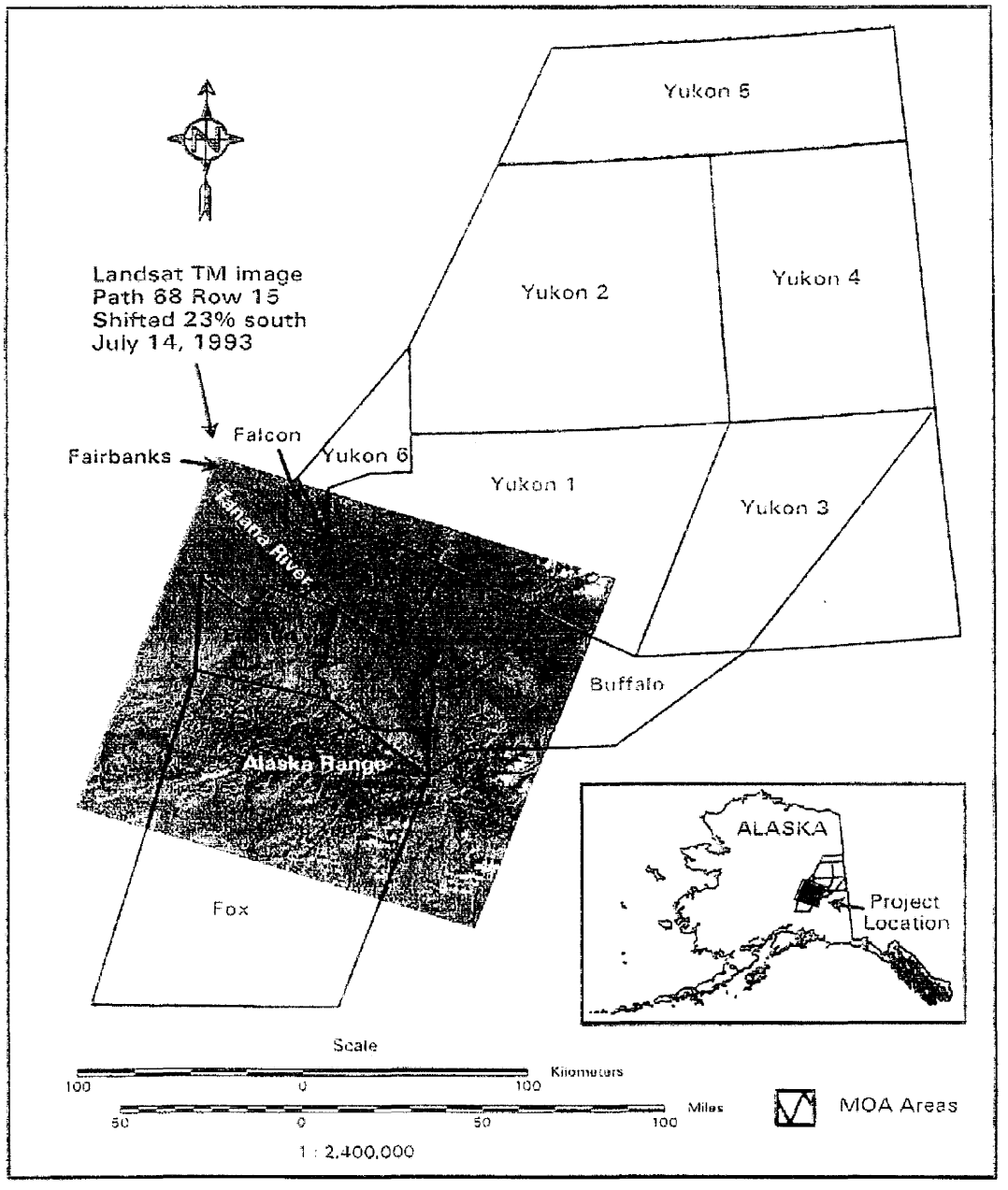


Figure 1. The Military Operational Area mitigation effectiveness study - habitat phase project area

Methods

Image Preprocessing

Under normal circumstances, an unsupervised clustering and seeding technique is used to initially generate spectrally unique areas within the study area (Kempka *et al.* 1995). These spectrally unique areas would then be refined and selected as sample sites for the field work, using aerial photography and a preliminary decision tree of the landcover classification. These sample sites would then be plotted on field maps for field verification, and the coordinates of their center points would be programmed into the aircraft global positioning systems (GPS). However, due to the schedule of this project, there was not adequate time before the field season to preselect the field sites. In fact, the Landsat TM scene used in this project was still being processed by EOSAT when the field data was being collected. The field maps were produced from two existing scenes. The southern area of the project was covered by an existing Ducks Unlimited (DU)/Bureau of Land Management (BLM) Landsat TM scene (Path 68 Row 16, August 23, 1987) and the northern area was covered by an existing Alaska State Department of Natural Rresources (DNR) Landsat TM scene that was the same image used in this project except that it was not shifted. Therefore, 55 - 1:63,360 scale quadrangle color infrared plots of the Landsat TM data were produced for the placement of field sample sites and for navigational purposes with the helicopters. In addition to the color infrared plots, a 30 class unsupervised classification was performed and plotted for each quadrangle to aid the navigator in selecting unique and homogenous sites.

Classification Scheme

The first step in any mapping project is the definition of a classification system, which categorizes the features of the earth to be mapped. Specifications of the system are driven by (1) the anticipated uses of the map information; and (2) the features of the earth that can be discerned with the data (e.g., aerial photography field or satellite imagery) to create the map. A classification system has two critical components: (1) a set of labels (e.g., forest, shrub, water); and (2) a set of rules, or a system for assigning labels. It is important that the set of rules or the system for assigning labels be both mutually exclusive and totally exhaustive. In other words, any area to be classified should fall into one and only one category or class. In addition, every area should be included in the classification.

The classification scheme for the landcover inventory was based on The Alaskan Vegetation Classification by Viereck *et al.* 1992. This classification scheme was chosen because of its acceptance in the scientific community and because it is specific to Alaska. Through a series of meetings with biologists familiar with the vegetation and previous work with Landsat TM classifications in Alaska (Kempka *et al.* 1995, Kempka *et al.* 1993), the Viereck classification was adapted to reflect the landcover classes that would be useful to district biologists and could indeed be mapped using satellite imagery. The classification scheme consisted of 10 major categories and 25 subcategories (Table 1). A classification decision tree was developed in order to eliminate any confusion in the classification (see Appendix A).

Table 1. The Classification Scheme

1.0 Forest	5.0 Wetlands
1.1 Closed Needleleaf	5.1 Wet Herbaceous
1.2 Open Needleleaf	5.2 Aquatic Bed
1.3 Closed Deciduous	6.0 Water
1.4 Open Deciduous	6.1 Snow
1.5 Closed Mixed	6.2 Ice
1.6 Open Mixed	6.3 Clear Water
2.0 Shrub	6.4 Turbid Water
2.1 Tall Shrub	7.0 Barren
2.2 Low Shrub	7.1 Sparsely Vegetated
2.3 Dwarf Shrub	7.2 Rock/Gravel
2.4 Alpine Tundra - Shrub	7.3 Mud/Silt/Sand
3.0 Sedge/Graminoid	8.0 Agriculture
3.1 Sedge Graminoid	9.0 Urban
3.2 Alpine Tundra - Graminoid	10.0 Other

Field Verification

The purpose of field data collection is to assess, measure, and document the on-the-ground vegetation variation within the project area. This variation will then be correlated with the spectral variation in the satellite imagery during the image classification process. In order to obtain a reliable and consistent field sample, a custom field data collection card (Kempka *et al.* 1994) was developed and used to record field information (Figure 2). Two, four-person helicopter crews were designated to perform the field assessment. Each crew consisted of a pilot, biologist, recorder, and navigator. The co-pilot seat was occupied by the navigator who runs the GPS equipment and interprets the satellite image derived field map. The biologist was the person most knowledgeable regarding the vegetation, and the recorder would verify vegetation composition and record landcover type percentages and other data on the field form. Low-level helicopter surveys are a very effective method of field data collection, since a much broader area can be covered with a view from above (similar to a satellite sensor), and for use in remote areas that are difficult to access (Figure 3).

Figure 3. Helicopters are the most effective way to complete the field work for this project.

Rev 8/18/95

Tanana Flats Field Form

Site Number ____ Soil Type ____	Observer ____	Obs. Date ___/___/___ (Mo Day Year)	1 2 3 4 Obs. Level	%Slope (Avg) ____	Aspect (Avg) ____ Elev'n ____
Air Photo: ____ (Flight Line) ____ (Photo)	___/___/___ (Photo Date)	Photo Source: ____	Image		1:64K Quad: ____
LAT (GPS) (Uncorrected) ____:____:____.____ (Degrees Min Seconds)		LONG (GPS) (Uncorrected) ____:____:____.____ (Degrees Min Seconds)		GPS/Local Time: ____:____:____.____ (Hr Min Seconds)	

FOREST	FOREST con't	SHRUBS	HERBACEOUS	BARREN	WETLAND	WATER	OTHER
Closed Needleleaf	Open Broadleaf	Tall	Dry Graminoid	Gravel/Rock	Wet herbaceous	Clear	
Open Needleleaf	Broadleaf woodland	Low	Mesic Graminoid	Sand	Aquatic bed	Turbid	
Needleleaf Woodland	Closed Mixed	Dwarf	Bryoid	Mud/Silt		Ice/snow	
Closed Broadleaf	Open Mixed			Other			

NOAA Coastwatch Classification	
Viereck Classification	

100%	300%	Height	TREES	
			Alaska Paper Birch	Betula papyrifera
			Larch, Tamarack	Larix laricina
			White Spruce	Picea glauca
			Black Spruce	Picea mariana
			Balsam Popular	Populus balsamifera
			Black Cottonwood	Populus trichocarpa
			Quaking Aspen	Populus tremuloides

100%	300%	Height	SHRUBS	
			Rose	Rosa acicularis
			Alder	Alnus spp.
			Bearberry	Arctostaphylos spp
			Dwarf Arctic Birch	Betula nana/glandulosa
			Mountain-avens	Dryas spp.
			Highbush cranberry	Viburnum edule
			Labrador-tea	Ledum spp
			Willow	Salix alaxensis
			Willow	Salix pulchra
			Blueberry	Vaccinium spp.

100%	300%	HERBACEOUS	
		GRASS	
		Pendant grass	Arctophila fulva
		Bluejoint	Calamagrostis canadensis
		SEDGE	Carex spp.
		Cottongrass	Eriophorum spp
		RUSH	Juncus spp
		LICHEN	Color
		MOSS	Bryoid spp.
		FERN	
		Subtotal % Cover	

100%	300%	HERBACEOUS con't	
		FORBS	
		Yarrow	Achillea borealis
		Monkshood	Aconitum delphinifolium
		Vetch	Astragalus spp
		Dwarf dogwood	Cornus canadensis
		Shooting star	Dodecatheon spp
		Fireweed	Epilobium angustifolium
		Horsetail	Equisetum spp
		Bistort	Polygonum spp
		Cinquefoil	Potentilla spp
		Wintergreen	Pyrola spp
		Cloudberry	Rubus spp
		Saxifrage	Saxifraga spp

100%	300%	AQUATIC	
		Mare's tail	Hippuris spp
		Buckbean	Menyanthes trifoliata
		Pond lily	Nuphar polysepalum
		Pondweed	Potomegetum spp
		Burreed	Sparganium minimum
		Sphagnum	Sphagnum spp

100%	300%	NON-VEGETATED	
		Clear water	
		Turbid Water	
		Snow/Ice	(circle one)
		Mud/Silt/Soil	(circle one)
		Sand	
		Gravel/Rock	(circle one)
		Subtotal % Cover	
		GRAND TOTAL % COVER	

C#	COMMENTS

Figure 2. Field data collection card for theTanana Flats project.

Initial sampling was performed with both crews on the ground to verify and standardize the classification and sampling techniques (Figure 4). After initial on-the-ground sampling, the majority of the sample sites were observed visually by helicopter to determine the percent cover for each species and an overall landcover class. Ground verification was performed when identification of dominant vegetation and/or species was uncertain.

Figure 4. Standardizing sampling techniques ensures consistent results.

Due to the large project area and the relatively short time for field sampling, the personnel collecting the field data were split into two crews. Each crew was responsible for collecting field data for their portion of the project area. The biologist responsible for calling the vegetation for both crews did initial sampling together on the first day of field work to calibrate their calls. This was done in order to help the biologist make consistent calls between field crews. The crew working out of Fairbanks was responsible for the areas south of Fairbanks to the Alaska range. The crew working out of Paxson was responsible for the area south of the Alaska range and through the Delta River corridor.

In order to more thoroughly describe the field site it was decided to add a 300% column to the field form (Figure 2). The 100% column describes all the vegetation as seen from a vertical view. The 300% column describes the tree canopy, shrub canopy and the ground cover canopy. Theoretically, a site could have 100% tree cover, 100% shrub cover under the tree canopy and 100% ground cover under the tree and shrub canopy, thus 300%. The 300% level would be very beneficial for wildlife modeling and detecting the subtle differences in the spectral variation in the imagery. However, the reliability of this 300% canopy call was suspect from a helicopter. It was very difficult to assess canopy especially in heavily forested or vegetated areas. In addition, the added time spent in determining this subcanopy was very time consuming and not cost effective. Thus, the 300% assessment was limited to only a small portion of the total 315 sites sampled in the field.

Field Data Analysis

The field sites were entered into a customized database developed by Pacific Meridian Resources. Each site was labeled with a class name by running the database through the classification decision tree (Appendix A). Any discrepancies between what the site was called in the field and the calculated call by the computer was reviewed in order to assure a correct classification of the site. The field sites were then stratified by class. Approximately one half of the field sites for each class were randomly selected for accuracy assessment purposes. If a class had less then 30

samples, 15 were used in the classification and the remaining samples were selected for the accuracy assessment.

Each site collected in the field was on-screen digitized into an ArcInfo coverage using the image, field maps and GPS points. The pertinent attributes from the database were then related to the ArcInfo coverage of field sites. A new attribute (Type) was added to the coverage to designate if the site was to be used as a training area or for accuracy assessment. Two separate coverages were created using the Type attribute to separate the training sites from the accuracy assessment sites. The coverage with all the field sites and the coverage with the accuracy assessment sites were stored in separate files. Only the coverage with the training sites were used in the classification process.

Classification

Every image is unique (e.g. spectral and spatial differences resulting from different dates of imagery and/or sensors) and presents its own special problems in the classification process. The approach that is used in this project has been used and proven to be successful over many years (Figure 5). The image processor's site-specific experience and knowledge in combination with high quality ancillary data can overcome some of the spectral differences to produce a high quality and extremely useful product. Therefore, the image processor should be actively involved in field data collection and gain first hand knowledge of every training site.

Generation of New Bands

Two new bands were generated from the raw Landsat TM data. A ratio of band 4 and band 3 was derived by dividing the digital number (DN) of band 4 by the DN of band 3. Many different band ratios could be derived and tested for spectral separability, however, from past experience in Alaska and other vegetation studies, the 4/3 ratio was chosen (Kempka *et al.* 1995, Congalton *et al.* 1993). The 4/3 ratio typically reduces the shadow effects and enhances the differences between vegetation types. The second band that was generated was the first principal component of the visible bands (1,2,3). Normally, the visible bands are highly correlated and therefore redundant data. The first principal component normally accounts for over 90% of the variation in the visible bands. These two new bands were subset with the six raw bands to produce an 8-banded file to be used in selecting the best band combination.

Seeding Process

The field sites that were designated as training areas were seeded (generate statistics from the imagery) in ERDAS Inc., Imagine software using spectral bounds as the limit for seed growth. The standard deviations of the seeded areas were kept to about 3 and all seeded areas were required to be over 15 pixels (approximately 3.75 acres) in size. Along with the field training areas, additional seeds were generated for the water, turbid water, ice, snow and urban areas. These classes were easily recognized on the imagery and aerial photography.

Band Selection

Following the supervised seeding process a spectral pattern analysis was run on the supervised seeds in order to determine optimal band combination. Spectral pattern analysis is a graphical method of comparing the mean digital number of the different classes across the different bands. A

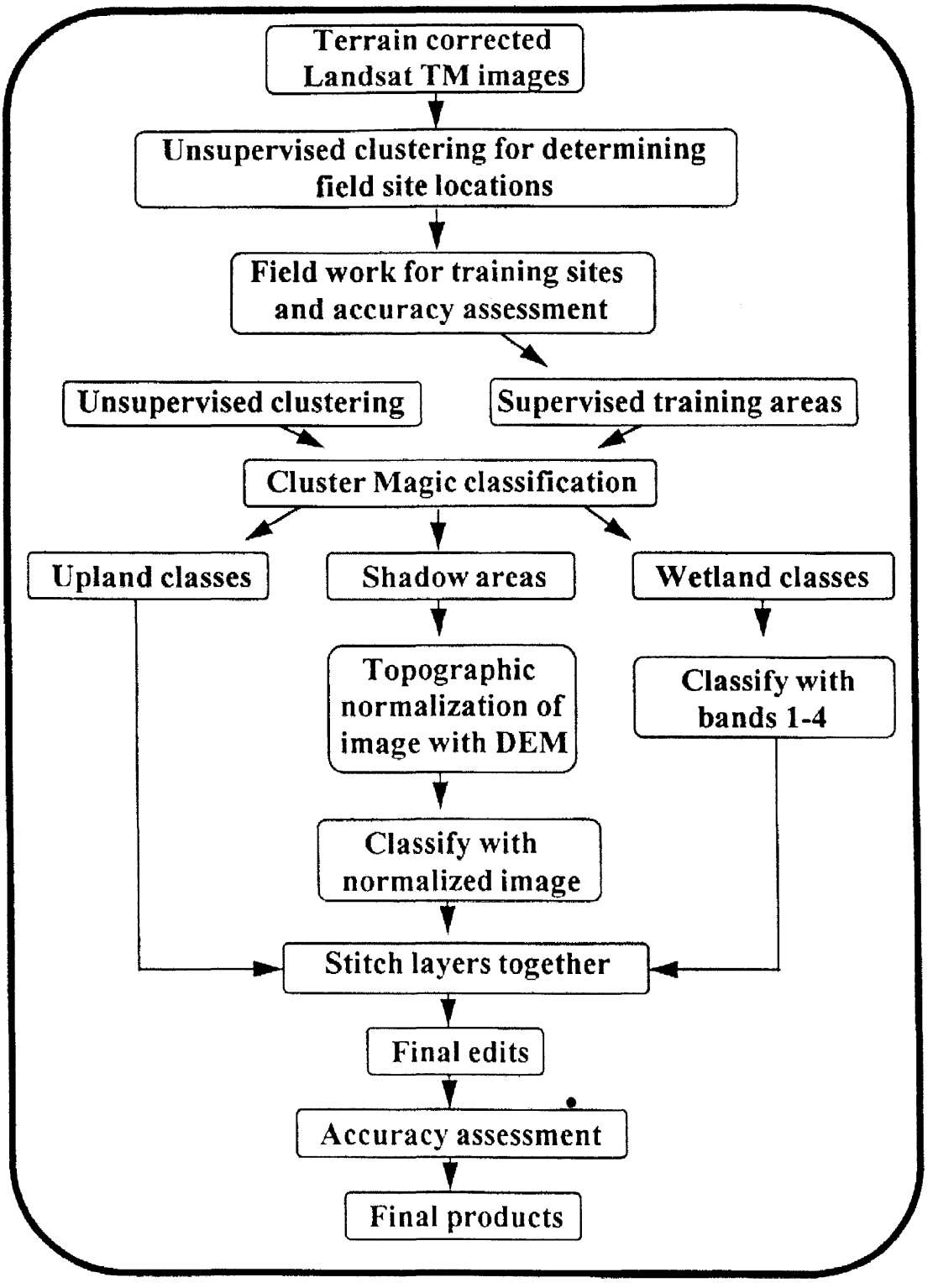


Figure 5. The classification process for the Tanana Flats project

statistical program was run in ERDAS Inc., Imagine (DIVERGE) that calculates the spectral distance between class signatures to determine the bands that have the maximum separability between classes. The procedure to select the optimal bands was performed to determine which bands would give the best separation between the classes and which bands had redundant and/or confusing information. The bands with redundant and/or confusing information were removed during the classification process to reduce the processing time and enhance the classification.

Modified Supervised/Unsupervised Classification

Before the actual classification was begun, the clouds and shadow areas associated with the clouds were removed through on-screen-digitizing. Since no landcover information could be obtained under the clouds or shadows caused by the clouds, these areas were removed from the classification process. The cloud and shadow areas were included as "Other" in the final classification.

A modified supervised/unsupervised classification approach (Chuvieco and Congalton 1988) was used for the classification. This approach uses a statistical program to group the spectrally unique clusters from the unsupervised signatures with the signatures of the supervised training areas. In this way, the spectrally unique areas are labeled according to the supervised training areas. Therefore, 250 unsupervised clusters were generated using a program in ERDAS Inc., Imagine (ISODATA) and grouped with the supervised signatures. The process was repeated until all of the spectral classes were adequately matched and labeled. The classification approach provides three major benefits: (1) it aids in the labeling of the unsupervised classes by grouping them with known supervised training sites; (2) it helps identify classes that posses no spectral uniqueness, (i.e. training sites that are spectrally inseparable); and (3) identifies areas of spectral reflectance present in the imagery that have not been represented by a training site.

After the first iteration, the wetland classes (including water) and the shadow areas (mountain shadows) were separated from the classification (Figure 5). The wetland classes were classified with bands 1-4. Bands 1-4 were used on the wetland classes because of the better penetration of water and separability of wet and highly saturated classes. The shadow areas were also classified separately using a topographic normalized image. The topographic normalized image was derived using the DEM and the solar azimuth and sun angle from the TM image. The resulting normalized image reduces the shadow effects by normalizing the shadow areas to the rest of the image. This enhancement will help to more accurately classify the shadow areas.

After the upland classes, mountain shadows, and wetland classes were classified, they were stitched together. The clouds and cloud shadows were then added to the "Other" class to produce a complete and final classification of the entire scene. The next step was to add the "Alpine Tundra" class. Any "Graminoid" class or "Shrub" class above 1000 meters was recoded to "Alpine Tundra". The 1000-meter delineation was determined from a previous study in the Steese White Mountains by Markon (1993). Using the DEM to determine the 1000-meter contour, the graminoid class was recoded to alpine tundra - graminoid above the 1000 meter

contour and the shrub classes were recoded to alpine tundra - shrub above the 1000 meter contour.

The final step of the classification process was to make final edits. The final edits consisted of: (1) viewing the classification and imagery to make sure there were no glaring errors such as mountain shadow areas being classified as water or urban areas being classified as barren; and (2) spot checks with aerial photography in areas of known confusion.

Accuracy Assessment

The purpose of quantitative accuracy assessments is the identification and measurement of map errors. There are two primary motivations for accuracy assessment: to understand the errors in the map (so they can be corrected), and to provide an overall assessment of the reliability of the map (Gopal and Woodcock, 1992). In order to give their clients a high quality and reliable product, Pacific Meridian Resources strongly advocates performing quantitative accuracy assessments for all mapping projects.

There are many factors to consider when designing an accuracy assessment. These include how to determine the sample size, how to allocate this sample and which sampling scheme to employ. Congalton (1991) suggests that 50 samples be selected for each map category, as a rule of thumb. This value has been empirically derived over many projects. A second method of determining sample size is using the multinomial distribution and specifying a given confidence in the estimate (Tortora 1978). The results of this calculation tend to favorably agree with Congalton's rule of thumb. Once the sample size is determined, it then must be allocated among the categories in the map. A strictly proportional allocation is possible. However, the smaller categories in aerial extent will have only a few samples, which may severely hamper future analysis. The other extreme is to force a given number of samples from each category. Depending on the extent of each category, this approach can significantly bias the results. Finally, a sampling scheme must be selected. A purely random approach has excellent statistical properties, but is practically difficult and expensive to apply. A purely systematic approach is easy to apply, but could result in sampling from only limited areas of the map.

Given the above mentioned constraints, Pacific Meridian Resources uses an accuracy assessment methodology that balances statistical validity with practical applicability for this project. Ideally, approximately 50 samples per class would be selected. However, this was not possible in this project. Instead, for the two to three largest categories, 50 samples per category were taken. For the two to three smallest categories, 15 samples per category were striven for. For the rest of the categories, 30 samples will be taken. This combines the rule of thumb sample sizes with some proportional allocation. The sampling scheme would also combine aspects of systematic and random sampling.

For this project, the time limitations for field work would not allow for adequate field samples to be withheld for the accuracy assessment. However, a complete set of aerial photographs (see Data Acquisition) that cover the entire project area were available. Therefore, random sample points stratified by class were generated using ERDAS Inc., Imagine. The points were stratified by class in order to obtain enough samples in each class. Each point was then

viewed over the raw image in order to locate the associated aerial photograph. The site was then photo interpreted, using a stereoscope, and given a class name.

As an additional quality check and consistency test, Pacific Meridian Resources checked the accuracy of the photo interpretation. This is accomplished by photo interpreting the field sites and comparing the photo interpretation with the field site. This gives the user of the classification additional confidence in the classification and may point out why and where the errors may occur.

The accuracy assessment sites are used to build a matrix that compares the reference data (field site or photo interpreted site) with the classification. The matrix (commonly referred to as an error matrix or confusion matrix) is a square array of numbers set out in rows and columns that express the number of pixels assigned to a particular category in one classification relative to the number of pixels assigned to a particular category in another classification. In most cases, one of the classifications is considered to be correct and may be generated from aerial photography, airborne video, ground observation or ground measurement. The rows usually represent this reference data while the columns indicate the classification generated from the remotely-sensed data. An error matrix is an effective way to represent accuracy in that the individual accuracies of each category are plainly described along with both the errors of inclusion (commission errors) and errors of exclusion (omission errors) present in the classification. A commission error occurs when an area is included into a category where it does not belong. An omission error is excluding that area from the category in which it does belong. Every error is an omission from the correct category and a commission to a wrong category.

In addition to clearly showing errors of omission and commission, the error matrix can be used to compute overall accuracy, producer's accuracy and user's accuracy (Story and Congalton 1986). Overall accuracy is simply the sum to the major diagonal (i.e., the correctly classified pixels or samples) divided by the total number of pixels or samples in the error matrix. This value is the most commonly reported accuracy assessment statistic. Producer's and user's accuracies are ways of representing individual category accuracies instead of just the overall classification accuracy.

In addition, a Cohen's coefficient of agreement (Kappa) analysis was performed on the error matrix as a further measure of accuracy (Congalton 1991). Kappa is a measure of overall agreement in the error matrix after chance agreement is removed from consideration. In other words, Kappa attempts to provide a better measure of agreement by adjusting the overall accuracy for chance agreement or that agreement that might be contributed solely by chance matching of the two maps.

In addition to calculating Kappa, confidence intervals can be calculated using the approximate large sample variance. The large sample variance can then be used to test if the agreement between the classification and reference data is significantly different from zero or a random classification with the Z statistic. The Z statistic in the Kappa analysis can also be used to test if a classification is significantly different from another classification.

Results

Field Verification

Field data were collected on a total of 316 sites (Table 2) during five days in late August 1995. Thirty percent of the field sites were set aside for accuracy assessment. With no major roads or transportation within the study area, helicopters were the most cost effective method of collecting field data.

The 300% level on the field form was only filled out during the first few field samples. The extra time needed to collect the 300% level was determined to be a poor use of the little remaining helicopter time due to mechanical difficulties (even without the 300% level recorded, there were inadequate field samples for the accuracy assessment). In addition, the reliability of calling vegetation under the canopy from an elevated position in the helicopter has yet to be determined.

Classification

The results of the spectral pattern and DIVERGE analysis in Imagine was the selection of bands 3, 4, 5, 4/3 ratio and the first principal component of the visible bands. This band combination was used in the maximum likelihood classification of the upland classes. The wetland classes had better results with bands 1-4. The visible bands and the near infrared band provide better separation between the wetland classes because water absorbs the infrared wavelength and thus produces low reflectance that makes it hard to separate the different vegetation types.

A limited number of digital elevation models (DEMs) were to be provided by the USGS. Therefore, the mountain areas within the MOAs were the highest priority and those areas outside of the MOA were the lowest priority. The DEMs received from the USGS covered all of the MOA and a portion of the areas outside the MOAs (Figure 6). The topographic normalization and alpine tundra modeling were not performed on the area outside the DEMs. Additional editing was performed on the area outside the DEMs to alleviate some of the problems where the DEMs did not exist.

Upon reviewing the 1000-meter contour for modeling the alpine tundra, it was determined to do further analysis within the DEMs. By plotting the 1000-meter contour over the imagery, it can be seen that much of the shrub lands surrounding the Alaska range would have been considered alpine tundra (Figure 7). Therefore 100-meter contour intervals were generated from 1000 meters to 1500 meters in order to determine the best contour interval to use in the alpine tundra modeling (Figure 8). In addition, the field sites that were labeled as alpine tundra in the field by the biologist were summarized with the DEM to provide assistance with assigning a contour interval. Various locations around the study area were viewed in detail and analyzed by a biologist to determine the proper interval. It was decided that the 1200-meter level would provide the best fit for the entire study area. The final classification resulted in a single GIS data layer composed of 24 unique landcover classes for the project area (Table 3). The percentage of each landcover

Table 2. The number of field sites collected and the number of accuracy assessment sites withheld.

Class Name	# of Field Samples	# of Field Samples for Accuracy Assessment
Closed Needleleaf	25	4
Open Needleleaf	34	12
Closed Broadleaf	42	22
Open Broadleaf	13	0
Closed Mixed	27	6
Open Mixed	0	0
Tall Shrub	27	8
Low Shrub	25	14
Dwarf Shrub	26	9
Graminoid	19	7
Bryoid/Lichen	10	2
Alpine Tundra	24	9
Sparsely Vegetated	10	3
Gravel/Rock	6	3
Mud/Silt/Sand	5	0
Wet Herbaceous	16	1
Aquatic Bed	3	0
Clear Water	2	0
Turbid Water	1	0
Ice/Snow	1	0
TOTAL	**316**	**100**

category within the project area is low shrub 20.2%, other 15.4%, open needleleaf forest 14.9%, closed mixed forest 6.3%, closed deciduous forest 5.7%, snow 5.0%, closed needleleaf forest 4.9%, tall shrub 4.2%, dwarf shrub 3.8%, graminoid 3.7%, alpine tundra shrub 2.8%, alpine tundra graminoid 2.5%, sparse vegetation 2.0%, bryoid/lichen 1.8%, open deciduous forest 1.6%, agriculture 1.3%, ice 1.2%, clear water 1.0%, turbid water 0.9%, wet herbaceous 0.4%, urban 0.3%, and aquatic bed 0.1% (Figure 9).

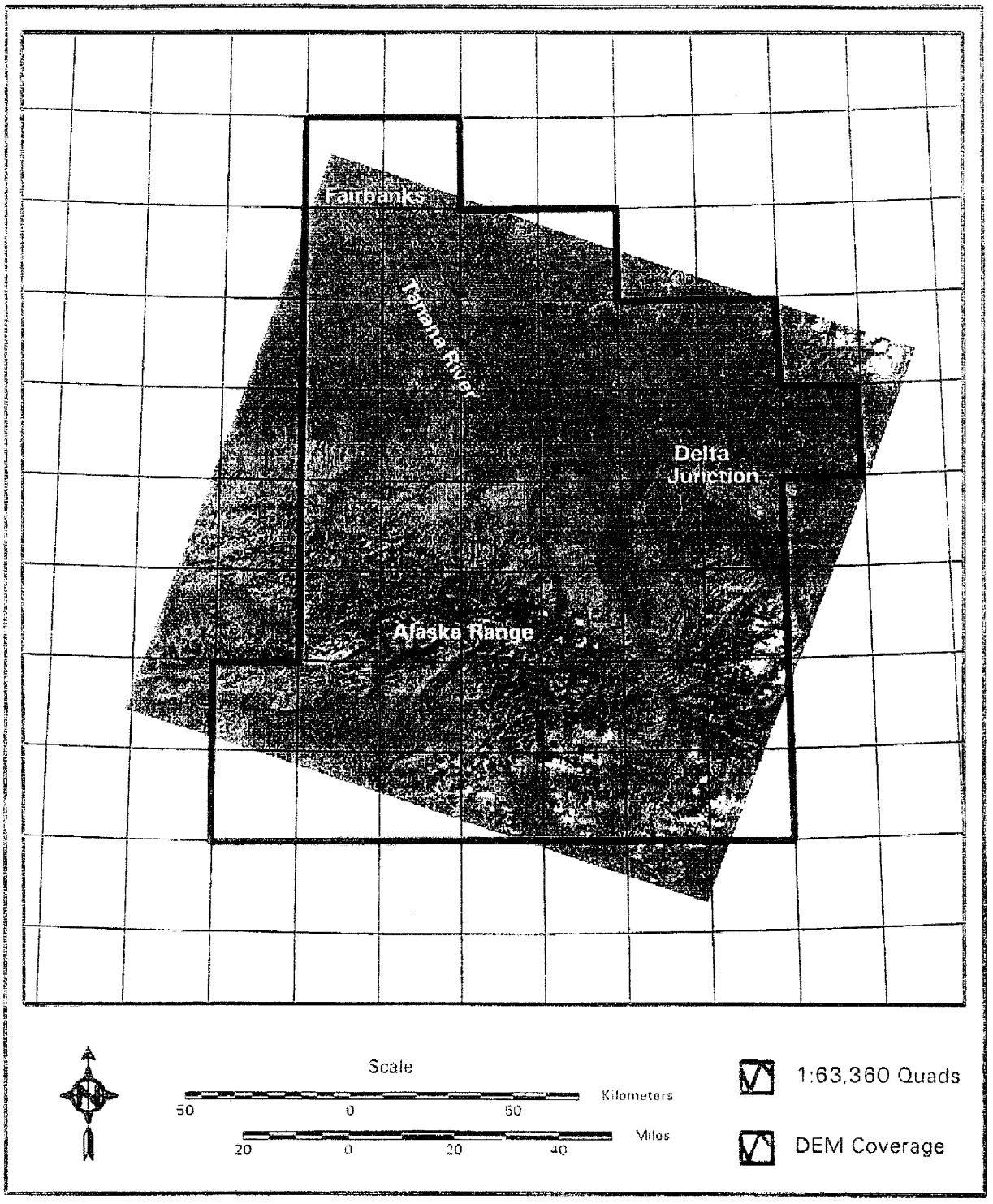


Figure 6. The digital elevation model coverage of the Tanana Flats project area

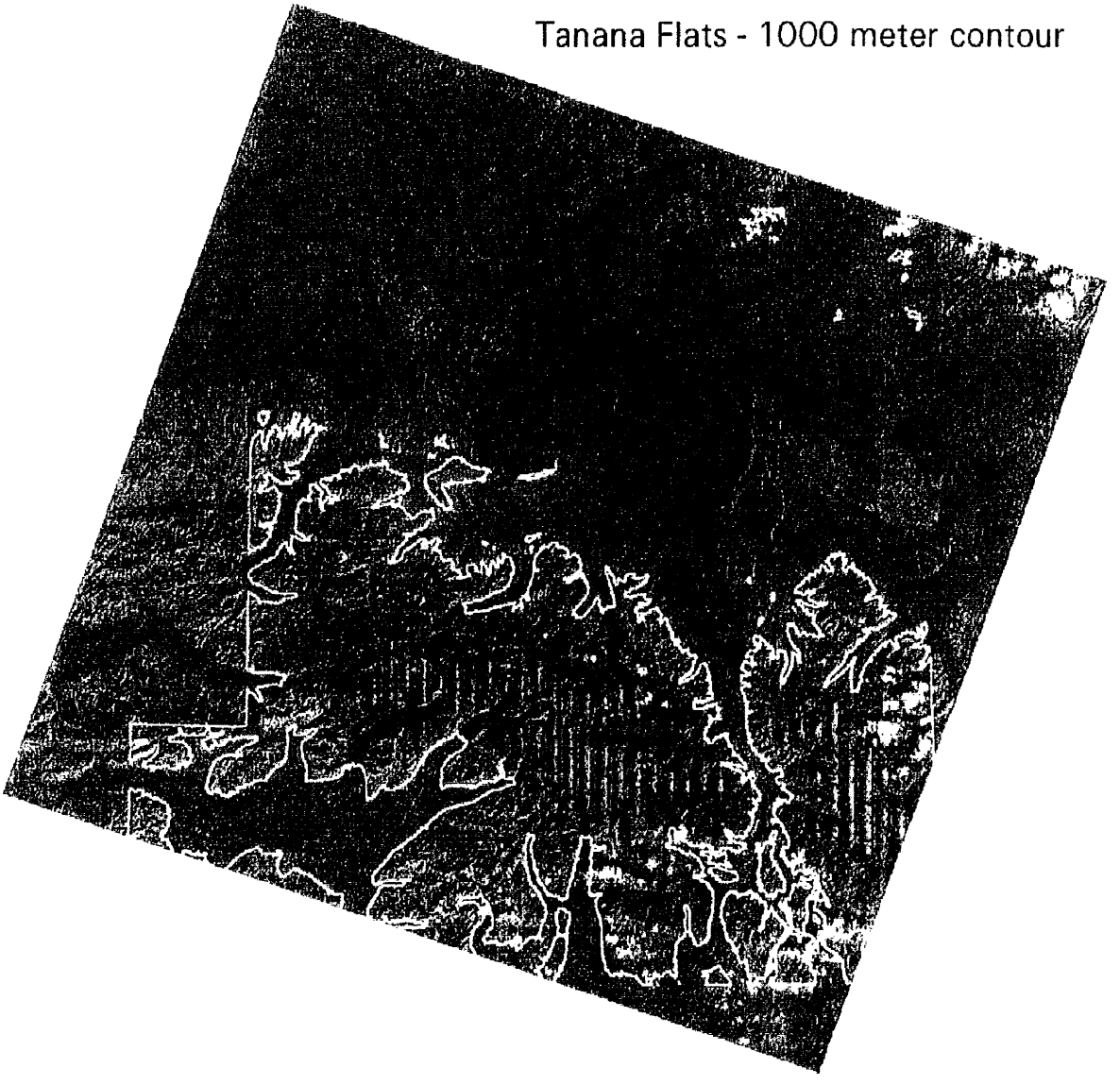


Figure 7. The 1000-meter contour for the Tanana Flats project

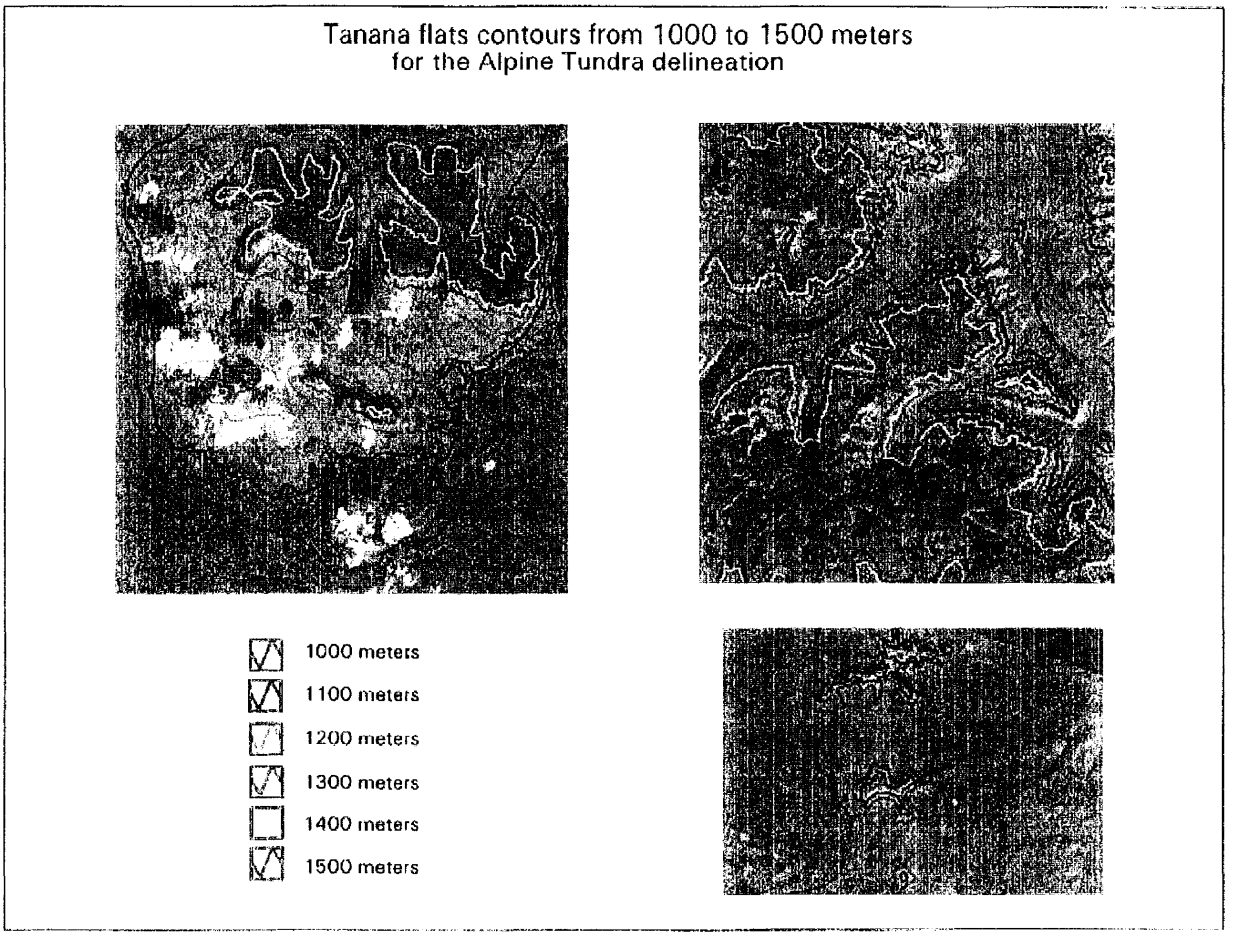


Figure 8. The 100-meter contour intervals from 1000 meters to 1500 meters

Table 3. The acreage statistics for each landcover category within the study area.

CLASS NAME	HECTARES	ACRES
Closed Needleleaf	164,430.36	406,316.5
Open Needleleaf	504,240.9	31,246,007.3
Closed Broadleaf	191,556.99	473,347.9
Open Broadleaf	54,581.85	134,874.7
Closed Mixed	212,799.42	525,839.1
Tall Shrub	141,105.87	348,680.4
Low Shrub	683,159.6	71,688,125.5
Dwarf Shrub	126,993.69	313,808.4
Graminoid	123,724.53	305,730.1
Bryoid/Lichen	60,606.27	149,761.4
Alpine Tundra - Graminoid	83,788.56	207,046.1
Alpine Tundra - Shrub	95,998.14	237,216.7
Wet Herbaceous	11,790.54	29,135.0
Aquatic Bed	3,392.28	8,382.5
Clear Water	34,161.39	84,414.6
Turbid Water	30,521.07	75,419.2
Ice	41,989.50	103,758.3
Snow	168,844.50	417,224.1
Sparse Vegetation	68,590.62	169,491.2
Gravel/Rock	220,206.42	544,142.3
Mud/Silt/Sand	193,231.98	477,486.9
Urban	10,462.68	25,853.8
Agriculture	45,050.04	111,321.1
Other	113,488.11	280,435.4
TOTAL	**3,384,715.4**	**81,363,819**

Accuracy Assessment

Due to the low number of field samples collected in the field, only 100 samples were set aside for the accuracy assessment. An additional 234 accuracy assessment samples were generated using color infrared aerial photography (Table 4). A few of the landcover classes were not included in the accuracy assessment due to inadequate field data and/or difficulty in photointerpreting. The "Open Needleleaf" class was not included because of the low number of field samples and the ability to interpret this class from aerial photography. The "Alpine Tundra – Graminoid" and "Alpine Tundra – Shrub" were merged into one class (alpine tundra) for the accuracy assessment, also due to the low number of field samples and the ability to interpret this class from aerial photography. The "Turbid Water", "Snow" and "Ice" classes were not included in the accuracy assessment due to temporal differences between the field data collection, aerial photographs and imagery. The "Agriculture" class was left out because agriculture was limited to one area of the project and 100% of the photos were used to classify the "Agriculture" class.

For the 15 remaining classes, the accuracy assessment showed an overall accuracy of

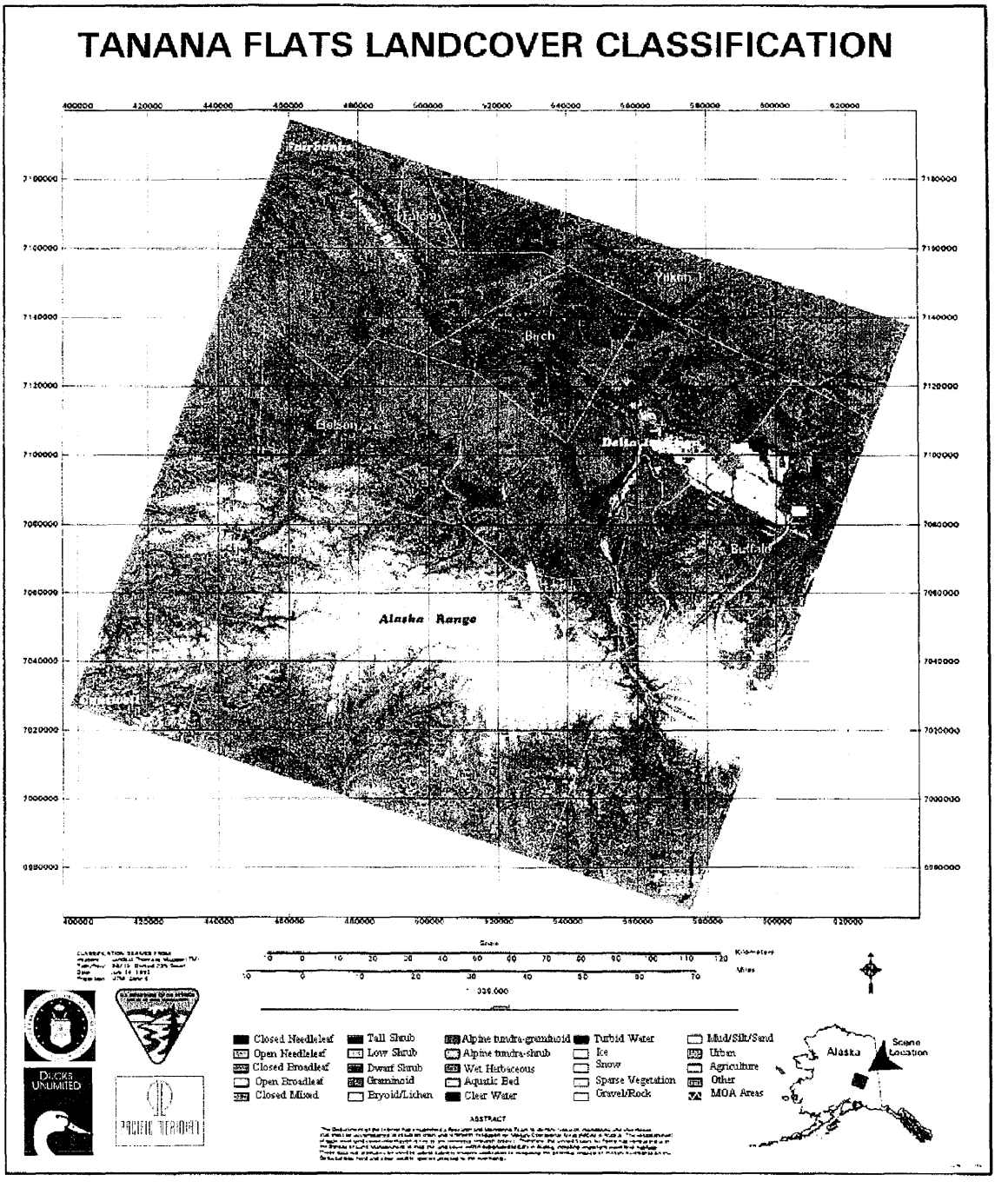


Figure 9. The final classification for the Tanana Flats project.

82% with a KHAT value of 0.81. The KHAT number is an evaluation of the overall accuracy that considers chance diagonal agreement. Landis and Koch (1977) characterized the possible ranges for KHAT into three groupings: a value greater then 0.80 (i.e., 80%) represents strong agreement; a value between 0.40 and 0.80 (i.e., 40 - 80%) represents moderate agreement; and a value below 0.40 (i.e., 40%) represents poor agreement. The Z statistic of Kappa analysis is 37.62. The Z statistic shows that the classification was significantly different from a random classification at the 99% confidence level. A Z statistic of 1.98 or less means that the classification is not significantly different from a random classification at the 99% confidence level.

Table 4. The number of photo interpreted sites used in the accuracy assessment.

Class Name	Number of Photo Intrepreted Accuracy Assessment Sites
Closed Needleleaf	16
Open Needleleaf	18
Closed Deciduous	9
Closed Mixed	17
Tall Shrub	17
Low Shrub	17
Dwarf Shrub	17
Graminoid	21
Bryoid/Lichen	14
Alpine Tundra	10
Wet Herbaceous	16
Aquatic Bed	8
Water	17
Sparse Vegetation	16
Other	21
TOTAL	234

As an additional quality control, 42 of the 100 field sites withheld for the accuracy assessment were photo interpreted. An error matrix was then generated for the photo interpreted field sites (Appendix B). It is important to note that this matrix is not statistically valid and should not be considered a reliable estimate of accuracy. However, the matrix is very valuable in determining which classes the photo interpreter had problems interpreting. In this case, the photo interpreter had trouble with the difference between "Closed Deciduous Forest" and the "Tall Shrub" and between the "Low Shrub" and "Tall Shrub" categories. It also appears from this analysis that the classification was biased towards the photo interpretation which may have

introduced some errors that will not be reported in the accuracy assessment. In other words, the image processor used the aerial photographs extensively in the classification process due to the limited amount of field work. If the image processor misinterpreted a class (e.g. called a "Closed Deciduous Forest" a "Tall Shrub" class because the height difference was not clear on the 1:60,000 CIR photo) then the computer could have been trained poorly for that site. However, this does not mean that the accuracy assessment for this project is not reliable, it just suggests that there may be confusion between some of the classes that are hard to separate (like "Closed Deciduous Forest" and "Tall Shrub"). In order to eliminate the confusion between these classes, there would need to be a considerable and expensive collection of field data that was not possible for this project.

Final Products

The primary product of this project is a digital database of the 24 landcover classes. for the Tanana Flats project area. Hard copy maps with acreage statistics were also created of the entire project area at a scale of 1:250,000. Selected 1:63,360 scale maps were also produced with acreage statistics. In addition, for viewing purposes and future analysis, the 316 field samples were digitized from the field maps into ArcInfo and the field site database was related to the ArcInfo coverage. The result was an ArcInfo coverage of the field sites with the attribute data. A final report describing the analysis and technical design was also delivered to the client.

Conclusions

The Department of the Interior (DOI) has established a Research and Monitoring Committee to identify research, monitoring and inventories that must be accomplished to establish short- and long-term mitigation for Military Operational Areas in Alaska. The establishment of base level landcover information is key to any terrestrial research project. This project has developed a baseline landcover inventory for the Tanana Flats area from Landsat Thematic Mapper satellite imagery in Geographic Information System form. This landcover database will be used for applications such as mitigating the potential impacts of military overflights on the Delta caribou herd.

The project area was categorized into 24 landcover classes with an overall accuracy of 82%. In the future, a change in the classification decision tree for the "Shrub" class may warrant consideration. Currently, a site is considered a "Shrub" class if it has greater then 25% shrubs. In some cases in this project there was confusion because a site would be composed mostly of graminoids (e.g., 70%) but would still be considered a "Shrub" class because it had 30% shrub species. A more reasonable percentage for the "Shrub" class might be 40%. An addition to the classification decision tree might also be considered if the project had a need to separate the ericaceous shrubs from willow/alder. The indications from this project are that these classes could be separated if adequate field data were available.

Literature Cited

Chuvieco, E. and R.G. Congalton. 1988. Using cluster analysis to improve the selection for training statistics in classifying remotely sensed data. Photogrammetric Engineering and Remote Sensing. 54:1275-1281.

Congalton, R.G., K. Green, J. Teply. 1993. Mapping old growth forest on national forest and park lands in the Pacific Northwest from remotely sensed data. Photogrammetric Engineering and Remote Sensing. 59:529-535.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment. 37:35-46.

Gopal, S. and C. Woodcock. 1992. Accuracy assessment of the Stanislaus vegetation map using fuzzy sets. In Remote sensing and natural resource management. proceedings of the fourth forest service remote sensing applications conference. American Society for Photogrammetry and Remote Sensing. pp. 378-394

Kempka, R.G., R.D. Macleod, J.F. Payne, F.A. Reid, D.A. Yokel, G. Balough. 1995. National Petroleum Reserve-Alaska landcover inventory: exploring arctic coastal plain using remote sensing. Proceedings of GIS94, Vancouver, British Columbia. February 1995. pp. 788-796.

Kempka, R.G., B.S. Maurizi, F.A. Reid, R.C. Altop, and J.W. Denton. 1994. Standardizing reference data collection for satellite landcover mapping in Alaska. Proceedings of GIS94, Vancouver, British Columbia. February 1994. pp. 419-426.

Kempka, R.G., F.A. Reid, R.C. Altop. 1993. Developing large regional databases for waterfowl habitat in Alaska using satellite inventory techniques: A Case Study of the Black River. Proceedings of GIS94, Vancouver, British Columbia. February 1993. pp. 1-11.

Landis, J. and G. Koch. 1977. The measurement of observer agreement for categorical data. Biometrics. Vol. 33, pp. 159-174.

Markon, C.J, 1993. Steese White Mountains landcover mapping project users guide. EROS Alaska Field Office, Hughes STX Corporation. Anchorage, AK. P.16.

Story, M. and R. G. Congalton. 1986. Accuracy assessment: user's perspective. Photogrammetric Engineering and Remote Sensing. Vol. 16, No. 5, pp. 529-535.

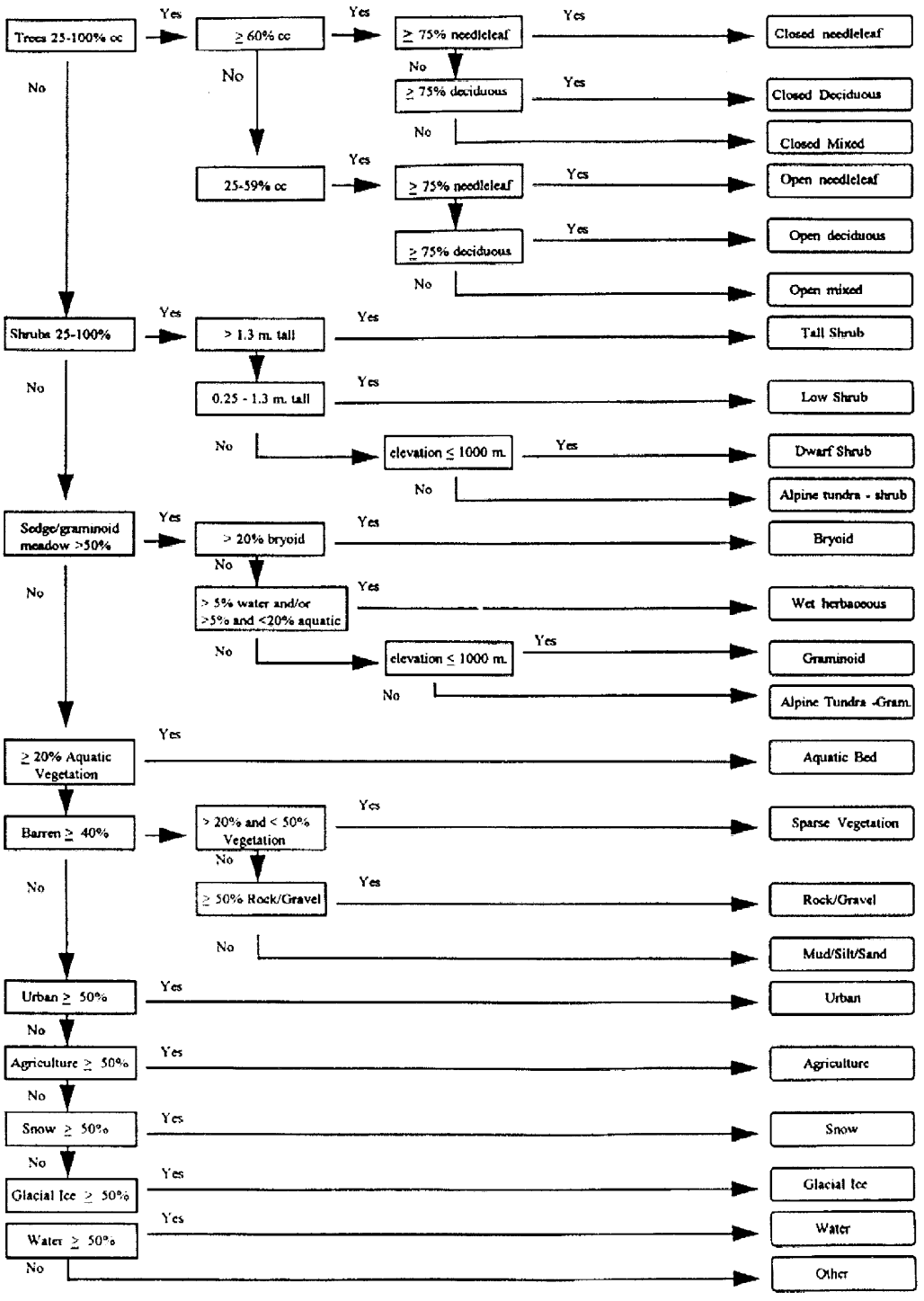
Tortora, R. 1978. A note on sample size estimation for multinomial populations. The American Statistician. Vol. 43, No. 9, pp. 1135-1137.

Viereck, L.A., C.T. Dyrness, A.R. Batten, and K.J. Wenzlick. 1992. The Alaska vegetation classification. Gen. Tech. Rep. PNW-gtr-286. Portland, OR:U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station. p.278.

Appendices

Appendix A

The decision tree for the Tanana Flats project



Appendix B

Error matrices for the Tanana Flats classification

TANANA FLATS
ACCURACY ASSESSMENT

CLASSIFICATION

REFERENCE	CNF	ONF	CDF	CMX	TSH	LSH	DSH	GRM	BLK	ATU	WHB	AQB	H2O	SVG	OTH	Total	USERS
CNF	18			1									1			20	90%
ONF	2	20		5		2	1									30	67%
CDF			26		3	2										31	84%
CMX			1	22												23	96%
TSH		2	1		20	2										25	80%
LSH		1			3	23	2	1	1							31	74%
DSH					1	2	18	4						1		26	69%
GRM						3	6	17	1					1		28	61%
BLK							1	2	11	1				1		16	69%
ATU										18				1		19	95%
WHB											17					17	100%
AQB												8				8	100%
H2O												1	16			17	94%
SVG							2							17		19	89%
OTH														2	22	24	92%
Total	20	23	28	28	27	34	30	24	13	19	17	9	17	23	22	334	
PRODUCERS	90%	87%	93%	79%	74%	68%	60%	71%	85%	95%	100%	89%	94%	74%	100%		

OVERALL ACCURACY =
273/334 =
82%

CNF = Closed needleleaf forest, ONF = Open Needleleaf forest, CDF = Closed Decedious forest, CMX = Closed Mixed Forest, TSH = Tall shrub, LSH = Low Shrub, DSH = Dwarf Shrub, GRM = Graminoid, BLK = Broid-Lichen, ATU = Alpine Tundra, WHB = Wet Herbaceous, AQB = Aquatic Bed, H2O = Water, SVG = Sparse Vegetation, OTH = Other

TANANA FLATS
JOB# 434

PACIFIC MERIDIAN RESOURCES
7/29/96

TANANA FLATS
ACCURACY ASSESSMENT

PHOTO INTERPRETATION

FIELD REFERENCE	CNF	ONF	CDF	CMX	TSH	LSH	DSH	GRM	BLK	ATU	WHB	AQB	H2O	SVG	OTH	Total	USERS
CNF	1															1	100%
ONF	1	2				1					1					5	100%
CDF			6	2	4											12	50%
CMX																0	100%
TSH																0	100%
LSH						6		1								7	86%
DSH						2	2		1							5	40%
GRM		1						2	1							4	50%
BLK																0	100%
ATU										5						5	100%
WHB											1					1	100%
AQB																0	100%
H2O																0	100%
SVG									1							1	0%
OTH															1	1	100%
Total	2	3	6	2	4	9	2	3	3	5	2	0	0	0	1	42	
PRODUCERS	50%	67%	100%	0%	0%	67%	100%	67%	0%	100%	50%	100%	100%	100%	100%		

OVERALL ACCURACY =
26/42=
62%

CNF = Closed needleleaf forest, ONF = Open Needleleaf forest, CDF = Closed Decedious forest, CMX = Closed Mixed Forest, TSH = Tall shrub, LSH = Low Shrub, DSH = Dwarf Shrub, GRM = Graminoid, BLK = Broid-Lichen, ATU = Alpine Tundra, WHB = Wet Herbaceous, AQB = Aquatic Bed, H2O = Water, SVG = Sparse Vegetation, OTH = Other

TANANA FLATS
JOB# 434

PACIFIC MERIDIAN RESOURCES
7/29/96

Appendix C

Contact information

The following additional data is available:

ArcInfo coverages
Final map classification in ERDAS Imagine format
Final map compositions in Imagine 8.2 format
Raw Landsat TM and DEM imagery
Field database files and FoxPro data entry program
ArcInfo coverage of aerial photograph flight lines

For more information please contact:

Bureau of Land Management
Alaska State Office
222 West 7th Avenue, #13
Anchorage, AK 99513-7599
907-271-3431

Ducks Unlimited, Inc.
3074 Gold Canal Drive
Rancho Cordova, CA 95670-6116
916-852-2000

United States Department of the Air Force
611CES/CEVP
10471 20th Street
STE 320
Elmendorf AFB
Anchorage, AK 99506-2200