

**ALBERTA GROUND COVER CHARACTERIZATION (AGCC)**

**TRAINING AND PROCEDURES MANUAL**

**BY**

G.A. Sánchez-Azofeifa, M. Chong, J.Sinkwich, S. Mamet

Earth Observation Systems Laboratory  
Department of Earth and Atmospheric Sciences  
University of Alberta  
Edmonton, Alberta  
Canada

**FINAL DRAFT**  
**APRIL 15<sup>TH</sup>, 2005**



## **Abstract**

The procedures necessary for a preliminary land cover classification of Landsat 7 ETM+ imagery according to the Alberta Ground Cover Characterization (AGCC) methodology are detailed in this manual. Land cover classifications completed following the AGCC methodology have been employed in studies involving wildfire and forest mapping. AGCC procedures have integrated a variety of requirements and initiatives from associated partners that has resulted in a beneficial planning tool for environmental conservation and sustainable development. The procedure documented here are such that researchers without a priori knowledge of land cover classification may perform a land cover classification in the province of Alberta unaided. In this manual, a brief description of the Landsat sensor and its data were given followed by detailed instructions regarding the search and acquisition of these data. An explanation of the hierarchical classification approach applied in the AGCC was presented followed by a detailed description of the procedures of classification. The procedures encompassed within this document include classification methodology, the planning and conducting of fieldwork, accuracy assessment on the classified product; exportation of AGCC data in terms of raster to vector conversion, metadata standards and procedures for creating metadata for the product, and lastly time management and allocation for completion of an AGCC classification.

## **Table of Contents**

### **1. List of tables and figures**

### **2. Disclaimer**

### **3. Introduction**

#### **3.1** Brief description of the organization of the manual and how to use it

#### **3.2** Definitions and sensors

- i. Sensor spatial resolution
- ii. Sensor spectral resolution
- iii. Temporal resolution
- iv. Data cost

#### **3.3** The AGCC land cover classification system

### **4. Remote sensing characteristics, Landsat TM data search and acquisition.**

#### **4.1** The Landsat Thematic Mapper (TM) sensor and its characteristics

#### **4.2** Data search and acquisition

### **5. Image processing**

#### **5.1** Hierarchical Classification Approach

#### **5.2** Comprehensive step-by-step procedure for classifying AGCC

- i. Data compilation and preparation (processing Landsat 7 ETM imagery)
- ii. Downloading and Compiling GeoTIFF File Types in PCI Geomatica
- iii. Applying Top of the Atmosphere Reflectance (TOAR)
- iv. Processing the normalized image

### **6. Classifying for AGCC**

#### **6.1** Hierarchical Classification Methods

- i. Extracting known features (roads and hydrography)
- ii. Subset of known clearcuts and burns
- iii. Isolating and extracting ancillary features
- iv. Isolating land cover types and internal classification
- v. Reassembly

## **7. Conducting fieldwork.**

### **7.1 Planning for fieldwork**

- i. Selection of field sites
- ii. Planning the route
- iii. Preparation of field maps

### **7.2 Preparation of field crew**

- i. Field safety
- ii. Field equipment

### **7.3 Description of how to conduct AGCC field calls**

### **7.4 Processing Field Data**

## **8. Accuracy assessment of AGCC data sets**

### **8.1 Descriptions of how to conduct an accuracy assessment on the AGCC Level 3.**

- i. Data Preparation
- ii. Image Analysis
- iii. Accuracy
- iv. Revision

## **9. Exporting AGCC data sets**

### **9.1 Description of steps and algorithms used to transform from raster to vector the information extracted from the TM imagery.**

### **9.2 Description of steps and algorithms used to produce a map using a given minimum mapping unit.**

- i. Area calculation
- ii. Minimum mapping unit

## **10. Metadata for AGCC**

### **10.1 Description of the metadata standards used by AGCC**

### **10.2 Description of how to create metadata**

## **11. Time management for the AGCC**

### **11.1 Descriptions of the time required to complete tasks and processes for AGCC classifications**

## **12. Appendices**

## 1. List of tables and figures

Table 3.1. 1 AGCC Classification scheme .....	9
Figure 3.3. 1 Relationship between AGCC levels and expected accuracy.....	14
Figure 4.1. 1 Historical information of Landsat launch programs .....	15
Table 4.1. 1 Radiometric characteristics of the ETM+ and TM sensors .....	16
Table 4.2. 1 Useful search and order/download sites .....	17
Figure 4.2. 1 Landsat path and row coverage of Alberta .....	18
Figure 4.2. 2 Effects of scan line corrector fault on Landsat imagery.....	19
Figure 5.1. 1 Flow diagram of AGCC hierarchical classification approach .....	20
Figure 5.2. 1 Importing band 1 of imagery using PCU Geomatica Focus module .....	22
Figure 5.2. 2 Layer transfer of remaining 5 bands using PCI Geomatica Focus module.....	22
Figure 5.2. 3 Adding six additional unsigned 8-bit channels using PCI Geomatica Xspace.....	23
Figure 5.2. 4 Script storage of Top of the Atmosphere Reflectance (TOAR) script.....	24
Figure 5.2. 5 TOAR intergace when applying TOAR using PCI Geomatica EASI module.....	25
Figure 5.2. 6 Subsetting atmospherically corrected channels using PCI Geomatica Xpace .....	26
Figure 6.1. 1 Select desired attributes from vector layers using ArcGIS ArcMap.....	27
Figure 6.1. 2 Exporting selected data to a new shapefile .....	28
Figure 6.1. 3 Clipping one layer based on another using the GeoProcessing Wizard.....	28
Figure 6.1. 4 Reprojecting the clipped data using Arc Toolbox.....	29
Figure 6.1. 5 Unsupervised classification using Erdas Imagine.....	30
Figure 6.1. 6 Recoding the classification using Erdas Imagine.....	30

Figure 6.1. 7 Recoding via vector and AOI tools in Erdas Imagine.....	31
Figure 6.1. 8 Selecting and copying vectors to AOI; selecting AOIs.....	32
Figure 6.1. 9 Recoding selected AOI to appropriate land cover class.....	33
Figure 6.1. 10 Masking road and hydrography layer out of satellite image .....	34
Figure 6.1. 11 Unsupervised classification using Erdas Imagine.....	35
Figure 6.1. 12 Adding the separate layers using the Model Maker in Erdas Imagine (Note that the number of layers added may vary .....	39
Figure 7.1. 1 Field map containing the satellite image, field points, and access information.....	43
Figure 8.1. 1Exporting the classified image to generic binary using the export function in Erdas .. Imagine .....	50
Figure 8.1. 2 Ground point data fields in Excel.....	51
Figure 8.1. 3 Image Analysis interface.....	52
Figure 8.1. 4 Accuracy interface, confusion matrix .....	53
Figure 9.1. 1 Converting raster to vector using ArcGIS ArcToolbox .....	55
Figure 9.2. 1 Adding an <b>area</b> field .....	55
Figure 9.2. 2 Calculating area using the field calculator .....	56
Figure 9.2. 3 Selecting by attribute, polygons of a minimum mapping unit of 1ha (10, 000 m <sup>2</sup> ) or .....	57
Figure 9.2. 4 Eliminating selected polygons .....	57
Figure 10.2. 1 Metalite application using the ArcGIS ArcCatalog interface .....	58
Figure 10.2. 2 Import metadata function in ArcCatalog.....	59
Figure 10.2. 3 Export metadata function in ArcCatalog.....	59
Figure 10.2. 4 Sample metadata in Metalite ArcCatalog interface.....	60
Figure 10.2. 5 Sample metadata in html format .....	60

## **2. Disclaimer**

The University of Alberta recognizes that research output will be used by the scientific community, industry, organizations, and the public in various ways. Reflecting on this, the University Research Policy Committee accepted, in February 1987, the following statement as a disclaimer on this use of research or products:

The University of Alberta encourages its staff to carry out and publish scientific and scholarly research of the highest quality and consistent with the best ethical standards. Nevertheless it does not and cannot endorse or guarantee the results or the products of research performed by its staff or students on or off campus, nor is it responsible for possible inaccuracies, errors or misinterpretations. In commercialization of its research by a company, the latter must agree to hold the University harmless in the event of legal action against the company. Approved by the University Research Policy Committee, February 1987.

Given the above policies regarding data use, the classifications produced using the methodology outlined in this document result in PRELIMINARY classifications that are not without error. Aspects of the satellite classification introduce uncertainty (ie. spatial resolution and spectral mixing) to the results. While the EOSL has taken reasonable measures to ensure that the data produced is accurate, the data may be deemed inaccurate depending on the application or purpose of the end user. These data should not be construed as absolute, but as a representation of the land cover at a given time for the primary purpose of regional land cover mapping. As such, the University of Alberta and the Earth Observation Systems Laboratory (EOSL) accepts no responsibility or liability for the misuse or misinterpretation of the data produced using the methodology described here.

The methodologies applied here are a part of the Alberta Ground Cover Characterization (AGCC) and the Alberta Earth Observations for the Sustainable Development of Forests (EOSD) Land Cover Mapping Initiative. The stakeholders to date consist of the Canadian Forest Service, Alberta Sustainable Resources Development, and the University of Alberta.

### **3. Introduction**

#### **3.1 Brief description of the organization of the manual and how to use it**

Assessing land cover on a regional scale is becoming increasingly important in Alberta where government, industry, and academia are looking to balance conservation and development interests. The Alberta Ground Cover Characterization (AGCC) is an initiative between the University of Alberta and partners in the provincial and federal government to map the ground cover for the province of Alberta using medium resolution satellite data. The AGCC has been applied as a component of a larger initiative by the federal government to map the land cover of Canada as a whole by translating the end classification product to accommodate the Earth Observation for the Sustainable Development of Forests (EOSD) criteria. These data have also been applied to wildfire mapping models by government at the provincial level. Furthermore, other federal partners such as Canadian Wildlife Service and non-profit organizations such as Ducks Unlimited have employed the data to aid in the detection of areas of interest including various wetland types. The products created using the AGCC method of classification have integrated a variety of needs and ideas from all partners that has resulted in a valuable planning tool for conservation and sustainable development.

The AGCC class scheme is organized into five categories. The first category involves anthropomorphic features that include urban and industrial areas, agricultural lands, clearcuts and burned regions. The second uplands category includes forested lands, shrubland, and grasslands. The third contains wetlands and water, the fourth is barren land for areas which are primarily rock covered or are permanently covered in ice or snow, and finally the fifth category is anything that remains unclassifiable such as areas that are obscured by clouds and shadows. These five categories are broken down further into ninety-nine land cover classes (see Table 3.1.1).

This manual will attempt to act as a guide for an individual with no prior knowledge of the methods used in the AGCC classification process on how to conduct a classification according to the AGCC classification scheme of an area within the province of Alberta. It will be arranged into 9 sections that follow a logical sequence of procedures for creating the AGCC classification. Beginning in section 3, the introduction, there will be discussion of remote sensing data types and sources for the acquisition of the data used in the AGCC. Section 4 will discuss Landsat data search and acquisition. The fifth section will give a brief overview of the hierarchical classification approach used in the AGCC as well as instructions on how to process raw imagery involving data compilation and atmospheric normalization. The following section will give a detailed account of the methodology employed for the classification of the image into the land cover classes outlined by the AGCC strata. Section 7 will look at conducting the



fieldwork involved and section 8 will concern the accuracy assessment of the classification based on the field data gathered. To follow will be discussion of exporting the finished classifications using a minimal mapping unit, conversion to third party classification schemes, and translating from raster to vector coverage. Section 10 will look at metadata standards that have been employed and describe the procedures to create useful metadata. Lastly, the final topic of discussion will concern time management and allocation in conducting the AGCC work from start to finish.

Table 3.1. 1 AGCC Classification scheme

## **ANTHROPOMORPHIC**

### **1. Urban and Industrial**

- 10 Undifferentiated Urban and Industrial Features
- 11 Urban (cities, towns - mostly residential and downtown core areas)
- 12 Commercial and industrial (industrial parks, heavy oil sand development, Refineries, hydro generating facilities)
- 13 Major roads, highway and railways
- 14 Cutlines and Trails.
- 15 Surface mines (coal) gravel pits, spoil piles
- 16 Farmstead and/or ranch (including shelter belts)

### **2. Agriculture**

- 20 Undifferentiated Agriculture
- 21 Cropland (including cereal crops and forage)
- 22 Irrigated Land
- 23 Agricultural Clearing (recently cleared land, often with windrows)

### **3. Clearcuts**

- 30 Undifferentiated clear-cut
- 31 Graminoid (grasses/sedges) dominated clear-cut
- 32 Tree/shrub dominated clear-cut
- 33 Tree (replanted - immature trees, <20 years old) dominated clearcut

### **4. Burns**

- 40 Undifferentiated burn
- 41 Graminoid (grasses/sedges) dominated burn
- 42 Tree/shrub dominated burn
- 43 Tree dominated burn
- 44 New Burn
- 45 Herbaceous Burn

## B. UPLANDS

### 5. Forested Land (>6% tree cover)

*Coniferous Dominated Forest (>80% coniferous cover based on occurrence)*

Closed (>50% crown closure)		Open (6 - 50% crown closure)	
50	Closed Fir	150	Open Fir
51	Closed Black Spruce	151	Open Black Spruce
52	Closed Pine	152	Open Pine
53	Closed White Spruce	153	Open White Spruce
54	Closed Undifferentiated Conifer	154	Open Undifferentiated Conifer

*Deciduous Dominated Forest (>80% deciduous cover based on occurrence)*

Closed (>50% crown closure)		Open (6 - 50% crown closure)	
55	Closed Aspen, Balsam Poplar and/or Birch	155	Open Aspen, Balsam Poplar and/or Birch
5	Riparian Poplar		

*Mixed Wood Dominated Forest (20 - 80% mixed wood cover based on occurrence)*

Closed (>50% crown closure)		Open (6 - 50% crown closure)	
56	Closed Coniferous Dominated Mixedwood (60-80% Coniferous Cover)	156	Open Coniferous Dominated Mixed wood (60-80% Deciduous Cover)
57	Closed Deciduous Dominated Mixedwood (60-80% Deciduous Cover)	157	Open Deciduous Dominated Mixedwood (60-80% Deciduous Cover)
58	Closed Coniferous and Deciduous Cover (20-60%)	158	Open Coniferous and Deciduous Cover (20-60%)

### 6. Shrubland (>25% shrub cover and <6% tree cover)

Closed shrubland (streams and coolies)

- 61 Closed Riparian Shrub
- 62 Closed Coulee Shrub Thicket
- 63 Closed Upland Shrub

Open shrubland

- 161 Open Riparian Shrub
- 162 Open Coulee Shrub Thicket (chokecherry/buffaloberry, rose, buckbrush,  
saskatoon)
- 163 Open Upland Shrub
- 164 Open Sagebrush Flat

**7. Grassland (<25% shrub cover and <6% tree cover) and Upland Fords (<6% graminoid)**

Graminoids (grasses and sedges)

- 71 Fescue Grassland
- 72 Mixed Grassland
- 73 Sandhill Grassland
- 74 Coulee Grassland (grassland on steep coulee slopes)

Upland Fords

- 75 Upland Ford Meadow

**C. WETLANDS AND WATER**

**8. Wetlands**

- 81 Emergent Wetlands (cattails)
- 82 Graminoid Wetlands (sedges/grasses/forbs)
- 83 Shrubby Wetlands (willow and birch)
- 84 Sphagnum Bog
- 85 Lichen Bog
- 86 Black Spruce Bog (sphagnum understory)
- 87 Black Spruce Bog (lichen understory)
- 88 Undifferentiated Wetlands
- 89 Woodland Fen (Larch Drainage Flow Patterns)

**9. Water**

- 90 Lake, pond, reservoir, river and stream

**D. BARREN LANDS**

**10. Barren (<6% Vegetation cover)**

- 101 Permanent Ice and Snow
- 102 Rock, Talus, and/or Avalanche Chute
- 103 Exposed Soil
- 104 Alkali Flat and/or Mud Flat
- 105 Upland Dune Field
- 106 Alluvial Deposit
- 107 Beach
- 108 Badland
- 109 Blowout Zone

**E. UNCLASSIFIED**

- 112 Cloud / Haze
- 113 Shadow

### 3.2 Definitions and sensors

Evaluating remote sensing data needs for a given project requires an understanding of remote sensor data availability, characteristics, and cost. This section provides an overview of the factors considered when selecting imagery data for conducting the AGCC classifications. Four major factors that should be considered when choosing data include spatial resolution of the sensor selected, spectral and temporal resolution of the data acquired by that sensor, and lastly data costs.

#### *i. Sensor spatial resolution*

*Definition:* Spatial resolution is a measure of the smallest angular or linear separation between two objects that can be resolved by the sensor.

In a practical sense, a sensor's spatial resolution is defined as the dimension of the ground-projected instantaneous field of view (IFOV), or in other words, the pixel dimensions. The sensor used in the AGCC is the Landsat TM sensor, which has a pixel size of 28.5m x 28.5m for six out of seven bands. A useful rule-of-thumb for detecting features is that the spatial resolution of the sensor system should be less than one-half the size of the feature measured in its smallest dimension. For Landsat TM, detecting a land cover type such as forest would require forest patches of at least 57m in diameter. As the spatial resolution increases, more spatial detail will be apparent in the image. More spatial detail provides better information, but finer resolution also increases the amount of data to be processed. As such, imagery used for any project must be of sufficient spatial resolution to detect the desired objects or features.

#### *ii. Sensor spectral resolution*

*Definition:* Spectral resolution refers to the number and dimension of specific wavelength intervals in the electromagnetic spectrum to which a remote sensing instrument is sensitive.

A wavelength interval, or "band" is a region of the electromagnetic spectrum for which a sensor measures the intensity of reflected energy. Multispectral sensors measure radiant energy in multiple broad bands (>10 nm each). Landsat TM records data for 7 discrete bands. For any given sensor, detectors are designed to capture radiant energy in bands that will provide important information on land, sea or atmospheric attributes. Landsat TM bands were selected for their value in providing information on plant pigmentation (bands 1-3), internal leaf structure (bands 4) and leaf moisture content (bands 5 and 7).

### *iii. Temporal resolution*

*Definition:* The temporal resolution of a remote sensing system refers to how often it records imagery of a particular area.

Temporal resolution naturally dictates frequency of data acquisition. Various sensors will acquire data at various time intervals. Temporal resolution of a sensor is an important consideration when time-varying phenomena, such as phenological cycles of forests and agricultural crops, are being monitored. It is also important to recognize that cloud cover may obscure all or portions of an image, further limiting the availability of quality data. The Landsat TM sensors selected for the AGCC classifications have a revisit time of 16 days. This allows for some choice when selecting imagery within the season timeframes to detect phenological cycles of forests and agricultural crops for land cover mapping.

### *iv. Data cost*

Data acquisition and processing costs are additional considerations for program planning. Improvements in computer technology have allowed vast reductions in data processing costs, but the price of image data has varied widely in the past, depending on the availability to which government agencies subsidizes production and allow use. At present one scene of Landsat 7 ETM imagery may be purchased online from the Canadian Centre for Remote Sensing (CCRS), a federal remote sensing database, (<http://ceocat.ccrs.nrcan.gc.ca>) for as low as \$240 CAD to \$800 CAD as raw data. Orthorectified, these data could cost anywhere from between \$540 CAD to over \$1200 CAD. Another resource is RADARSAT International (RSI), a private spatial data and products distributor ([http://www.rsi.ca/products/sensor/landsat7/lsat7\\_pricelist\\_cad.asp](http://www.rsi.ca/products/sensor/landsat7/lsat7_pricelist_cad.asp)). One scene of Landsat 7 ETM imagery may be purchased online for \$800 CAD raw and \$1150 orthorectified. Photosat is another online Canadian resource for data purchase ([http://www.photosat.ca/products/landsat7/landsat\\_prices.shtml](http://www.photosat.ca/products/landsat7/landsat_prices.shtml)). Photosat is a private data distributor sublicensed by CCRS. A raw Landsat 7 ETM scene may be purchased for \$750 CAD and \$1100 CAD orthorectified. However given the above options for data purchase, Landsat 7 ETM data is becoming increasingly more available through free download sites such as GeoGratis subsidized by the federal government (<http://www.geogratis.ca>).

Ultimately, choice of dataset will be project-specific and must necessarily be a compromise between spatial, spectral and temporal resolution, as well as availability, continuity and cost, all within the context of the project in question, in this case the AGCC.

### 3.3 The AGCC land cover classification system

The AGCC in nature is a hierarchical land cover classification system that allows for different interpretation levels represented as AGCC levels 1, 2, 3, and 4. Each level will depend on the nature of the final product extracted from the remote sensing platform, the level of economic investment used on the generation of the dataset and most importantly, the limitations of the sensor to differentiate between one specific class and another. AGCC Level 1 is defined as the more simplistic extraction and involves in discrimination of broad features such as forest and non-forests. AGCC level 2 is more detailed than Level 1 but still simple in the class grouping: forest (deciduous, conifers, mixed wood) and non-forest (agriculture and urban land). AGCC level 3 presents a more comprehensive classification; the classification for each class involves the extraction of different crown closure levels in the dominant species groups (eg. open and closed deciduous). Finally AGCC Level 4 involves a comprehensive mapping of species. It should be noted that this level of classification is best undertaken using aerial photo interpretation and in future, may be achieved using hyperspectral sensors. Figure 3.3.1 presents the relationship between the various AGCC levels and the expected accuracy for each. A thorough description of the AGCC classification scheme is given in Table 3.2.1.

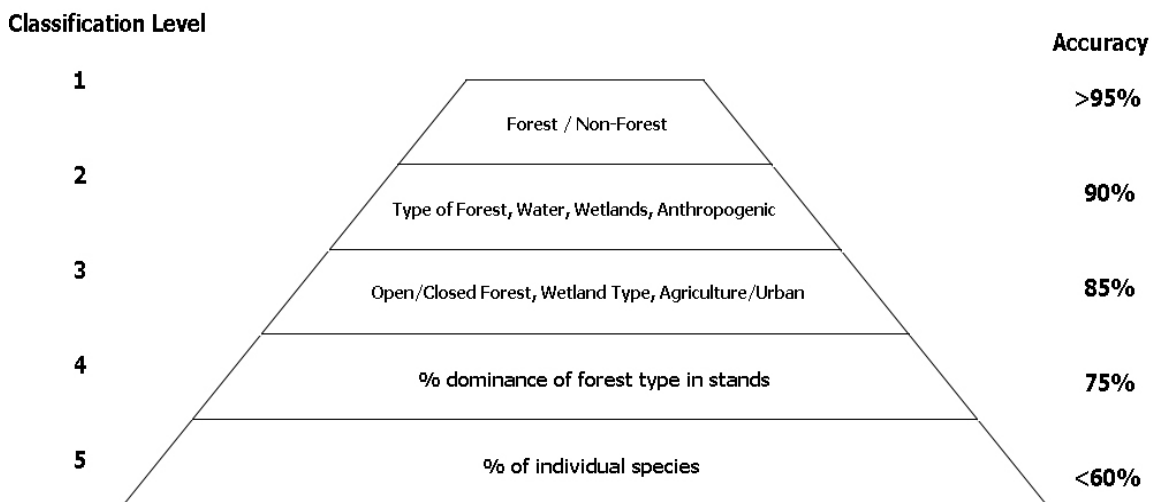


Figure 3.3. 1 Relationship between AGCC levels and expected accuracy

## 4. Remote sensing characteristics, Landsat TM data search and acquisition

### 4.1 The Landsat Thematic Mapper [TM] sensor and its characteristics

The Landsat project is the longest-running program instigated for the acquisition of moderate resolution imagery of the Earth. Landsat 1 was launched in July 1972. However, it was not until March 1984 that the launch of Landsat 5 resulted in an operational sensor that could provide continuous useable imagery of the Earth. The Multispectral Scanner (MSS), present on board each of the spacecrafts launched throughout the Landsat program provided information of 80 m pixel resolution and four spectral bands. This sensor provides information on the Earth's surface in the visible and near infrared portions of the spectrum. The U.S. Geological Survey (USGS) Earth Resources Observation Systems (EROS) Data Centre holds complete image data archives.

In July 1982, the launch of Landsat 4 saw the inclusion of the Thematic Mapper (TM) sensor with a 30 m pixel resolution and 7 spectral bands. Both sensors are now on Landsat 5 currently still acquiring and providing usable satellite imagery. Figure 4.1.1 gives a brief overview of the historical information of the Landsat launch programs and sensors on board.

Satellite	Launch Date	Notes
Landsat 1	23 July 1972	Decommissioned 6 January 1978
Landsat 2	22 January 1975	Decommissioned 25 February 1982
Landsat 3	5 March 1978	Decommissioned 31 March 1983
Landsat 4	16 July 1982	Decommissioned June 2001
Landsat 5	1 March 1984	Operational
Landsat 6	October 1993	Failed on launch
Landsat 7	15 April 1999	Operating in SLC-Off mode after May 2003

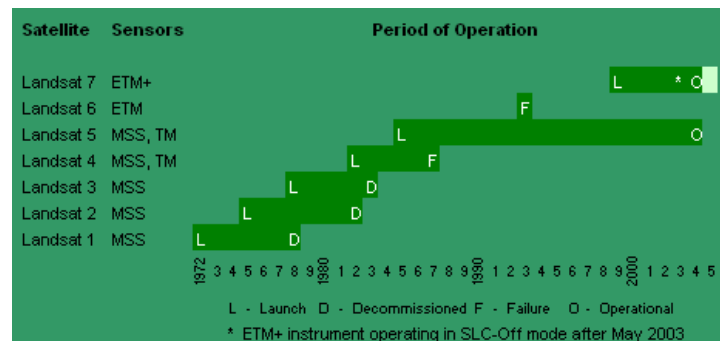


Figure 4.1. 1 Historical information of Landsat launch programs

The newest in this series of remote sensing satellites is Landsat 7, which was launched on 15 April 1999. Landsat 7 has the new Enhanced Thematic Mapper Plus (ETM+) sensor. This sensor has the same 7 spectral bands as its predecessor TM, but has an added panchromatic band with 15 m resolution and a lower resolution 60 m thermal band. The ETM+ sensor also has a five percent absolute radiometric calibration. Landsat 7 imagery is currently the imagery of choice for the AGCC project. The ETM+ product provides greater spectral variation allowing for more detailed analysis of the land cover present within the image.

The wide spectral range of both the ETM+ and the TM sensor allow for the identification and isolation of various types of land cover. As such, the Landsat 4/5 and 7 suite of imagery are ideal in the application of vegetation and land cover analyses. Table 4.1.1 shows the radiometric characteristics of the ETM+ and TM sensors and describes the application of each band as it is used to identify global land cover and water bodies.

Table 4.1. 1 Radiometric characteristics of the ETM+ and TM sensors

**Radiometric Characteristics of the ETM+ and TM Sensors**

Band Number	Spectral Range (in nm)	EM Region	Generalised Application Details
1	450 - 520	Visible Blue	Coastal water mapping, differentiation of vegetation from soils
2	520 – 590	Visible Green	Assessment of vegetation vigor
3	630 - 690	Visible Red	Chlorophyll absorption for vegetation differentiation
4	760 - 900	Near Infrared	Biomass surveys and delineation of water bodies
5	1550 – 1750	Middle Infrared	Vegetation and soil moisture measurements; differentiation between snow and cloud
6	10400- 12500	Thermal Infrared	Thermal mapping, soil moisture studies and plant heat stress measurement
7	2080 - 2350	Middle Infrared	Hydrothermal mapping
8	520 - 900 (panchromatic)	Green, Visible Red, Near Infrared	Large area mapping, urban change studies



## 4.2 Data search and acquisition

Landsat imagery is ideal for use in resources applications and land cover analyses, as Landsat's Global Survey Mission has successfully established a data acquisition strategy that ensures repetitive acquisition of the global landmass and water bodies. Furthermore, with its 16-day revisit period, Landsat imagery is readily available. Landsat imagery is now available to the public via a number of government and private distribution websites. Table 4.2.1 gives a number of search and order/download sites in addition to those mentioned in section 2.2 in the data cost discussion.

Table 4.2. 1 Useful search and order/download sites

<b>Earth Explorer</b>  Earth Explorer is a search and order tool for satellite imagery, aerial photography, and various other cartographic products distributed by the U.S. Geological Survey (USGS) ( <a href="http://edcsns17.cr.usgs.gov/EarthExplorer/">http://edcsns17.cr.usgs.gov/EarthExplorer/</a> ).
<b>GéoGratis</b>  GéoGratis is a Canadian government web and transfer protocol (ftp) site that distributes a geospatial data of Canada. Data is available for download, without charge. Vector data and satellite imagery are available in a variety of formats and scales ( <a href="http://geogratias.cgdi.gc.ca/clf/en">http://geogratias.cgdi.gc.ca/clf/en</a> ).
<b>Global Land Cover Facility</b>  The Global Land Cover Facility (GLCF) is an Earth science data distributor based out of the University of Maryland. The GLCF develops and distributes remotely sensed satellite data and products concerned with land cover on a local to global scale ( <a href="http://glcf.umiaccs.umd.edu/index.shtml">http://glcf.umiaccs.umd.edu/index.shtml</a> ). These data are available for download, without charge.
<b>GloVis</b>  The USGS Global Visualization Viewer (GloVis) is an interactive browse-based search and order tool for Landsat 7 ETM+, Landsat 4/5 TM, and Landsat 1-5 MSS USGS archived data ( <a href="http://glovis.usgs.gov/ImgViewer/ImgViewer.html?lat=&amp;lon=&amp;sensor=ETM">http://glovis.usgs.gov/ImgViewer/ImgViewer.html?lat=&amp;lon=&amp;sensor=ETM</a> ).

The Landsat 7 spacecraft follows the World Reference System – 2 (WRS – 2), a global notation used in cataloging Landsat data. This path row system makes it easy to find the desired imagery. Scenes within path 39 to 49 and row 18 to 26 cover the province of Alberta. Figure 4.2.1 illustrates the path and row coverage for the province. The path of a given scene runs from the east to the west and the row runs from the north to the south.

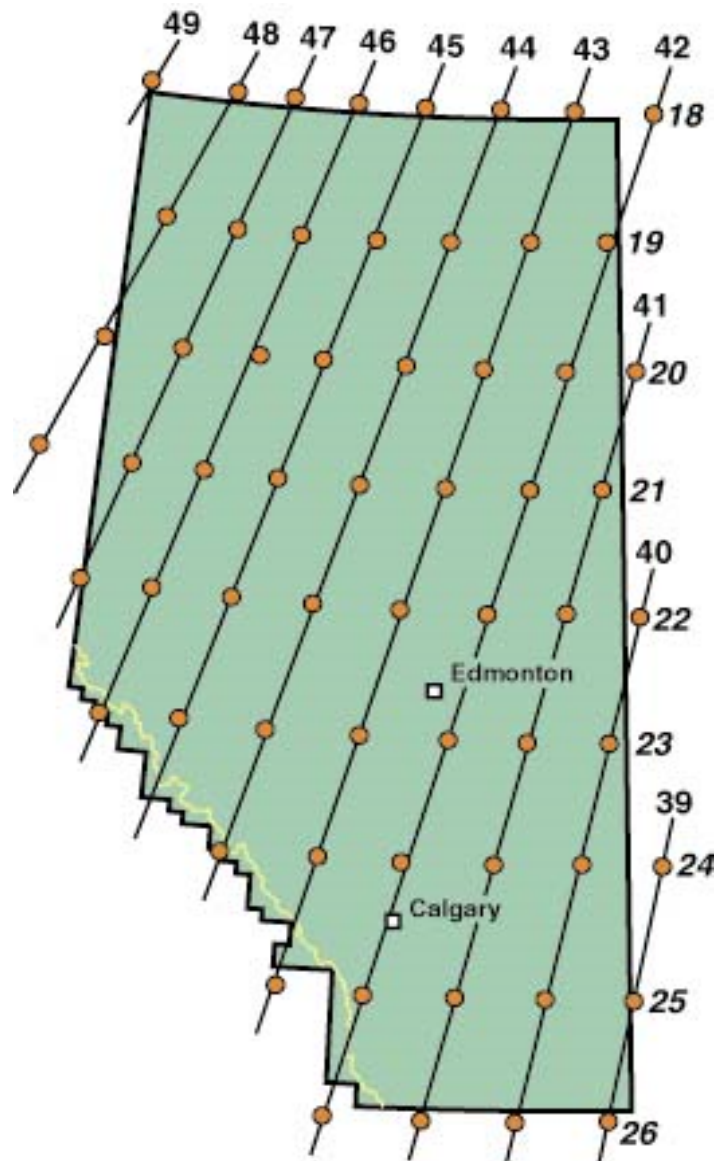


Figure 4.2. 1 Landsat path and row coverage of Alberta

It should be noted that on 31 May 2003, a fault with the Landsat 7 Scan Line Corrector (SLC) occurred resulting in continuous gaps in the processed products. However, corrected or “Gap-filled” data may now be obtained from the U.S. Geological Survey (USGS). These data have also become available for purchase and download from the various data distributors mentioned within this manual. For more details regarding the processes involved in correcting or filling the data refer to USGS Landsat 7 ETM+ SLC-Off Products website ([http://landsat.usgs.gov/gapfilled\\_off2off\\_release.php](http://landsat.usgs.gov/gapfilled_off2off_release.php)). Figure 4.2.2 illustrated the extent of the error and the effect on the imagery acquired.



Figure 4.2. 2 Effects of scan line corrector fault on Landsat imagery

## **5. Image processing**

### **5.1 Hierarchical Classification Approach**

The AGCC classification method employs a Hierarchical Classification/Hyperclustering approach in extracting land cover types. Hierarchical Classification refers to the process of systematically eliminating known features in an image to facilitate the classification process. Known features such as roads and hydrography are extracted using provincial datasets provided by Alberta Sustainable Resource Development (ASRD). With these features extracted, spectral confusion may be reduced during the classification process. Additional known features including clearcut and burn data are also provided by ASRD. These data are subset, further analyzed and classified to describe land cover change as a result of these anthropogenic activities. Figure 5.1.1 presents a comprehensive flow diagram that briefly describes the hierarchical method used to produce the AGCC land cover datasets in accordance to the AGCC Hierarchical Classification approach.

## Hierarchical Classification Method

The principle behind this method is to improve the separability of the signatures during the classification process; this is done by removing known or defined features from the landscape. In each level, a more undifferentiated feature than the last is processed and removed from the image. The final output is the assembly of the various layers that were extracted systematically.

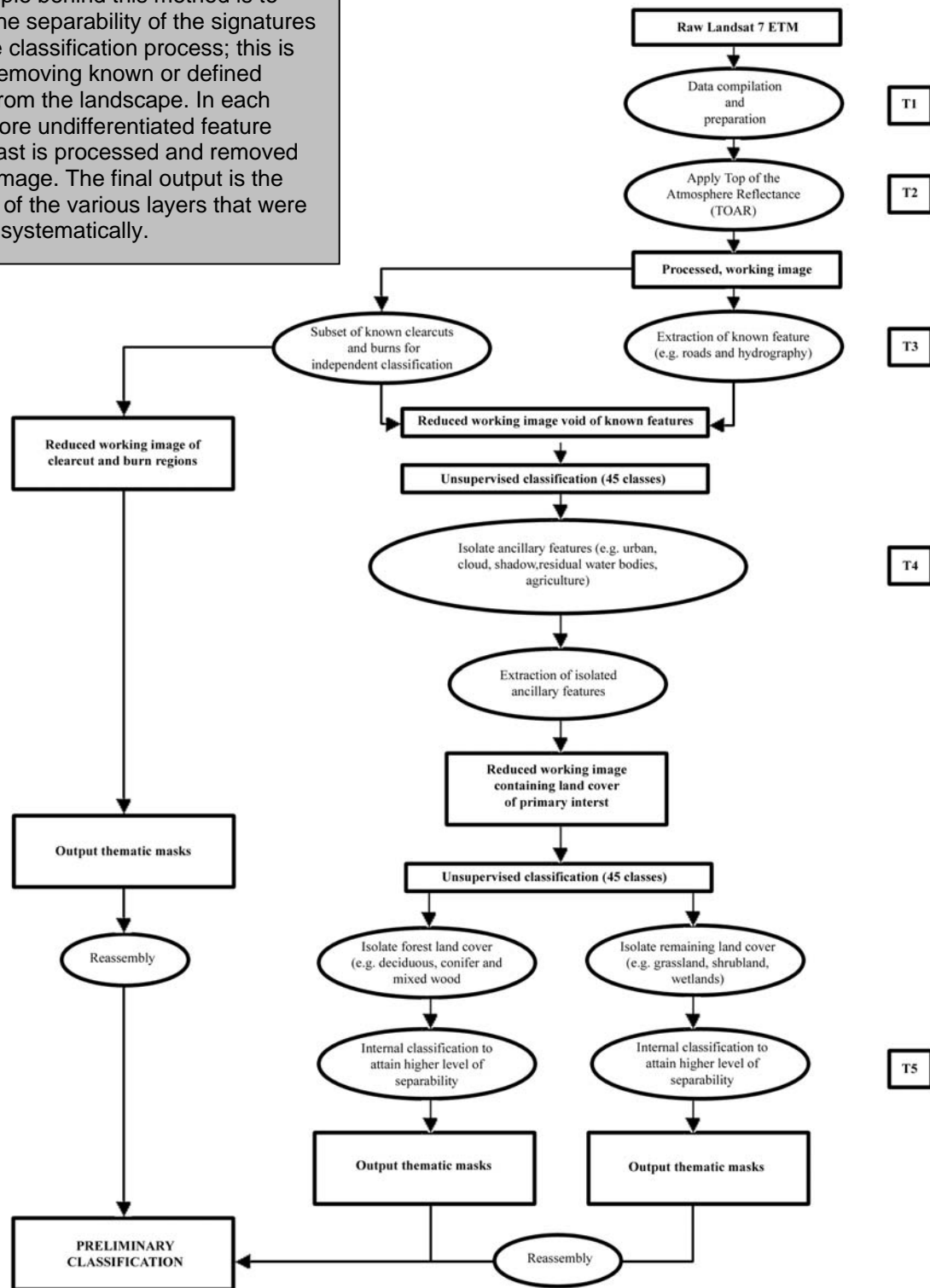


Figure 5.1. 1 Flow diagram of AGCC hierarchical classification approach

## 5.2 Comprehensive Step-by-Step Procedure for Classifying AGCC

### *i. Data Compilation and Preparation (Processing Landsat 7 ETM Imagery)*

A large portion of the Landsat 7 ETM imagery used in the AGCC classifications were obtained from a collection of imagery made available by the Centre for Topographic Information through GéoGratis, a web and file transfer protocol (FTP) site that distributes free spatial data for Canada (<http://geogratias.cgdi.gc.ca/clf/en>) or from the Global Land Cover Facility (GLCF) at the University of Maryland (<http://glcf.umiaccs.umd.edu/index.shtml>). When imagery was unavailable from these free download sites, data used was purchased from the distributors previously mentioned in this manual. Landsat 7 ETM imagery is currently available through GéoGratis and the GLCF already rectified and may also be purchased in this form. These data are downloaded one band at a time and must be compiled to comprise a full scene. Atmospheric normalizations must also be applied to these data once compiled. This normalization is carried out using a remote sensing software package PCI Geomatica 8.2. Data downloaded from both GéoGratis and GLCF are GeoTIFF file types. As such, the data must be imported and converted to the PCI file format (\*.pix) to allow for atmospheric normalization using the Top of the Atmosphere Reflectance routine (TOAR). The TOAR process undertakes normalization in solar illumination and transformation from digital numbers to TOA reflectance values. The following is a step-by-step description of the data compilation and atmospheric correction procedures.

### *ii. Downloading and Compiling GeoTIFF File Types in PCI Geomatica*

Download bands 1 through 5 and 7, as well as all associated metadata files and reports for the image of interest from the appropriate data distributor. Using PCI Geomatica, these data may be compiled and processed to result in a full scene ready for processing. The PCI Geomatica software suite contains a number of data processing and analyses modules. As seen in Figure 5.2.1, using the PCI Geomatica Focus module, import band 1 and save the output file using the full path and row name of the scene (ex. p47r21\_raw.pix). The remaining five bands are to be saved to this initial file to make up the full Landsat scene.

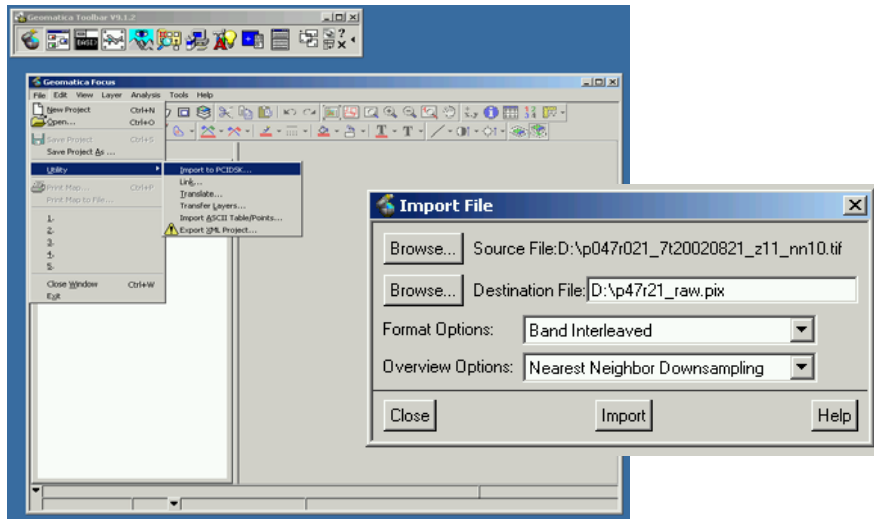


Figure 5.2. 1 Importing band 1 of imagery using PCU Geomatica Focus module

In PCI Geomatica Focus, the remaining five bands are transferred to the file created by importing band 1 (p47r21\_raw.pix). The various layers that will make up the full Landsat scene will be displayed in the Destination Layers window as seen in Figure 5.2.2. Ensure that the transferred layers are in ascending order and that band 1 is the imported layer.

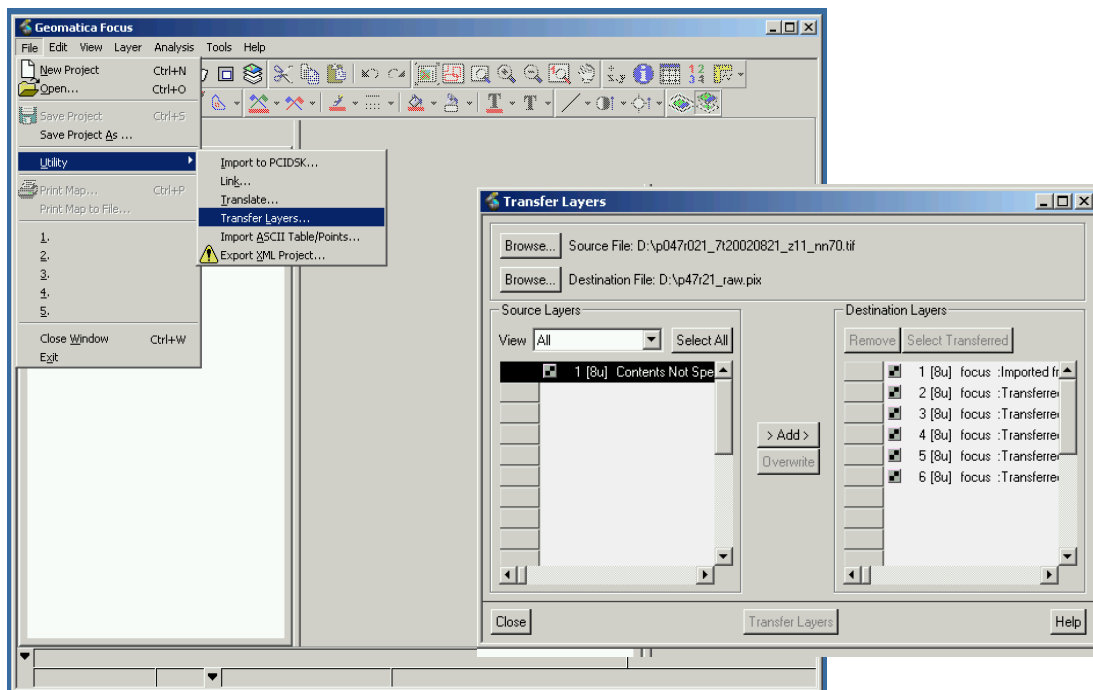


Figure 5.2. 2 Layer transfer of remaining 5 bands using PCI Geomatica Focus module

Once the data has been compiled, additional unsigned 8-bit channels must be added to the resultant image to accommodate the atmospherically corrected layers created. The atmospheric correction used in the AGCC is the Top of the Atmosphere Reflectance (TOAR). The TOAR is a module developed by Robert Landry at the Canadian Centre for Remote Sensing. This application is used to reduce or eliminate factors that affect the reflectance of a target that are unrelated to the target itself. Examples of such factors include variations in solar illumination and atmospheric transmission. Using the PCI Geomatica Xspace module, the six additional unsigned 8-bit channels required could be added (Figure 5.2.3). Open the full image using File Utility to enable channels to be added. The resultant product will be ready for atmospheric correction.

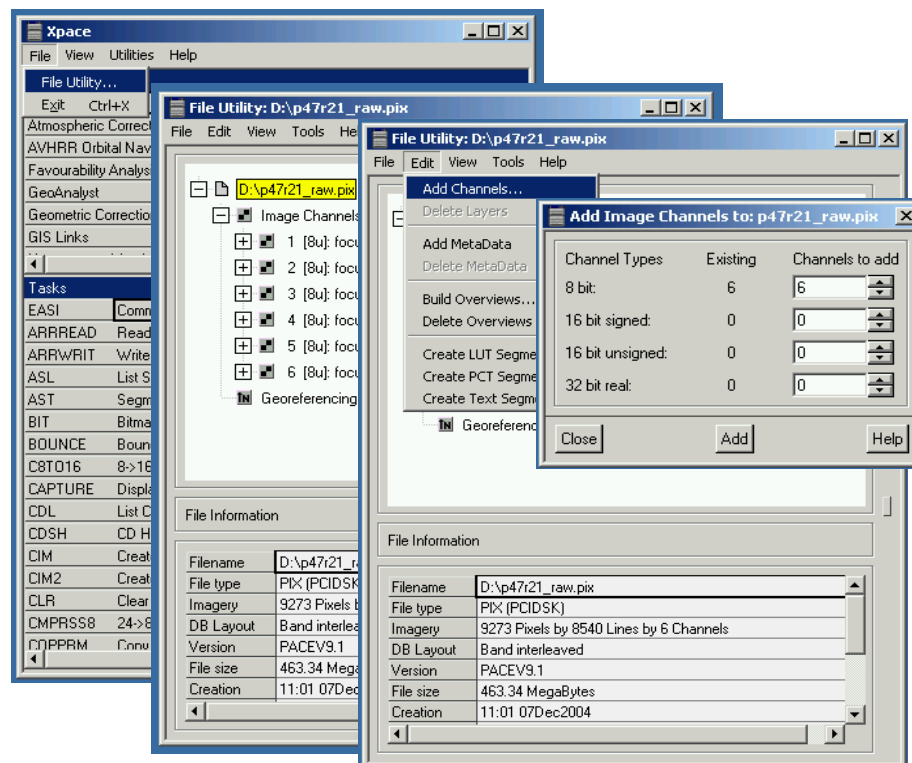


Figure 5.2. 3 Adding six additional unsigned 8-bit channels using PCI Geomatica Xspace

### iii. Applying Top of the Atmosphere Reflectance (TOAR)

To apply the Top of the Atmosphere Reflectance (TOAR) module, the script must first be downloaded to the local workstation. This script may be obtained from the Natural Resources Canada file transfer protocol site (<ftp://ftp.ccrs.nrcan.gc.ca/ftp/ad/landry>). Be sure to download the current version of the script that is compatible with the version of PCI Geomatica that is installed on the local workstation. The script must be stored in the “exe” directory within the Geomatica program files as seen in Figure 5.2.4. No installation is required.

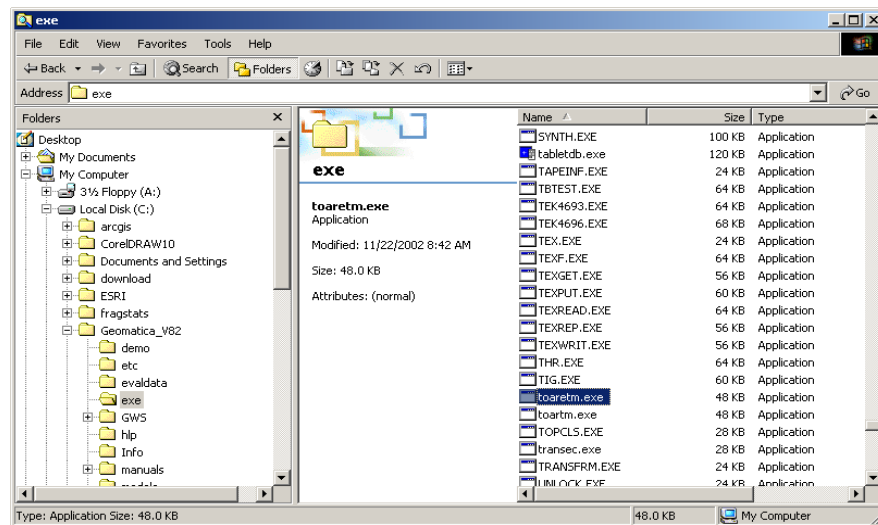


Figure 5.2. 4 Script storage of Top of the Atmosphere Reflectance (TOAR) script

To prepare all data for the TOAR process, the full scene image and its corresponding MTL file must be saved to the same working directory. The MTL file is a metadata file describing the processing parameters and information regarding the time and conditions at which the image was acquired. This MTL file must be in \*.LIG format. Simply modify the file extension if necessary. The MTL file will provide values and coefficients necessary to carry out the algorithms involved in the correction process.

The TOAR process is executed in the PCI Geomatica EASI module. From the prompt input: **s toaretm**. This will display the status of the TOAR module. The module will retain the information of the last image processed. The following is an example of the command prompts that will appear, though it will likely be void of information if the program is being executed for the first time (Figure 5.2.5).



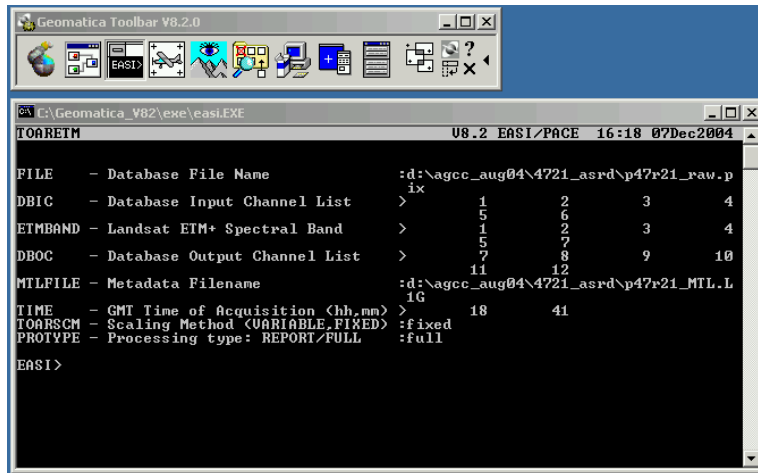


Figure 5.2. 5 TOAR intergace when applying TOAR using PCI Geomatica EASI module

The command prompt requires the input of eight lines of information.

1. FILE is the path and name of the image being processed.
2. DBIC or Database Input Channel is a list of the channels at which the individual bands of data to be processed is located.
3. ETMBAND is the Landsat ETM+ spectral bands that correspond to each of the input channels. Be sure to enter the band numbers in the correct order and note that band 6 was omitted (ie. 1,2,3,4,5,7).
4. DBOC or Database Output Channel is an ordered list of what channels the output image layers are to be assigned to. These channels are the additional empty channels created previously during the data preparation process.
5. MTLFILE is the path and name of the metadata file which must be of the file type \*.L1G.
6. TIME is the scene acquisition or scene center time of the image, which should be contained within the MTL metadata file or pgs report. Be sure not to confuse the scene acquisition time with the product creation time. If no metadata is available, the scene acquisition time may be found with a query for the image on the USGS Earth Explorer website (<http://edcsns17.cr.usgs.gov/EarthExplorer/>).
7. TOARSCM is a more advanced function that is detailed in the Landry paper. For most applications “fixed” is selected.
8. PROTYPE allows the user to run either a test report or a complete process. It is a good practice to run a report first as this allows the program to check the MTL file to be sure it will work.

It is important to be mindful of syntax and notations when inputting the parameters. The **s toaretm** command may be input at any time to check parameters. Once the parameters have been entered and verified, the TOAR module may be executed by entering **r toaretm** at the prompt. A log file will automatically be created and saved to the working directory. This file should be saved and included in the final metadata that will accompany the corrected image.

#### *iv. Processing the Normalized Image*

The image has now undergone the Top of the Atmosphere Reflectance process. The corrected layers must now be exported. Recall the six unsigned 8-bit channels that were added to accommodate the new atmospherically corrected bands creates. These bands must be subset and exported to an ERDAS Imagine (\*.img) file format as the AGCC classification process is conducted using ERDAS Imagine software. Using File Utility in the PCI Geomatica Xspace module open the corrected image (Figure 5.2.6). Be sure that all twelve channels are occupied. Selecting channels 7 through 12, subset and save the output file in Erdas Imagine (\*.img) file format.

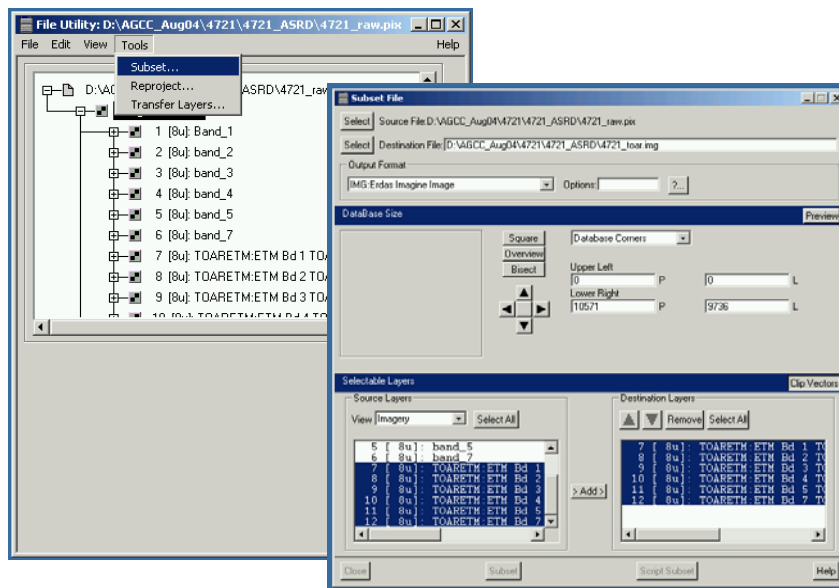


Figure 5.2. 6 Subsetting atmospherically corrected channels using PCI Geomatica Xspace

## 6. Classifying for AGCC

### 6.1 Hierarchical Classification Methods

#### i. Extraction of Known Features (Roads and Hydrography)

Recalling Figure 5.1.1, the first step in processing imagery according to the hierarchical classification approach is to extract all known or defined features. Where available, provincial vector data provided by ASRD for the AGCC allow road and hydrography features to be extracted from the imagery to reduce spectral confusion and better facilitate the classification process. If vector data is not available, the features must be extracted spectrally. The following is a detailed description of how to extract roads and hydrography from a satellite image using the GIS software package ArcGIS; and a remote sensing software package Erdas Imagine.

ArcGIS is a suite of GIS analyses and computation components. Using ArcGIS ArcMap, clip and reduce the vector data to decrease processing time.

Reducing Vector Data:

- open attribute table of the vector layer; select the following attributes accordingly (Figure 6.1.1):

- select the following attributes, for **roads**:

GRAVEL-2L, PAVED-DIV, PAVED-UNDIV-2L, PAVED-UNDIV-4L,  
WINTER-ROAD, TRUCK-TRAIL

- select all attributes, for **hydrography**:

**Note:** hydrography data should be in a seamless format. This will ensure that polygon topology is created and polygons such as rivers, lakes, bars, and islands will already be defined. If not, select the following and build polygon topology and create the polygons:

LAKE-INTERMITTENT, LAKE-PERRENNIAL, RESERVOIR, RIVER/STRM-  
MAJOR-L-BANK, RIVER/STRM-MAJOR-R-BANK, SAND BAR-WITH-RIVER

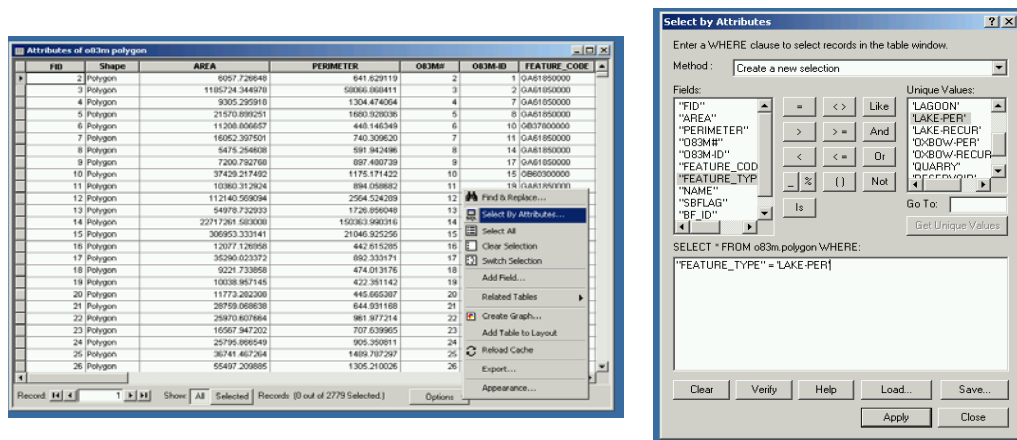


Figure 6.1. 1 Select desired attributes from vector layers using ArcGIS ArcMap

- export selected data as a new shapefile [right-click on the layer in your Table of Contents  
Display->Data->Export Data (Figure 6.1.2):

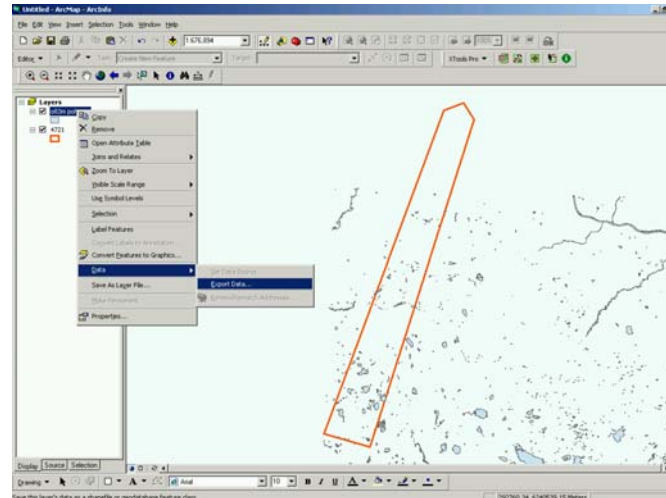


Figure 6.1. 2 Exporting selected data to a new shapefile

#### Clipping Vector Data:

To further reduce processing time, these data may be clipped to the area of interest. As seen in Figure 6.1.3., using an outline shapefile of the area of interest clip the vector layer (road and hydrography shapefiles).

- go to Tools -> GeoProcessing Wizard -> Clip one layer based on another
- the process will prompt the user for inputs
- this process will clip the vector layers based on the outline shapefiles

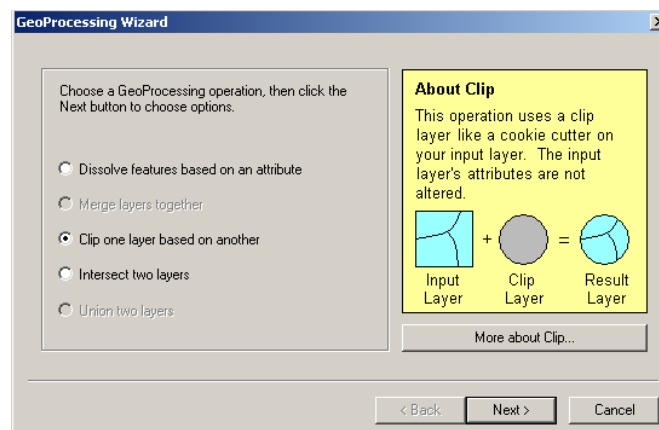


Figure 6.1. 3 Clipping one layer based on another using the GeoProcessing Wizard

- note that all new shapefiles created may need to be reprojected (Figure 6.1.4); using the Arc Toolbox component reproject the layers to UTM zone (10/11/12), Nad83; the process will prompt the user for inputs

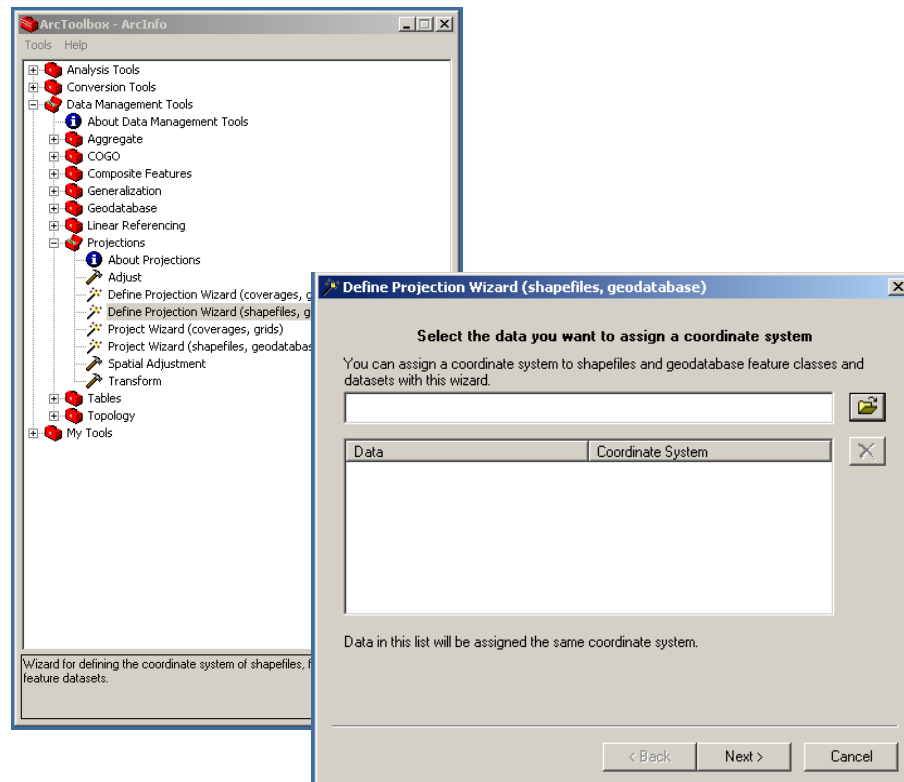


Figure 6.1. 4 Reprojecting the clipped data using Arc Toolbox

To extract the road and hydrography vectors from the satellite image, a mask must first be created to accommodate the data. Creating a blank thematic image in Erdas Imagine, themes may be recoded to represent the vector layers. This image will act as a mask to extract the road and hydrography data from the satellite image thereby reducing spectral confusion. A helpful tip is when extracting known features such as roads and hydrography is to first extract the hydrography layer, as features such as bridges will be overwritten if extracted first.

#### Creating a Blank Thematic Image (Mask):

- as seen in Figure 6.1.5 run a quick unsupervised classification with enough classes to accommodate one unclassified class and the number of vector layers that are to be extracted (ie. 2 roads and hydrography)

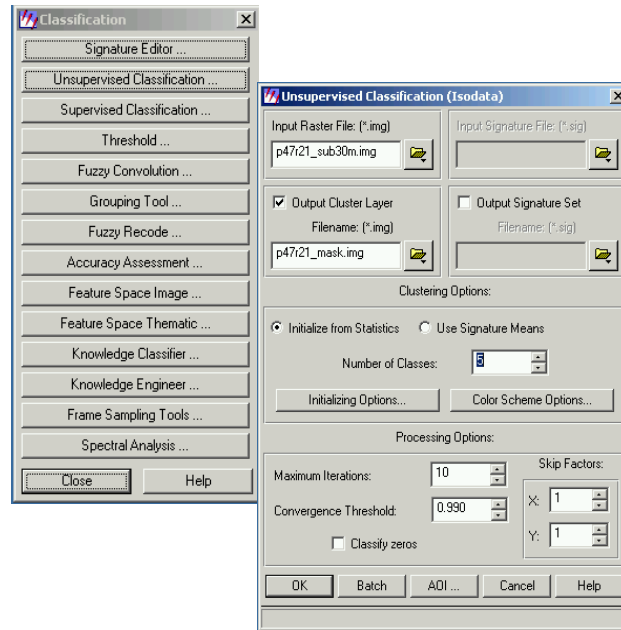


Figure 6.1. 5 Unsupervised classification using Erdas Imagine

- when extracting roads and hydrography, have at least a three class classification – it is good practice to create more classes than needed in the case that new known features are identified and may need to be add to the mask
- opening the classified image, RECODE all values to the highest number [ie. following the example of a three class classification, recode to 3] – this will allow you to have two blank bins to place the vector data in (Figure 6.1.6)
- once recoded be sure to save and close the image to ensure statistics are computed

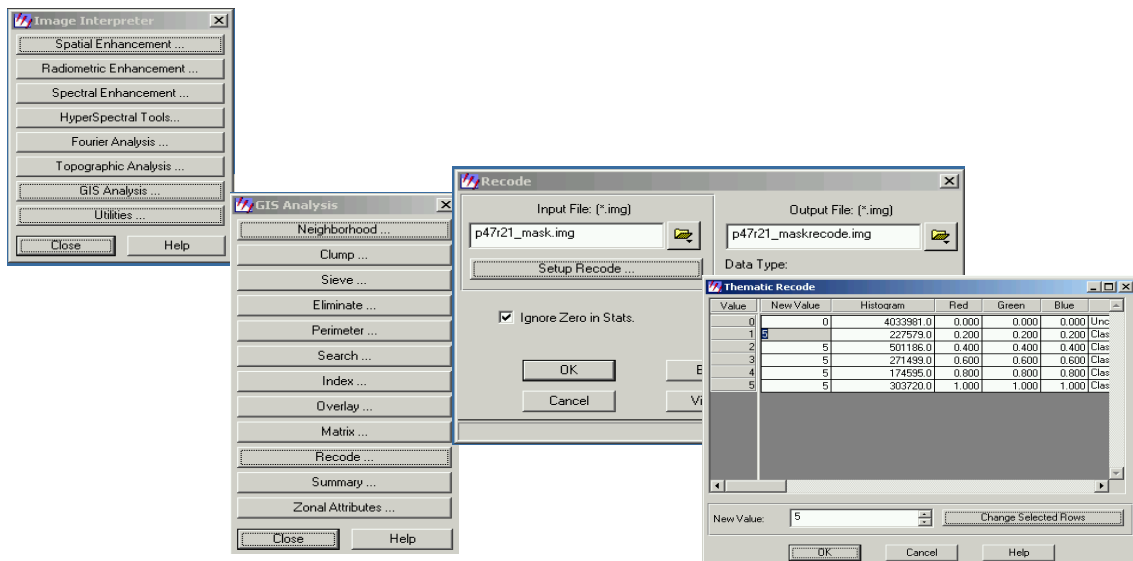


Figure 6.1. 6 Recoding the classification using Erdas Imagine

### Extracting the Vector Data:

- in the following order - open the blank image and vector files, respectively (**Note:** it is best to do so, one at a time to reduce processing time)
- items have to be opened in a certain order as Erdas may be finicky
- open the *Recode* table, *AOI* tools and *Vector* tools (these will only be available if an AOI layer and a vector layer is in the viewer) (Figure 6.1.7)

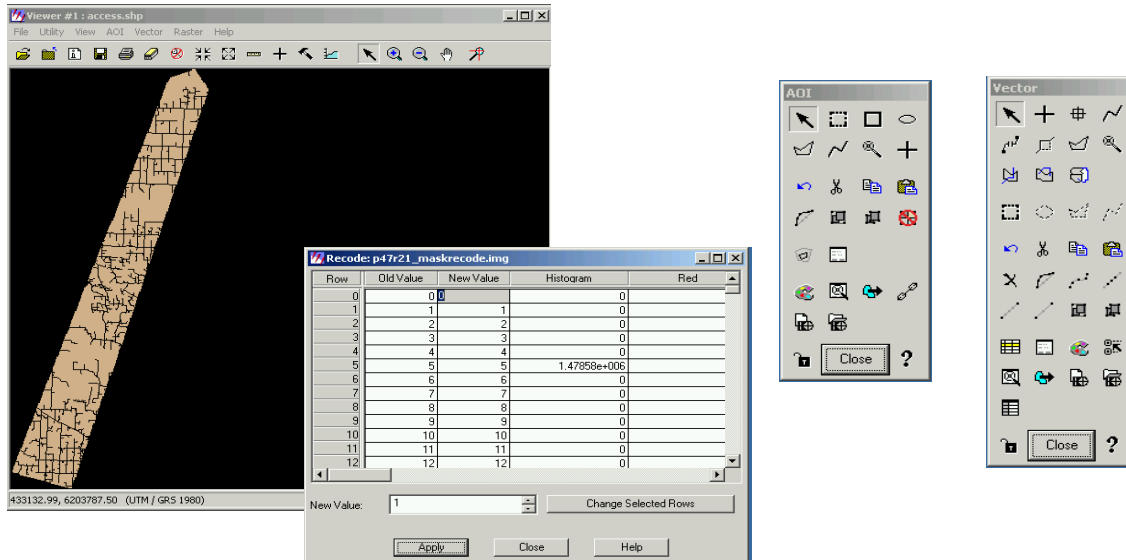


Figure 6.1. 7 Recoding via vector and AOI tools in Erdas Imagine

- select all the vectors using the vector tools, go to AOI(at the top of the view)->Copy to AOI
- this will take all the selected vectors and create AOIs (refrain from performing other functions on the workstation at this point, as the system will redraw and deselect the vectors)
- once all the vectors are copied to AOI, select all the AOIs using the AOI tools

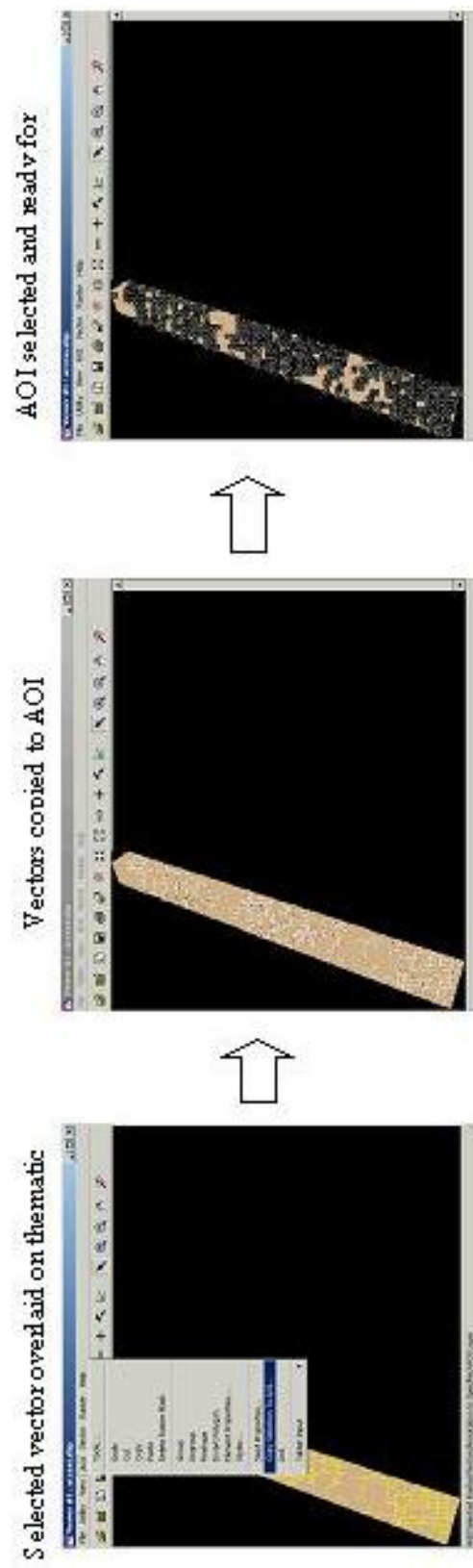


Figure 6.1. 8 Selecting and copying vectors to AOI; selecting AOIs



- in the recode table, Select All (rows), Change Selected to the desired class assigned to accommodate the data [ie. 1-water, 2-roads], hit Apply (Figure 6.1.9)
- this operation will change all selected vectors by the AOIs to the assigned bin (Ex. bin 1)
- it is a good practice to save frequently, after each vector layer is recoded
- once all vectors are recoded into the mask, the mask may be used to extract the road and hydrography layers from the imagery prior to classification

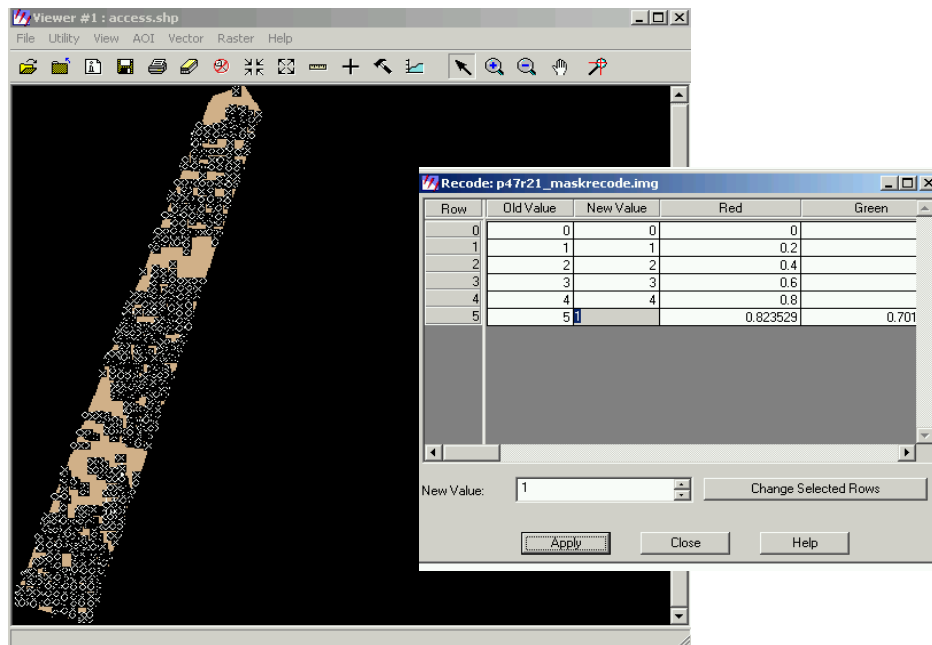


Figure 6.1. 9 Recoding selected AOI to appropriate land cover class

### Masking

Using the mask created, the road and hydrography layer may be extracted from the satellite image to decrease spectral confusion and better facilitate the classification process. As displayed in Figure 6.1.10, using Erdas Imagine in Interpreter -> Utilities -> Mask, recode the road and hydrography bins to zero. Input the satellite image file and input the mask file. To set up the recode function, right click to select all occupied rows with the exception of 0 -> Change Selected Rows: 0 -> give the zero row a value of 1. This process will result in a reduced working image for classification.

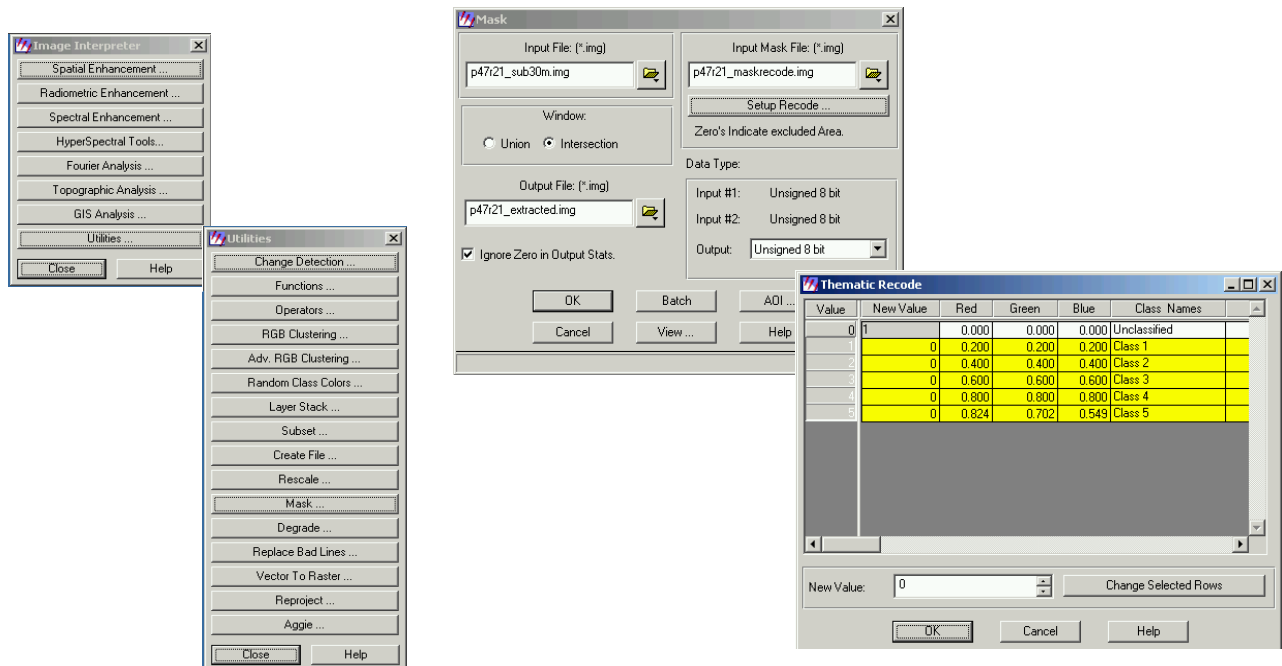


Figure 6.1. 10 Masking road and hydrography layer out of satellite image

## ii. Subset of Known Clearcuts and Burns

Known clearcuts and burns may be extracted in the same manner as the roads and hydrography to further reduce spectral confusion within the satellite image to be classified. Using data provided by the Forest Protection Division (FPD), a branch of ASRD these features may be identified and removed. However, as the clearcut and burned regions will contain various land cover types, a satellite scene containing only the clearcut and burn regions should be created. This may be accomplished by inverting the selection in the final step of masking. When masking the satellite image, the recode should be set to give the clearcuts and burns a value of 1 and all remaining classes a value of 0. This mask will result in a reduced working image of only the clearcut and burn regions that may be independently classified. Once classified, the resultant image may be recoded to coincide with the AGCC classification codes and added to the output thematic masks created at the various levels of the hierarchical classification to make up the preliminary classification.

### iii. Isolating and Extracting Ancillary Features

With the image ready for initial classification, a 50-class unsupervised classification may be performed to isolate ancillary features such as clouds, shadow, and residual water bodies, primarily lakes. Upon trial and error, 50 classes was determined to be sufficient in separating spectral ranges to accommodate the various classes of interest within the AGCC classification scheme. More than 50 classes may result in redundancies and increased processing time and fewer than 50 classes may result in spectral mixing. Figure 6.1.11 illustrated the parameters required for input to run a 50-class unsupervised classification.

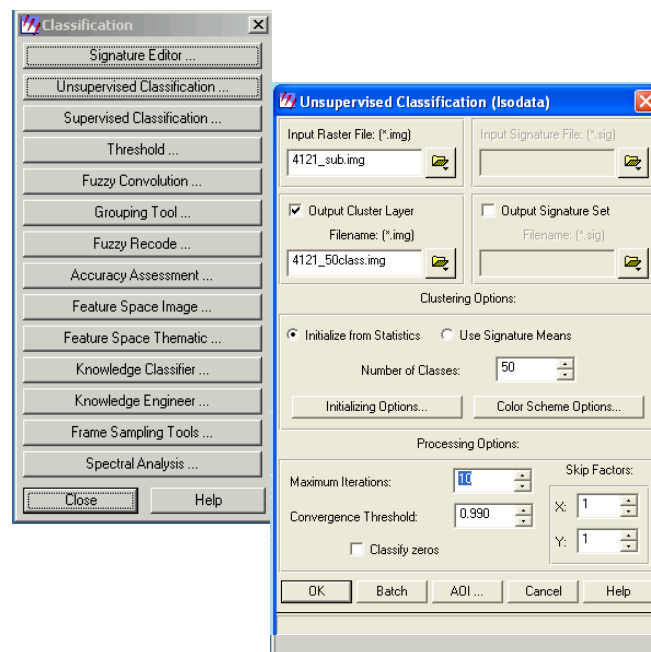


Figure 6.1. 11 Unsupervised classification using Erdas Imagine

The resultant unsupervised classification the analyst interprets the 50 classes produced using expert knowledge and resources to identify these ancillary features. The analyst also identifies features such as roads, industrial facilities, and agricultural land use. Using Table 6.1.1 as a guideline, various features may be identified depending on the band combination chosen when viewing the image. If the analyst is unable to extract these classes, it may be necessary to digitize and manually isolate these features. This may be very time consuming and will often take up the majority of the classification time. The end result is a single thematic layer consisting of all ancillary features. According to the hierarchical classification approach, this layer is extracted

from the satellite imagery using the same masking technique as applied to the extraction of the known features. The output image will result in an image with further reduced spectral confusion to be classified further to isolate the various land cover types present in the AGCC classification scheme.

Table 6.1. 1 Reflectance response of land cover types using various Landsat 7 band combinations.

Quinn, James W. "Band Combinations." April 5, 2005  
<http://web.pdx.edu/~emch/ip1/bandcombinations.html>

<b>R, G, B</b>	<b>Reflectance Response</b>
<b>4,3,2</b>	The standard "false color" composite. Vegetation appears in shades of red, urban areas are cyan blue, and soils vary from dark to light browns. Ice, snow and clouds are white or light cyan. Coniferous trees will appear darker red than hardwoods. This is a very popular band combination and is useful for vegetation studies, monitoring drainage and soil patterns and various stages of crop growth. Generally, deep red hues indicate broad leaf and/or healthier vegetation while lighter reds signify grasslands or sparsely vegetated areas. Densely populated urban areas are shown in light blue. This TM band combination gives results similar to traditional color infrared aerial photography.
<b>3,2,1</b>	The "natural color" band combination. Because the visible bands are used in this combination, ground features appear in colors similar to their appearance to the human visual system, healthy vegetation is green, recently cleared fields are very light, unhealthy vegetation is brown and yellow, roads are gray, and shorelines are white. This band combination provides the most water penetration and superior sediment and bathymetric information. It is also used for urban studies. Cleared and sparsely vegetated areas are not as easily detected here as in the 4 5 1 or 4 3 2 combination. Clouds and snow appear white and are difficult to distinguish. Also note that vegetation types are not as easily distinguished as the 4 5 1 combination. The 3 2 1 combination does not distinguish shallow water from soil as well as the 7 5 3 combination does.
<b>7,4,2</b>	This combination provides a "natural-like" rendition, while also penetrating atmospheric particles and smoke. Healthy vegetation will be a bright green and can saturate in seasons of heavy growth, grasslands will appear green, pink areas represent barren soil, oranges and browns represent sparsely vegetated areas. Dry vegetation will be orange and water will be blue. Sands, soils and minerals are highlighted in a multitude of colors. This band combination provides striking imagery for desert regions. It is useful for geological, agricultural and wetland studies. If there were any fires in this image they would appear red. This combination is used in the fire management applications for post-fire analysis of burned and non burned forested areas. Urban areas appear in varying shades of magenta. Grasslands appear as light green. The light-green spots inside the city indicate grassy land cover - parks, cemeteries, golf courses. Olive-green to bright-green hues normally indicate forested areas with coniferous forest being darker green than deciduous.

<b>4,5,1</b>	<p>Healthy vegetation appears in shades of reds, browns, oranges and yellows. Soils may be in greens and browns, urban features are white, cyan and gray, bright blue areas represent recently clearcut areas and reddish areas show new vegetation growth, probably sparse grasslands. Clear, deep water will be very dark in this combination, if the water is shallow or contains sediments it would appear as shades of lighter blue. For vegetation studies, the addition of the Mid-IR band increases sensitivity of detecting various stages of plant growth or stress; however care must be taken in interpretation if acquisition closely follows precipitation. Use of TM 4 and TM 5 shows high reflectance in healthy vegetated areas. It is helpful to compare flooded areas and red vegetated areas with the corresponding colors in the 3 2 1 combination to assure correct interpretation. This is not a good band combination for studying cultural features such as roads and runways.</p>
<b>4,5,3</b>	<p>This combination of near-IR (Band 4), mid-IR (Band 5) and red (Band 3) offers added definition of land-water boundaries and highlights subtle details not readily apparent in the visible bands alone. Inland lakes and streams can be located with greater precision when more infrared bands are used. With this band combination, vegetation type and condition show as variations of hues (browns, greens and oranges), as well as in tone. The 4,5,3 combination demonstrates moisture differences and is useful for analysis of soil and vegetation conditions. Generally, the wetter the soil, the darker it appears, because of the infrared absorption capabilities of water.</p>
<b>7,5,3</b>	<p>This band combination also provides a "natural-like" rendition while also penetrating atmospheric particles, smoke and haze. Vegetation appears in shades of dark and light green during the growing season, urban features are white, gray, cyan or purple, sands, soils and minerals appear in a variety of colors. The almost complete absorption of Mid-IR bands in water, ice and snow provides well defined coast lines and highlighted sources of water within the image. Snow and ice appear as dark blue, water is black or dark blue. Hot surfaces such as forest fires and volcano calderas saturate the Mid-IR bands and appear in shades of red or yellow. One particular application for this combination is monitoring forest fires. During seasons of little vegetation growth the 7 4 2 combination should be substituted. Flooded areas should look very dark blue or black, compared with the 3 2 1 combination in which shallow flooded regions appear gray and are difficult to distinguish.</p>
<b>5,4,3</b>	<p>Like the 4 5 1 combination, this combination provides the user with a great amount of information and color contrast. Healthy vegetation is bright green and soils are mauve. While the 7 4 2 combination includes TM 7, which has the geological information, the 5 4 3 combination uses TM 5 which has the most agricultural information. This combination is useful for vegetation studies, and is widely used in the areas of timber management and pest infestation.</p>
<b>5,4,1</b>	<p>This will look similar to the 7 4 2 combination in that healthy vegetation will be bright green, except the 5 4 1 combination is better for agricultural studies.</p>
<b>7,5,4</b>	<p>This combination involves no visible bands. It provides the best atmospheric penetration. Coast lines and shores are well defined. It may be used to find textural and moisture characteristics of soils. Vegetation appears blue. If the user prefers green vegetation, a 7 4 5 combination should be substituted. This band combination can be useful for geological studies.</p>
<b>5,3,1</b>	<p>This combination display topographic textures while 7 3 1 may display differences in rock types.</p>

#### *iv. Isolating Land Cover Types and Internal Classification*

Another unsupervised classification of 50 classes is run on the new output image void of all known and ancillary features. The resulting thematic image is used to distinguish between forested and non-forested classes. The analyst interprets the satellite image using expert knowledge and resources to identify the various land cover types (see Table 6.11). Clusters of confused spectral signatures are isolated and masked out for internal classification when necessary. This process allows for enhanced separability of confused or like spectral signatures, thus improving the classification. When available, point data is employed to verify the analyst's interpretation of the land cover.

This method of systematically reducing the dataset is beneficial in that it enhances the separability of less distinct classes. By successively masking layers out of the satellite image, sections of histogram values are omitted allowing remaining histogram values or pixels to be stretched further with each subsequent iteration of an unsupervised classification process. This technique reduces confusion and enhances the capabilities of performing better visual labeling of land cover classes. The labelling process is the most difficult and time-consuming portion of the hierarchical classification method. Using available field data and aerial photos will help to identify land cover types that the analyst may be uncertain or unfamiliar with. Modifying band combinations and image contrasts may also assist in the classification process.

#### *v. Reassembly*

Once the forest and non-forest layers have been classified and labelled, these layers are compiled with previously extracted layers to form the preliminary classification. This can be done by add all the separate layers together using the Modeler tool in Erdas Imagine (Figure 6.1.12). These layers include the road and hydrography layer, the clearcut and burn layer, the layer containing all ancillary features and the various layers of land cover types extracted.

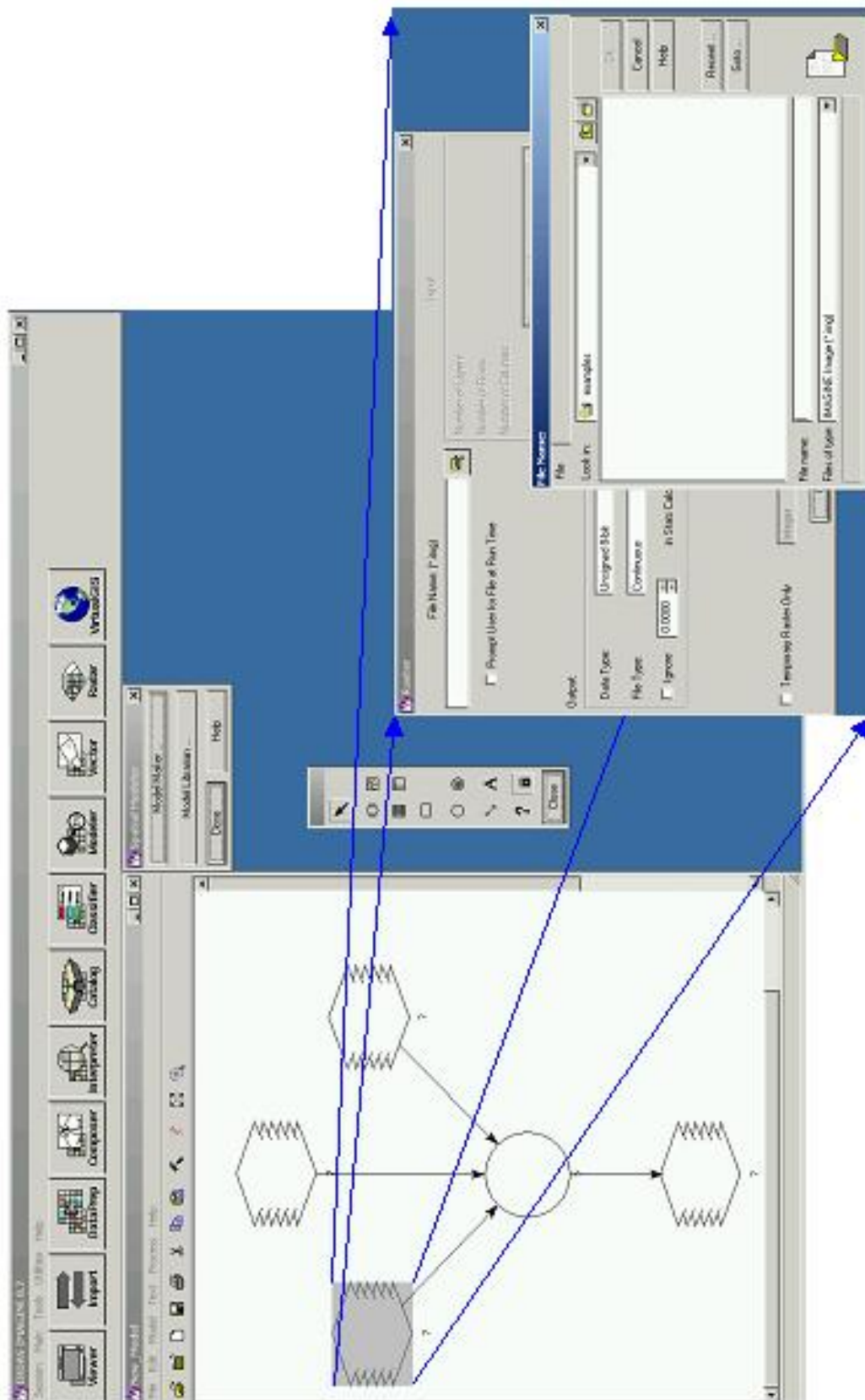


Figure 6.1. 12 Adding the separate layers using the Model Maker in Erdas Imagine (Note that the number of layers added may vary)

Prior to adding the various layers together, all thematic classification layers must be rescaled to contain pixel values that are consistent with the values of the corresponding AGCC class codes (Table 3.1.1). However, as clearcuts and burns are not necessarily land cover as much as a state of the land at a given time, these classes require special attention during the recode process.

For burns and clearcuts within the study area, the polygons obtained from the provincial Forest Protection Division (FPD) are used for the recoding of the preliminary classification. As the FPD datasets are very extensive, historical burns dating back to the 1800s are presented. A rule employed by ASRD indicates that burns greater than 20 years will be considered insignificant and will be classified according to land cover. However, burns or clearcuts within this timeframe that are no longer visible on the satellite image are also disregarded according to the AGCC methodology, as the information becomes obsolete if the feature is no longer visible. The following criteria are considered during the recode process. Note that the date of the burn or clearcut should be considered in reference to the date of image acquisition.

- Burns less than 3 years are recoded to 41

- Burns less than 5 years and greater than 3 years are recoded to 42

- Burns greater than 5 years are recoded to 43

- All wetland treed pixels are recoded to 45

- Clearcuts less than 3 years are recoded to 31

- Clearcuts less than 5 years and greater than 3 years are recoded to 32

- Clearcuts greater than 5 years are recoded to 33

For burns and clearcuts older than 5 years from the image acquisition date and that are still visible, the following rules are applied. Pixels containing islands or remnants retain regular AGCC land cover codes.

- Graminoid pixels are recoded to 41

- Shrub pixels are recoded to 42

- Treed pixels are recoded to 43

- All wetland treed pixels are recoded to 45

- Clearcuts with graminoid pixels are recoded to 31

- Clearcuts with shrub classification are recoded to 32

- Clearcuts with treed pixels are recoded to 33



Given the above procedures, the result of the AGCC hierarchical classification is a land cover map that represents a large area at a small scale. It incorporates forest species information as well as natural and anthropogenic features. With the preliminary classification complete, the image is ready for quality control.

## **7. Conducting fieldwork**

The field campaign is one of the most important phases of the AGCC project, as it familiarizes the research team with the landscape of which they are classifying. Direct involvement in the field will prove to be beneficial as it allows the team to appreciate how features or land cover types are represented on a satellite image. Team members will gain a sense of the affects of natural environmental conditions on the response of the sensor. This will aid in approaching atypical or unfamiliar spectral signatures encountered during the classification process. Furthermore an understanding or awareness is attained, of which features are distinguishable and relevant according to the capabilities of the sensor. But, above all the field campaign provides data that can be used to assess the accuracy of the classifications produced.

### **7.1 Planning for fieldwork**

To carry out a field campaign efficiently, much preparation is required prior to going out into the field. A major consideration when planning fieldwork is the approach taken in gathering the data. The area to be sampled for the AGCC project is quite extensive, as such aerial survey was the primary method employed in gathering field data to allow for the maximum data collected in the minimum amount of time. Planning for fieldwork should begin with the selection of field sites to be visited.

#### *i. Selection of field sites*

Selection of field sites is the most important process of fieldwork. The quality of the field sites selected will reflect directly on the quality of the data collected. In selecting field sites, it is ideal to have a preliminary classification already completed. This will allow for the identification of problem areas of high spectral mixing that require further analysis. As well, the accuracy of interpretations made by the analyst during the classification process may be assessed. As such, the following criteria should be employed when selecting field sites.

1. Locate all the major vegetation types interpreted within the classification (Ex. trees, shrubs, wetlands, grasslands, etc.) to be sure that all land cover types of interest are identified.
2. Locate complex problem areas of high spectral confusion for which distinctive vegetation classes were indistinguishable.
3. Locate pure vegetation sites to establish a baseline spectral signature for verification of major vegetation types.
4. Ensure accessibility to sites selected when conducting ground work. When conducting helicopter surveys, consideration should be given to fuel caches and flight time. The sample areas should be located at 60 - 100 meters from roads, well pads, or other anthropogenic features in order to avoid edge effects. The sample area should be of significant size from the imagery to enable easy identification in the field.

#### *ii. Planning the route*

A transportation route should be established before going out into the field to promote efficiency and to avoid getting lost. Accessibility to sites must be acknowledged, particularly when conducting groundwork. Driving or hiking under adverse conditions may be necessary. As such, four-wheel drive vehicles are preferred under such circumstances.

Special permission may be necessary to access private land areas. Information regarding this issue may be obtained from the provincial office in charge of recognizance.

#### *iii. Preparation of field maps*

A list of GPS coordinates should be generated for ready use in a GIS and to be displayed on hardcopy maps to be taken into the field. The simplest way to compile a list of GPS points is to create a shapefile when selecting field points directly from the classified image or satellite image. The geographic coordinates may then be exported from the shapefile. It is handy to have digital overlays of the satellite image, classified image, and field points for easy navigation and visual display in a portable GIS such as ArcPad while in the field.

Hardcopy maps of satellite images at 1:50,000 or 1: 75,000 scale should be created to directly record observations and notes made while in the field. For individual training sites, it is ideal to obtain aerial photos where available for more detailed analysis. Access information including roads, cutlines, seismic lines, and hydrography should be displayed on the maps to aid in orientation and allow for easy navigation. Figure 7.1.1 is an example of the field maps produced for the 2004 AGCC field campaign. This map was produced using ArcMap.

GP D1 Map #7

AGCC September 2004

1:75,000

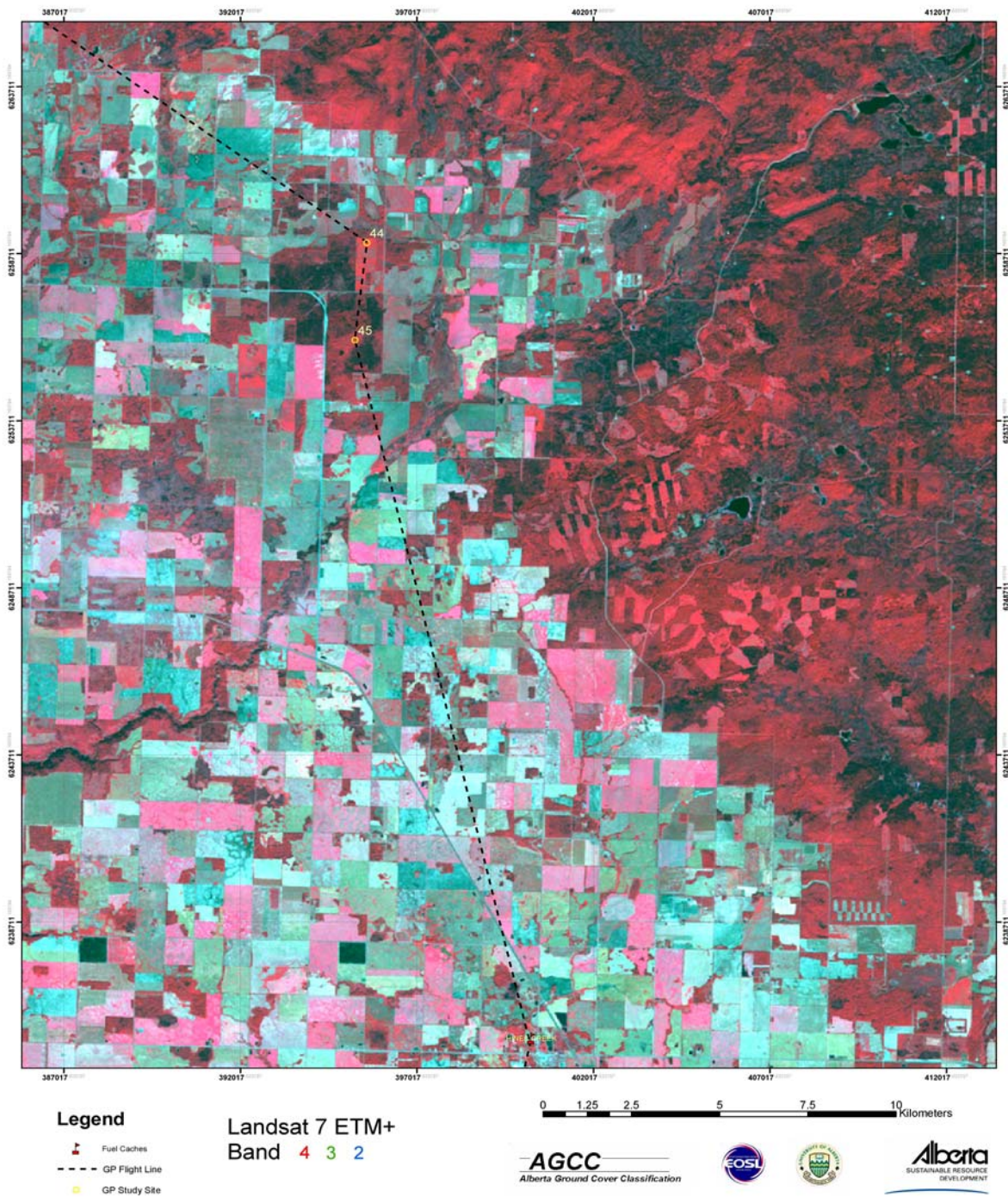


Figure 7.1. 1 Field map containing the satellite image, field points, and access information

## **7.2 Preparation of field crew**

The field crew should consist of individuals trained in vegetation identification particularly in a boreal environment, as well as First Aid. A minimum of two people is required to conduct field survey. All members of the field crew should have a working knowledge of the project and understand the procedures of data collection. Crewmembers will never be asked to work alone, as fieldwork may involve traveling and lodging in remote areas under adverse conditions. As such field safety must be addressed prior to going into the field.

### *i. Field Safety*

Prevention is the key to safety. All field members should approach each day with:

- attitudes and behaviors conducive to safe field practices
- good footwear, preferably hiking boots or sturdy running shoes if not boots; open-toed shoes are undesirable. Rubber boots may also be considered dependant on the conditions.
- clothing appropriate for the complete 'normal' range of local weather. May include long trousers, sweater, and wind/rain gear.
- sun protection: hat, sun glasses, and sun screen
- extra food and water when appropriate
- medication for chronic illnesses such as diabetes, asthma etc.
- mosquito and fly repellent

It is important to inform crewmembers of any conditions that might require emergency medical attention, such as bee sting allergy. Most fatal field accidents are related to vehicle travel. As such the following points should be noted:

- only approved/authorized drivers may drive field vehicles
- field vehicles may not be used for recreational purposes
- drivers will use common sense and drive defensively, thus they will stop if too tired to continue safely
- all Highway Traffic Act and local by-law, rules and regulations must be followed

### *ii. Field Equipment*

To properly conduct fieldwork, appropriate tools must be available for use. All members of the field crew should be familiar in the use of the equipment. It is important to ensure that all equipment is accounted for at the end of each sampling session as well as at the end of the day. The following is a list of equipment useful in conducting vegetation surveys.

#### Equipment List:

1-30 meter measuring tape  
1-Compass  
2-GPS units  
1-Digital Camera and accessories  
1-Laptop  
Plenty of flagging tape  
Extra batteries for the GPS  
First Aid Kit  
Waterproof field forms  
Writing instruments  
Calculator  
Bear Spray  
Mosquito repellent

### **7.3 Description of how to conduct AGCC field calls**

Field calls must be made in a manner that will help to identify the various land cover classes present in the AGCC classification scheme within a given satellite scene. The goal is to scale what is observed on the ground to what is observed on a 30m resolution satellite image. It is vital to record all information gathered in the field, whether it is a formal field call or simply observations made during the excursion. The AGCC field data is recorded on field forms provided in Appendix A. Given that all preparations are complete; the following is a description of how to conduct field calls to accommodate the AGCC classification scheme.

1. Navigate to the proposed site using GPS and map/photo information. While walking into each location mark the trail using flagging tape. This will aid navigation out of the sampling site for safety and efficiency. However, it is important to remove the flagging tap on the way out, as the plastic material could be harmful to wildlife and the environment.
2. Each location should be investigated before measuring commences. Identify the center point and explore up to 100 meters in various directions from the center. During the excursion evaluate whether the initial geographical position is ideal for the location of the center of the polygon. If not, move the location and ensure the new position is measured with the GPS and documented on the field form.
3. Once the location is accepted, there are 3 different field activities to be performed at each site:

a) Photography

Photographs should be taken at each field site as a visual record representative of the land cover present at the given sampling point. Using a compass, the cardinal direction of the photos should be documented. More than one photo should be taken to describe a complex site. For all treed stands oblique angle photographs should be taken at 45 degrees to the horizontal or ground. These photographs may act as indicators of crown closure of the stand when the sites are reexamined at a later date.

b) Crown Closure

A densitometer is a device used to determine crown closure by estimating the amount of overstory foliage in the crown. It uses permanent points or transects to calculate the relative density of plant growth above and surrounding the sampling point. The densitometer will have a concave mirrored surface with 37 grid intersections forming 24 squares. To take readings, hold level at elbow-height at a distance of 12 to 18". Imagine 4 uniformly spaced dots in each square of the grid and count the number of dots covered by vegetation. Do not break the squares into 4 imaginary squares and try to decide if each one has greater or less than 50% cover. The 4 squares within a square method introduce bias because individuals tend to decide closed calls more frequently in one direction than another. Deciding whether or not an imaginary dot is covered is a simple objective decision, and is less susceptible to bias. Avoid counting trunks as cover when possible, particularly when they make up a substantial portion of the cover. Step away from the trunk if necessary. The number of dots counted may be converted to percent canopy using the following formulae:

$$\text{Percent canopy closure} = (\# \text{ dots} * 1.04)$$

c) Data Collection

Site characteristics are recorded and entered into the AGCC field forms to give a representation of the conditions during the time of data collection (see appendix). Once the site characteristics have been recorded, AVI and AGCC calls are made. AVI calls represent 100% of the vegetation at up to three levels at a sample site. For the purposes of AGCC, a maximum of two levels are

considered. As such, AVI considers treed vegetation as well as ground vegetation that may not be detected by the sensor depending on the canopy closure. AGCC is different in that the calls represent all the vegetation present in the sample site as a whole. The combined signatures of all the vegetation are detected by the sensor and expressed by a single value in each spectral band. These data make up the spectral signature of the training site. The challenge with the AGCC call is to make a single vegetation call for an area encompassing a 30m x 30m pixel. It is useful to consider both AVI and AGCC field calls to gain an understanding of all the components that contribute to a spectral signature. Though the sensor is essentially detecting surface reflectance from the canopy, vegetation from below the canopy can ultimately affect the reflectance spectra as a whole. As anomalous spectral signatures are often observed, using all field data gathered, one can make decisions on how to label pixels of a given spectral signature and account for any irregular characteristics.

#### *Alberta Vegetation Inventory (AVI) Calls*

Note that for the purposes of AGCC, the AVI call will look at a maximum of 2 levels or stories in a given stand. The moisture regime, crown closure, height, species, and shrubs are measured for each level. It should be noted that certain species could be used as indicators for wet or dry conditions. For example, balsam poplar, devil's club, and horsetail are indicative of wet mesic to wet sites. Pine generally occurs under drier conditions. Each level will add up to 100% (Appendix B). A comprehensive description of the AVI calling criteria can be found in the AVI Standards Manual, which may be obtained from Alberta Sustainable Resource Development, Resource Data Division. In addition to these individual level measurements, another overall account of tree, shrub, and herbaceous species will also be taken for the sample site as a whole that will add up to 100%. In this case, the objective is to describe the percent coverage of each vegetation type (i.e. trees, shrubs, herbaceous) over the entire sample site. Though on the ground, this call is taken with the perspective of looking down onto the canopy through any open spaces within that may expose other levels of vegetation. This call will aid in making the appropriate AGCC call.

#### *AGCC Calls*

Recall that AGCC field calls accounts for the entire coverage of a single 30m x 30m pixel. As such, all of the vegetation types in the field site must be considered. Based on the combination of vegetation types within a 30m x 30m pixel, the true AGCC class as well as the effect of the various vegetation types on the reflectance of the site must be determined. For example, if a site

were predominately open deciduous based on canopy cover, but had a thick leafy shrub understorey, it is likely that the spectral reflectance may result in a closed deciduous call during interpretation on a satellite image. However in such a case, this thick understorey would be noted and the AGCC code recorded would likely be open deciduous. In doing this, the true vegetation class is identified but the effect of the understorey will be accounted for should a discrepancy occur between the field call and the satellite interpretation. All calls should be made based on the AGCC classification scheme (recall Table 3.1.1).

The full process of making field calls: entering a location, exploring, gathering measurements, and making a call should take less than 30 minutes to complete depending on the size of the field crew. A protocol should be determined ahead of time so that duties are distributed in a way that maximizes efficiency. Data should be collected during travel in the vehicle as well. Any information collected during field reconnaissance will help with the classification process. Information regarding the characteristics of the understory, maturity, the number of canopy levels, the moisture regime at the site, the dominant species, or any details that can aid in the classification and/or may affect spectral signatures of the vegetation types should be recorded. Be sure to mark in legible script features about the landscape that are identifiable on the hard copy maps. These additional features will help in the initial classification and in the quality control process of the final classification. Make sure that the forms are completed and legible before leaving the plot. Also, be sure to account for all equipment.

#### **7.4 Processing Field Data**

Once all the field data has been collected, the data should be entered into a spreadsheet to be assessable in digital format to be applied in a GIS. All information recorded, including comments and observations made in transit, should be entered. A separate table can be made containing a unique identifying code, the spatial coordinates, and the AGCC class code. This table may be applied to a GIS to overlay on the satellite image or classification. The data should be separated into two groups, one to be used for training during the classification process and the other for accuracy assessment. Ideally, 25% of the dataset should be applied to training and 75% should be applied to the accuracy assessment.



## 8. Validation of AGCC data sets

### 8.1 Descriptions of how to conduct an accuracy assessment on the AGCC Level 3.

Obtaining accurate ground point information serves the purpose not only of assisting in the classification of the satellite image, but in verification of the classified image. Thus, it is necessary to quantify errors in the classification to clarify the limitations of the final product. Validation of the AGCC data sets consists of using ground data acquired from fieldwork to produce and conduct an error analysis of the classification. To avoid biasing the validation, the collected data should first have been divided into two groups: 25% for training of the classification, and 75% for accuracy assessment. Two programs, created by Tomas de Camino, were designed for the assessment process: *Image Analysis* and *Accuracy*. The classified image is inputted to *Image Analysis* and, using the ground point data for assessment, a table of AGCC codes is derived from the corresponding point location on the classification. This table is imported into the *Accuracy* program to assemble an error matrix to identify class confusion and compute overall accuracy. The results are used to amend the classification. The following is a brief outline of the processes.

#### Data Preparation

1. Convert to Generic Binary (BSQ) with header file
  - a. ERDAS export function
2. Obtain ground point set and reduce to X, Y, and AGCC Code
  - a. Save as tab delimited text (\*.txt)

#### *Image Analysis*

3. Add points and Generic Binary file (not header)
4. Buffer points by 3 pixels (ie. 1ha)
5. Click 'Get Pixel Value' button
  - a. 'Center Value' and 'Buffer' values will appear
6. Save the analysis point table and open in Excel
  - a. Inspect for erroneous points, extract and note
  - b. Reduce to AGCC code and Center Value
  - c. Save as CSV (\*.csv)

#### *Accuracy*

7. Only 2 columns required AGCC and the Center Value

#### *Revision*

8. Modify classification according to results

### i. Data Preparation

To identify corresponding AGCC class codes and ground points, the classified image must first be converted to a format compatible with the *Image Analysis* program. As shown in Figure 8.1.1, using the export feature of ERDAS Imagine, the image is exported to a Generic Binary - Band Sequential (BSQ) with a header file:

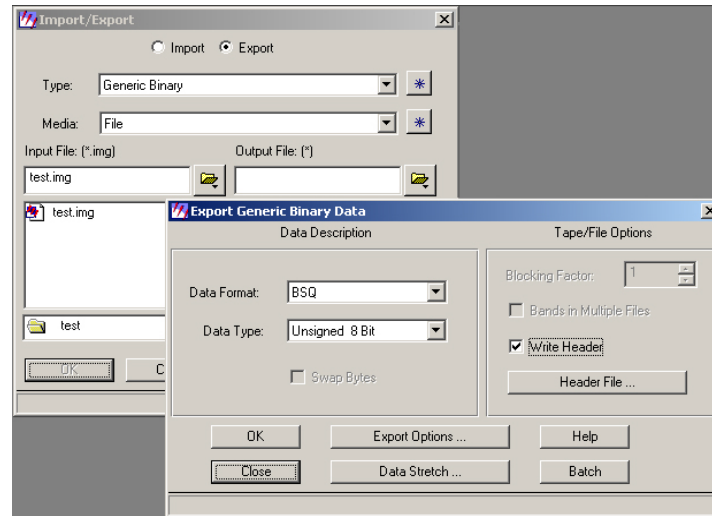


Figure 8.1. 1Exporting the classified image to generic binary using the export function in Erdas Imagine

The ground point data must be saved in tab delimited text format and may only contain three fields in the following order: X, Y, AGCC. Thus, some data manipulation may be required within the Excel environment to reduce and export only the necessary fields from the data table. Be sure to save the output as a tab delimited (\*.txt) file. Once the data has been reduced to contain only the necessary fields, erroneous points such as the ones that are highlighted in Figure 8.1.2 can be removed. Before proceeding, the data point values must be transformed to be consistent with the classified image in terms of classification detail. Calls made on the ground are often more specific than those differentiated during the classification. In order to correctly assemble the error matrix, classes that exist in the point data must also exist in the classified image. Discrepancies often occur as a result of varied levels to which the image is classified. For example, an AGCC Level 4 call of 51, 52, or 53 representing specific closed coniferous classes may have to be grouped into class 54, representing undifferentiated conifer in the classification likely of an AGCC Level 2 or 3 classification. Without this modification, accuracy results will be negatively represented.

	A	B	C	D	E	F
1	UTM_X	UTM_Y	AGCC			
2	487111.99090999900	6191828.00363000000	21			
3	487194.85864599900	6182298.15187000000	21			
4	472568.61378199900	6170945.20708999000	21			
5	470911.24678699900	6171773.88191000000	21			
6	470828.37799000000	6172561.13136999000	21			
7	471905.66581999900	6170157.95799000000	21			
8	470289.73673399900	6169204.96643999000	21			
9	482383.33583200000	6095978.77326999000	155			
10	475192.99445200000	6106211.43938000000	155			
11	478889.42236999900	6104751.28548000000	151			
12	452207.39206300000	6107990.96102000000	55			
13	447575.02189099900	6109199.60995000000	251			
14	439771.86176399900	6126993.80946000000	155			
15	451021.03650500000	6115490.41359999000	0			
16	450766.82018799900	6115490.41504000000	0			
17	449177.95393800000	6123116.97807000000	0			
18	387735.46020999900	6128967.11366999000	55			
19	405578.48315799900	6169886.34028000000	57			
20	409521.61915699900	6171229.17585000000	55			
21	411174.68989300000	6169761.49139000000	55			
22	409575.76253100000	6174025.72707000000	55			
23	398290.21496700000	6151787.66129999000	0			
24	395104.65387500000	6150574.53488999000	58			
25	394415.91265399900	6152700.09814000000	155			
26	394550.24281099900	6173902.16344999000	0			
27	419931.49899499900	6221829.25155999000	56			
28	420051.06384999900	6219580.74416999000	155			
29	423846.74632500000	6217962.51085999000	251			
30	424745.36211099900	6222225.19759000000	57			

Figure 8.1. 2 Ground point data fields in Excel

## ii. Image Analysis

After the image has been exported to a BSQ generic binary image and the ground data has been saved as tab delimited text, the two datasets can be brought into *Image Analysis* to generate a confusion matrix. Load the data by opening the binary image and import the corresponding tab delimited text file as seen in Figure 8.1.3. A table will appear and provide the option of buffering the point. As the AGCC project calls for a minimum mapping unit of 1ha, a buffer radius of 3 pixels should be employed. Pixel values surrounding the center of the ground point can be used to judge whether a point is classified correctly. Clicking on the “Get Pixel Value” button will then display the pixel values within the desired buffer.

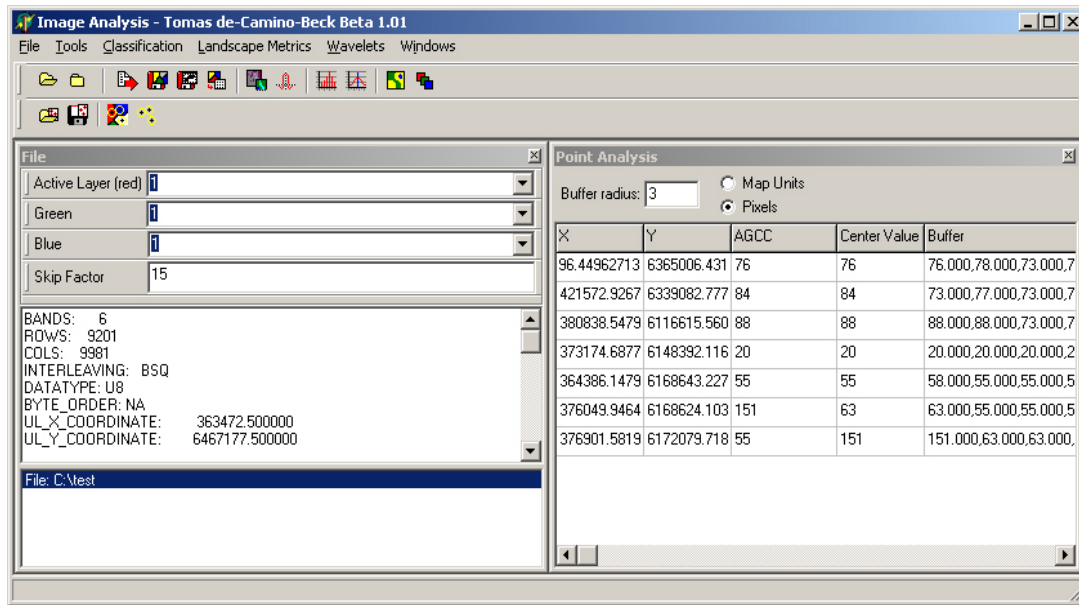


Figure 8.1. 3 Image Analysis interface

The point analysis table can now be exported to Excel for further analysis of the points and their surrounding pixels. The mode and incidence of the buffered values should be calculated and the excel file saved for later applications. These values will be used later to identify inappropriate point values that may affect accuracy assessment. Before saving for use with *Accuracy*, the table may need to be manually edited to remove relicts of the file format including commas and quotation marks that interfere with column separation. The fields of interest are the “AGCC” and the “Center Value” field. Save only these fields, without the headings, and any edits, to a new worksheet and export as a comma separated value file (\*.csv) to be applied to the *Accuracy* program.

### iii. Accuracy

The \*.csv data file is used by the *Accuracy* program to create a confusion matrix, and calculate statistics. It should be noted that this program may only be used on a workstation that has a D drive, or has a hard drive assigned as D. A confusion matrix is an array of class codes comparing the ground data and the classification data. It is an effective visual and statistical tool used to assist in identifying trends in the error and class confusion. For example, open deciduous classes can be confused with shrub classes and similarly, agricultural classes may be confused

with grass or herbaceous classes. Using the ground data, these areas may be identified within the image and modifications may be made to improve the correctness of the classification.

Importing the \*.csv file into *Accuracy* and pressing the Kappa ( $\kappa$ ) button will yield statistics calculated to describe the class accuracy and the overall accuracy. The overall accuracy is the simplest descriptive accuracy statistic and is calculated by dividing the total correct, or matching, values by the total number of points in the matrix. However, this value may be unrepresentative of where and to what degree errors are occurring. In the following example, 146 points were used in an accuracy matrix and an overall accuracy of 82.2% was calculated (Figure 8.1.4). However, with closer analysis, it can be determined that a large portion of the errors are primarily occurring between open and closed deciduous and with mixed classes. Further amending or reclassification of these classes could lead to improved results.

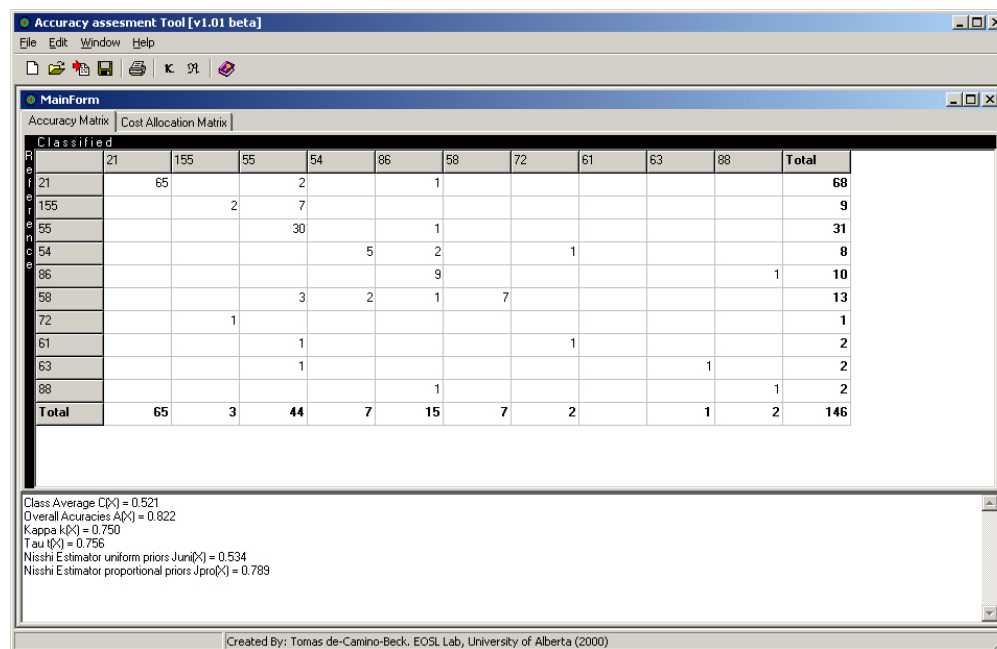


Figure 8.1. 4 Accuracy interface, confusion matrix

Producer accuracy, or omission error per class can be calculated by dividing the correct number of points by the total number of ground data points of each individual class. Producer's accuracy measures the probability of a reference point being correctly classified. User's accuracy, or commission error, is calculated when the total number of correct points in the class is divided by the total number of classified points in that category. User's accuracy is indicative of the probability that a pixel classified on the image actually represents that point on the ground. AGCC level three standards call for an acceptable overall accuracy of >80%.

#### *iv. Revision*

Recalling the mode and incidence values calculated from the *Image Analysis* results, the weight of correctness for each buffered ground point might be assessed. The incidence is the number of times the AGCC code of the ground point appears within the buffer pixels. A further measure of the accuracy is the weight of a point in comparison to these buffer values. The incidence is weighted with a significance value of 1 or 0 depending on whether the AGCC code appeared 10 or more times within the buffer. This is best calculated manually as it may be recoded should there be notable class confusions. For example, if the calculated value for the mode has an AGCC code of 72 (grass) but the incidence has a value of 21 for agriculture; there is reasonable confusion in the area to decide that the point is not incorrect, and thus a weighted value of 1 is given. Using the sum of the significance column and the total number of points used, an overall accuracy can be measured. Using the information from both the buffers and the error matrix, the classifications can then be corrected accordingly.

It should be noted that all AGCC classifications are considered preliminary classifications, even when the accuracy assessment is carried out. All classifications are submitted to Alberta Sustainable Resource Development and the Canadian Forest Service for further evaluation and auditing. Aerial photos and expert knowledge are used to assess the quality of the classifications and to identify any modifications that need to be made.

## **9. Exporting AGCC data sets**

### **9.1 Descriptions of steps and algorithms used to transform from raster to vector the information extracted from the TM imagery.**

#### **Raster to Vector Conversion**

A vector file of the classified image may be necessary when delivering data to various groups with limited software access. Open the ArcToolbox module and navigate to the *Raster to Polygon* function (Figure 9.1.1).

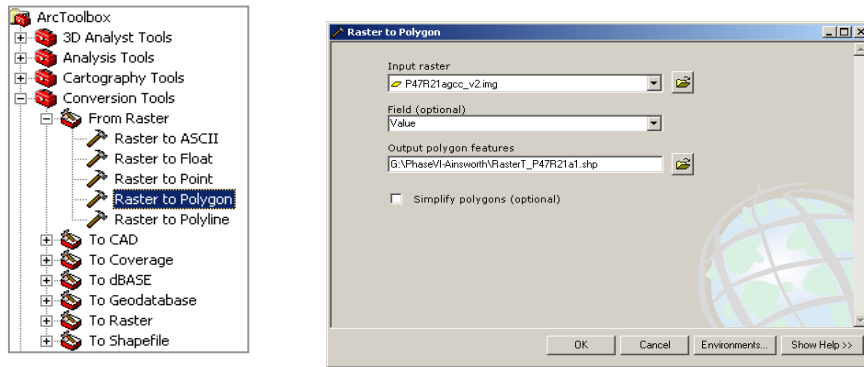


Figure 9.1. 1 Converting raster to vector using ArcGIS ArcToolbox

Opening the *Raster to Polygon* function, specify the input raster and output polygons. Execute the process by clicking the *OK* button. You can choose to simplify the resultant polygons by checking that option in the conversion interface. Simplification of linear features or polygon boundaries is a type of generalization operation that removes extraneous bends and small intrusions and extrusions from a line or polygon boundary without destroying its essential shape. However, this will modify the data and will alter any further analyses performed and is therefore not recommended.

## 9.2. Description of steps and algorithms used to produce a map using a given minimum mapping unit.

### *i Area Calculation*

It may be desirable to eliminate all polygons below a certain area, that is, apply a minimum mapping unit. To do so, add the shapefile to a layout in ArcMap and open the attribute table. If there is no area field it will have to be added (Figure 9.2.1).

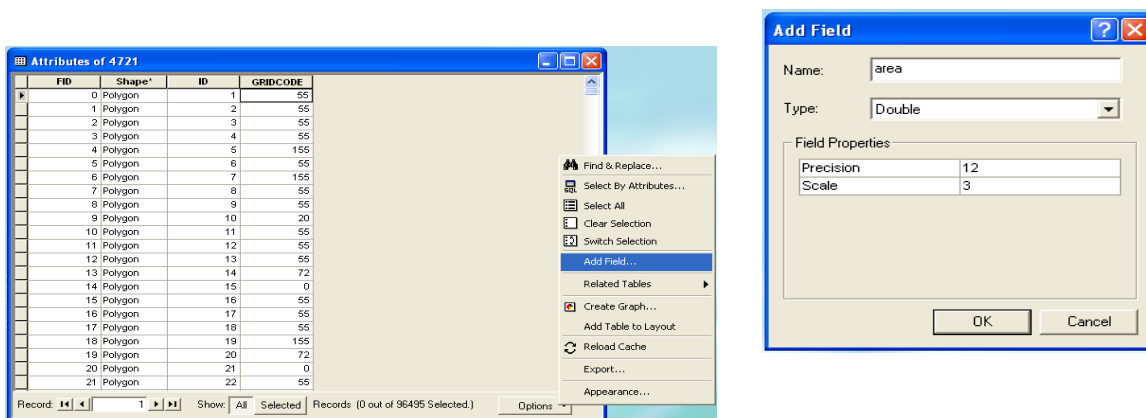


Figure 9.2. 1 Adding an area field

Once an area field is added, the area must be calculated. Making sure that the “Advanced” box is checked on; enter the Help file. Scroll down to locate the area calculation VBA script code. Once located, copy and paste this code into equation box to calculate the area. Lastly be sure to enter “output” to assign the output to the area field (Figure 9.2.2).

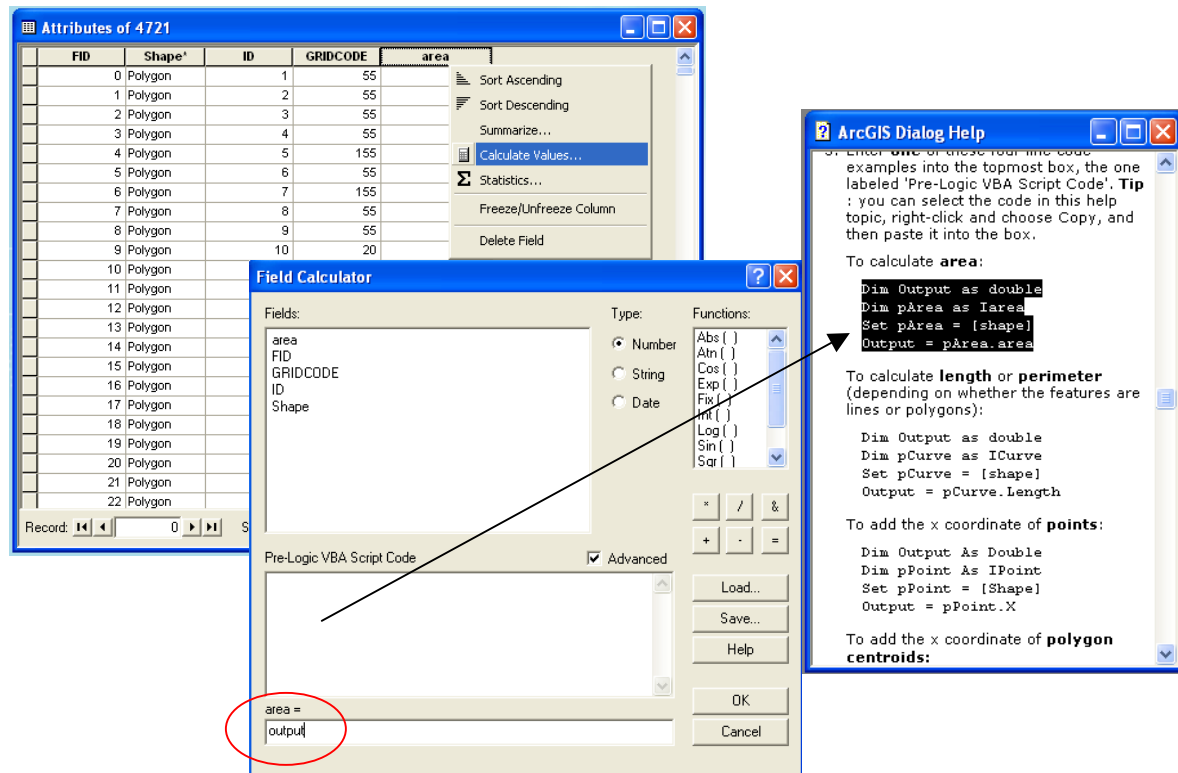


Figure 9.2. 2 Calculating area using the field calculator

## ii. Minimum Mapping Unit

Now that the shapefile contains an area field, it is possible to select and eliminate these polygons and incorporate into larger neighbouring polygons. In the attribute table options and choose *Select by Attributes*. An SQL window will open for the user to enter a formula (Figure 9.2.3). Enter a formula, similar to that illustrated below to select polygons of a desired size, for example 1ha, where 1ha = 10, 000 m<sup>2</sup>.



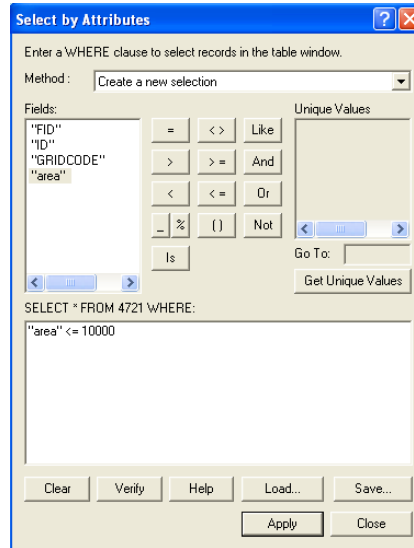


Figure 9.2. 3 Selecting by attribute, polygons of a minimum mapping unit of 1ha (10, 000 m<sup>2</sup>) or less

Once the undesired polygons are selected, the Eliminate operation may be applied to incorporate the smaller undesired polygons in with larger neighbouring polygons. In ArcToolbox, navigate to the Eliminate function (Figure 9.2.4). Be sure to uncheck the option to eliminate polygons by borders, as the intention is to eliminate polygons based on the calculated area, which is the default. Once the selected polygons of 1ha or smaller have been eliminated, the minimum mapping unit has been successfully applied. Note that area should be recalculated to ensure that the area of the eliminated polygons are incorporated into the larger neighbouring polygons.

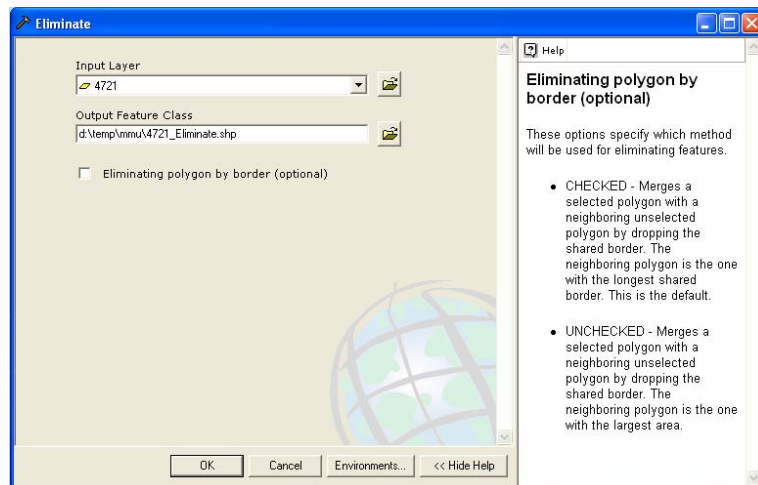


Figure 9.2. 4 Eliminating selected polygons

## 10. Metadata for AGCC

### 10.1 Description of the metadata standards used by the AGCC

The metadata standards employed by the AGCC must comply with international and provincial standards of metadata production to ensure accessibility to users. The current standard is defined by the US Federal Geospatial Data Committee (FGDC) and includes the following components:

- Identification Information
- Data Quality Information
- Spatial Organization Information
- Spatial Reference Information
- Entity and Attribute Information
- Distribution Information
- Metadata Reference Information

The FGDC standards were derived from the user's perspective of relevant information needed to assess the usability of a given dataset. Furthermore, it allows the producer and user to keep track of the source of the data as well as any modifications made to the product and document status.

### 10.2 Description of how to create the metadata.

Using ArcCatalog from the ESRI ArcGIS suite of software, the Metalite application may be used to create FGDC standard metadata (Figure 10.2.1).

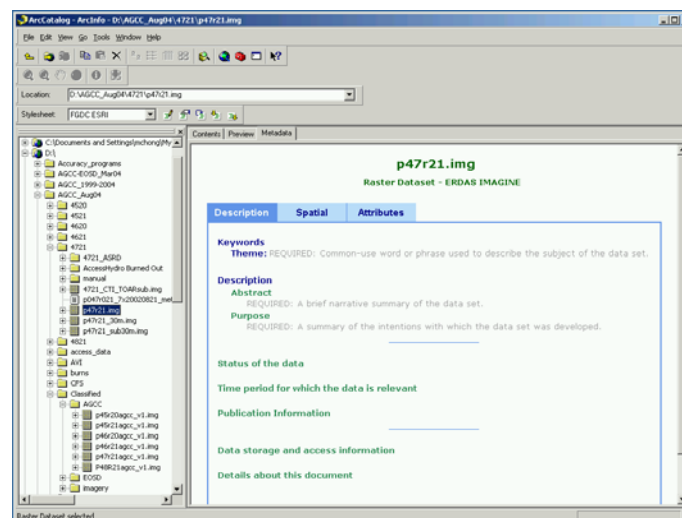


Figure 10.2. 1 Metalite application using the ArcGIS ArcCatalog interface

Selecting the *Metadata* tab, will display any metadata that may already be attached to the data. If not, information may be added by either importing existing metadata available in another format or by manually entering the information. To import metadata, the metadata must be available in \*.txt, \*.xml, or \*.sgml format all corresponding with the FGDC metadata standards. If these data are available, simply import using the import button seen below in Figure 10.2.2.

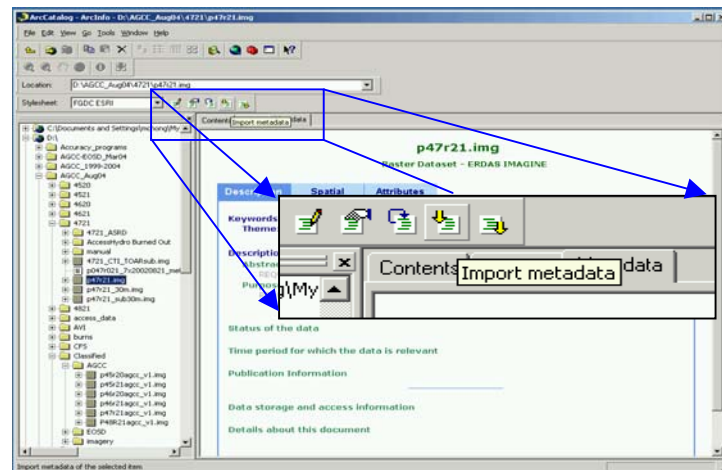


Figure 10.2. 2 Import metadata function in ArcCatalog

If metadata is not available in any of the aforementioned formats, using the Edit metadata function, metadata may be created. This function will also allow the modification of existing metadata. Systematically going through the tabs, fill in all relevant information according to the FGDC standard components. Once the metadata information is completed, it may be exported using the Export metadata button to allow the information to be saved in a format that may be imported and modified if necessary at a later date (Figure 10.2.3).

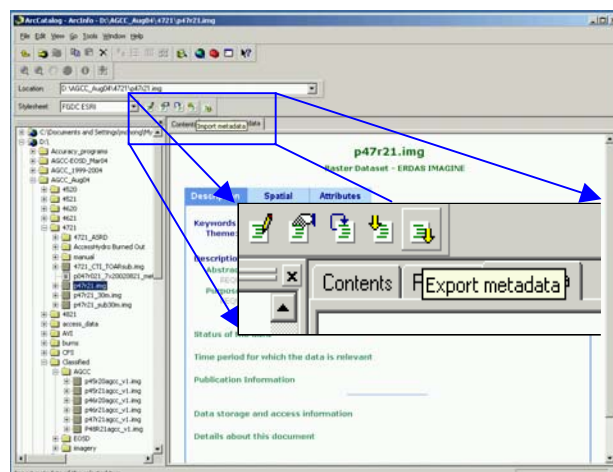


Figure 10.2. 3 Export metadata function in ArcCatalog

It is also good practice to export the metadata in html format to allow users with no access to ArcCatalog accessibility to the metadata. It is important to always provide informative documentation of the geospatial data produced. The following is an example of completed metadata in the ArcCatalog environment and in html format, respectively (Figure 10.2.4, Figure 10.2.5).

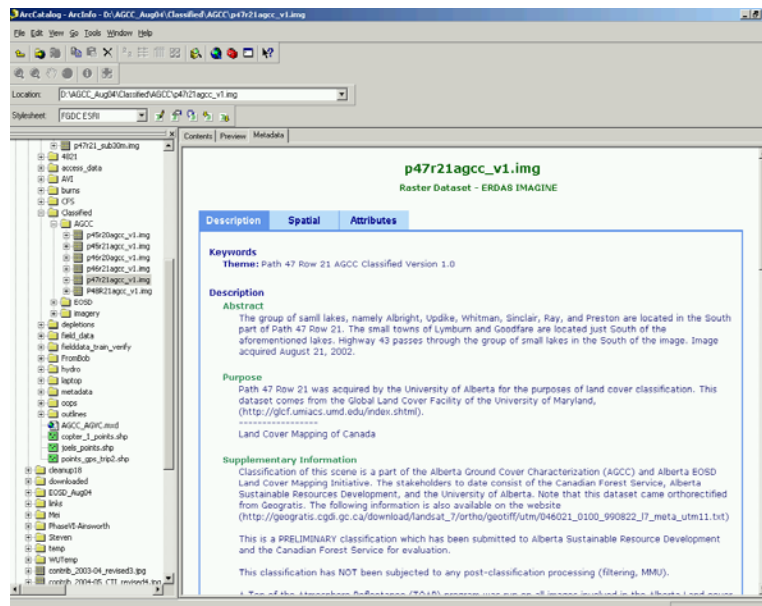


Figure 10.2. 4 Sample metadata in Metalite ArcCatalog interface

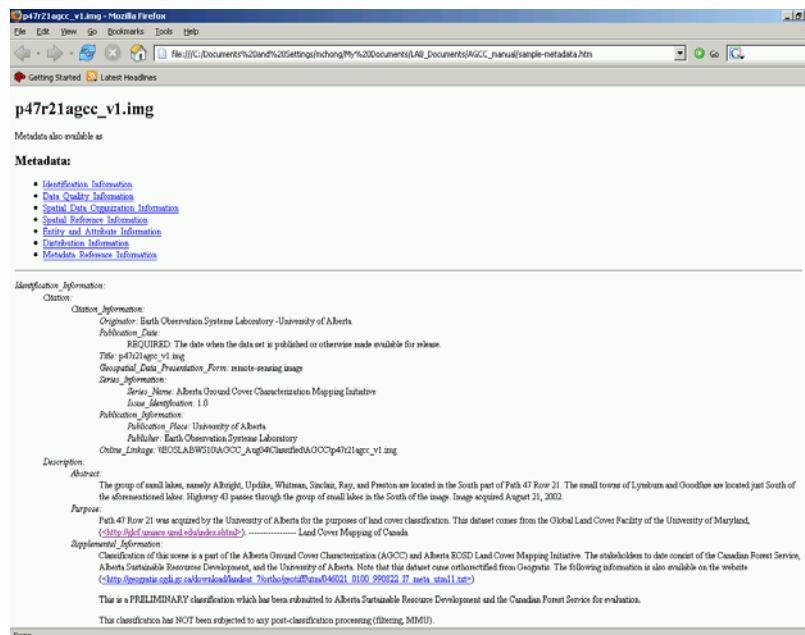


Figure 10.2. 5 Sample metadata in html format

## **11. Time Management for the AGCC**

### **11.1 Description of time required to complete tasks and processes for AGCC classifications**

Time management for the AGCC will vary depending on a number of factors. These factors include the level of AGCC classification wished to attain, the quality of the data acquired, the availability of manpower, software, hardware, etc. Recalling Figure 5.1.1, the time required to complete each task and process for the AGCC classification may be described. At stage T1, the data compilation and preparation process may take up to 30 minutes depending on the traffic of data from that site during the time of download. Once downloaded, the compilation and preparation process may take an additional 30 minutes totaling 1 hour for the T1 stage of processing.

At T2, the steps outlined in section four of the manual may take anywhere from 20 minutes to 1 hour depending on complications encountered with the MTL or metadata files. Problems with the syntax and format of the metadata files are common hindrances to ready completion of the TOAR process. However, the most common error or problem to arise is having an incorrect metadata file type. To avoid problems, ensure that the MTL file is of L1G file type.

Stage T3 involves the subset and extraction of known features including the roads, hydrography, clearcuts, and burns. This stage may take up to 2 days to complete due to the computing time required to process the numerous shapefile and coverage files used in creating the mask. Once the mask is completed, and a reduced working image void of all known features is created, stage T4 can commence. Stage T4 involves running a preliminary 50-class classification to isolate and extract ancillary features including urban areas, cloud, shadow, residual water bodies, and agriculture. Often times these various land covers are easy to identify, however it is the spectral separation of these classes that causes difficulty. Systematically isolating and extracting each ancillary feature may take up to 2 or 3 days depending on the degree of spectral confusion encountered.

Once ancillary features are isolated and extracted and a further reduced working image is created, another 50-class classification may be run on the image and stage T5 may commence. In stage T5 of the AGCC classification process, the primary land cover classes of interest are isolated and extracted. These classes include the various forest and vegetation classes encompassed in the AGCC classification scheme. This process may take up to 2 or 3 days depending on the level of spectral confusion present among the various classes and the AGCC level of classification desired. Once all land cover types are isolated and extracted, the reassemble process involved in stage T6 may begin. The reassemble stage will take no more than 1 day. The

various layers must be clustered, organized, and recoded to accommodate the AGCC class codes before all the individual layers may be added together. Upon reassembly, the preliminary classification is completed and ready for the accuracy assessment process.

The accuracy assessment process requires very little time. The various routines and programs described in section 7 of the manual can be completed in less than 30 minutes. However, the time required to make modifications and adjustments to the preliminary classification may be more time consuming. Depending on the level of adjustments needed, it may take 1 to 2 days to make all necessary adjustments. As such a full AGCC classification may be completed over period of 7 to 10 days depending on the level of spectral confusion encountered and the level of AGCC classification desired.

### **Suggested Readings**

Congalton, R. G. (1991). "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data." Remote Sensing of Environment **37**: 35-46.

ESRI, 1996. Technical paper, ESRI Inc., "Automation of Map Generalization: The Cutting-Edge Technology". White Papers Section, ArcOnline. Website. Accessed 12/9/04.  
<http://gis.esri.com/library/userconf/proc01/professional/papers/pap1060/p1060.htm>.

12. Appendices
Appendix A – AGCC Field Form

AGCC Field Form

Site Number, Obs. Date, Obs Level, Helicopter, Ground

Crew Members

1:50,000 Map Sheet, Elevation

Proposed Way Point, Actual Way Point

Northing (GPS), Easting (GPS)

UTM, Spheroid, GPS Error

Moisture Regime, Weather

Photo, Ob, Rep, Sky

Forest	Shrub	Herbaceous	Aquatic
Closed Conifer	Closed Tall	Dry Graminoid	Aquatic Bed
Open Conifer	Open Tall	Dry Forb	Emergent
Closed Deciduous	Closed Low	Wet Graminoid	Turbid Water
Open Deciduous	Open Low	Wet Forb	Clear Water
Closed Mixed		Lichen	
Open Mixed		Moss	
Other	Other	Other	Barren
Bog	Cropland	Rock	Sparse Veg
Fen	Pasture	Gravel	Rock / Gravel
Marsh	Meadow	Sand	Sand/Silt/Mud
Beaver Pond	Cut-block	Silt	
Lake	Fire	Mud	
River / Stream	Recent Fire	Deadfall	

AVI call Level 1

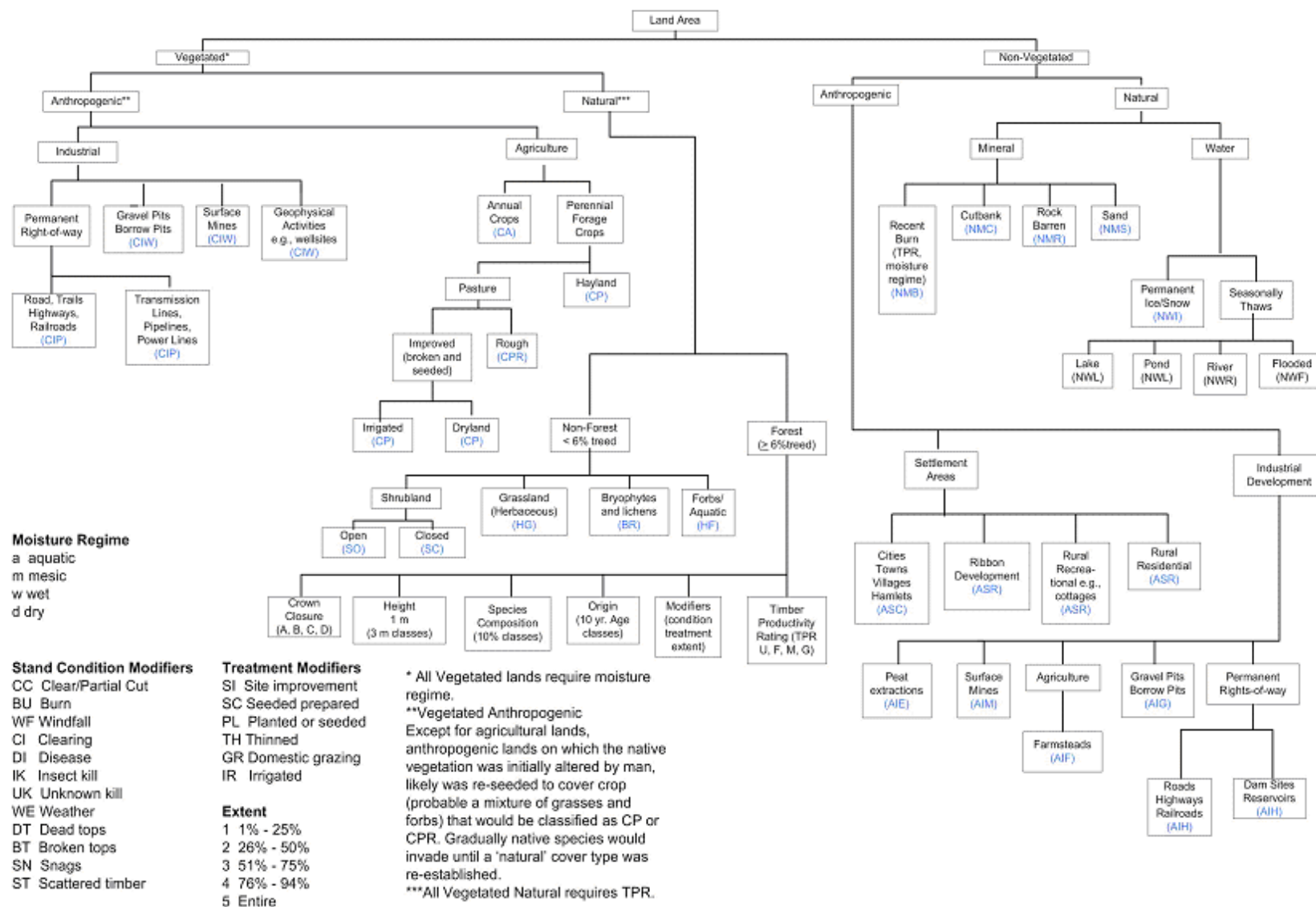
AVI call Level 2

AGCC Code, EOSD Code, GB Code

Comments

Tree Species	Sp. %	Ht	Shrubs	Sp. %	Ht	Herbaceous %	Sp. %
Jack Pine						Forbs	
White Spruce			Kinnikinnick			Horsetail	
Black Spruce			Willow			Fireweed	
Larch			Alder			Bryoid	
Balsam Fir			Bog Birch			Moss	
Aspen			Labrador Tea			Lichen	
Balsam Poplar						Graminoid	
Birch						Aquatic	Sp. %
Lodgepole Pine							
Total Crown Closure Measured by the Densiometer							
	C		NE		NW		SE
North							SW
East							
South							
West							
Crown Closure		A B C D		Average Densiometer Reading			
	Species	DBH	Ht (m)	SAP	Age	HLC	Species
1						16	
2						17	
3						18	
4						19	
5						20	
6						21	
7						22	
8						23	
9						24	
10						25	
11						26	
12						27	
13						28	
14						29	
15						30	

## Appendix B – Alberta Vegetation Inventory (AVI) flow chart



### Moisture Regime

a aquatic  
m mesic  
w wet  
d dry

### Stand Condition Modifiers

CC Clear/Partial Cut  
BU Burn  
WF Windfall  
CI Clearing  
DI Disease  
IK Insect kill  
UK Unknown kill  
WE Weather  
DT Dead tops  
BT Broken tops  
SN Snags  
ST Scattered timber

### Treatment Modifiers

SI Site improvement  
SC Seeded prepared  
PL Planted or seeded  
TH Thinned  
GR Domestic grazing  
IR Irrigated

### Extent

1 1% - 25%  
2 26% - 50%  
3 51% - 75%  
4 76% - 94%  
5 Entire

\* All Vegetated lands require moisture regime.

\*\*Vegetated Anthropogenic  
Except for agricultural lands, anthropogenic lands on which the native vegetation was initially altered by man, likely was re-seeded to cover crop (probable a mixture of grasses and forbs) that would be classified as CP or CPR. Gradually native species would invade until a 'natural' cover type was re-established.

\*\*\*All Vegetated Natural requires TPR.