TREE SPECIES MAPPING AROUND RECLAIMED OIL AND GAS WELLS SITES USING HYPERSPECTRAL AND LIGHT DETECTION AND RANGING (LIDAR) REMOTE SENSING

JAMES BANTING
Bachelor of Science, University of Lethbridge, 2009

A Thesis
Submitted to the School of Graduate Studies
of the University of Lethbridge
in Partial Fulfillment of the
Requirements of the Degree

MASTER OF SCIENCE

Department of Geography
University of Lethbridge
LETHBRIDGE, ALBERTA. CANADA

© James Banting 2016

TREE SPECIES MAPPING AROUND RECLAIMED OIL AND GAS WELLS SITES USING HYPERSPECTRAL AND LIGHT DETECTION AND RANGING (LIDAR) REMOTE SENSING

JAMES BANTING

Date of Defence: December 10, 2015

Dr. K. Staenz Co-Supervisor	Professor	Ph.D.
Dr. N. Rochdi Co-Supervisor	Adjunct Professor	Ph.D.
Dr. C. Coburn Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. A. Smith Thesis Examination Committee Member	Adjunct Professor	Ph.D.
Dr. S. Kienzle Chair, Thesis Examination Committee	Professor	Ph.D.

Dedication

I dedicate my thesis work to my parents, Gord Banting and Aruna Marathé, whose steadfast dedication to self-improvement have helped guide me through my life. Thank you for encouraging me to higher education and instilling a sense of curiosity in me from a very young age.

I also dedicate this work to my wife Justine. Your unwavering support has helped guide me through this process. I love you.

Abstract

Oil and gas activities in Alberta require disturbing forested lands, among other ecosystems, in order to extract resources. Due to the number of oil and gas sites requiring reclamation, monitoring can be problematic. Remote sensing provides cost-effective, timely, and repeatable data of these areas in support of monitoring efforts.

Support Vector Machine (SVM) and Multiple Endmember Spectral Mixture Analysis (MESMA) were tested in order to identify tree species around reclaimed and abandoned well sites near Cold Lake, Alberta using CHRIS satellite imagery with and without airborne LiDAR data. A hierarchical classification approach was employed, which achieved an accuracy of 83.4 % when using SVM together with CHRIS imagery and LiDAR. This positive result indicates the ability of remote sensing to support reclamation management and monitoring objectives within Alberta's forested areas.

Acknowledgements

I would like to thank my supervisors Karl Staenz and Nadia Rochdi for their support, insight and many meetings regarding what exactly I wanted to do.

I would like to thank my committee members, Anne Smith and Craig Coburn for their many helpful comments and discussions.

I'm very gratefully for the financial support provided through a Natural Science and Engineering Research Council of Canada (NSERC) CREATE scholarship (Advanced Methods, Education and Training in Hyperspectral Science and Technology; AMETHYST) and additional funding and resources from the University of Lethbridge (Department of Geography, School of Graduate Studies), the Alberta Terrestrial Imaging Centre (ATIC). This research was supported by TECTERRA funding through their University Applied Research Funding Program. I would also like to thank the Oil Sands Research and Information Network (OSRIN) as well as Alberta Environment and Sustainable Resource Development (ESRD), and in particular Shane Patterson from ESRD, for their general support, access to LiDAR data and aerial imagery, and shared knowledge throughout my thesis.

A special thanks to my fellow students, Devin Cairns, Tanya Byrne, Logan Pryor, Jurjen Van Der Sluijs, Peter Kennedy and Kevin Riddell.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AER – Alberta Energy Regulators
AESRD – Alberta Environment and Sustainable Resource Development
AISA – Airborne Imaging Spectrometer for Applications
AOT – Aerosol Optical Thickness
AVI – Alberta Vegetation Inventory
AVIRIS – Airborne Visible / Infrared Imaging Spectrometer
CARMS – Centre for Applied Remote Sensing, Modelling, and Simulation
CCM – Canopy Cover Model
CCMS – Committee on the Challenges of Modern Society
CHM – Canopy Height Model
CoB – Count-based Endmember Selection
CHRIS – Compact High Resolution Imaging Spectrometer
CS – Canopy Surface
CWV – Columnar Water Vapour
DBH – Diameter at Breast Height
DEM – Digital Elevation Model
DN – Digital Number
DTM – Digital Terrain Model
EAR – Endmember Average Root Mean Square Error
ELC – Equivalent Land Capability
EM – Endmembers
EMR – Electromagnetic Radiation
EOSD – Earth Observation for Sustainable Development
ESA – European Space Agency
ESRD – Alberta Environment and Sustainable Resource Development
ETM+ – Enhanced Thematic Mapper Plus
CGP – Ground Control Points
GIS – Geographic Information Systems
GPS – Global Positioning System
GSD – Ground Sampling Distance
IMU – Inertial Measurement Unit
ISODATA – Iterative Self-Organizing Data
LAI – Leaf Area Index
LCRC - Land Conservation and Reclamation Council
LiDAR – Light Detection and Ranging
LRD – Land Resource Development
LUT – Look-up Table
MASA – Minimum Average Spectral Angle
MASL – Meters above Sea Level
MESMA – Multiple Endmember Spectral Mixture Analysis
MODTRAN – Moderate Resolution Atmospheric Transmission
MSS – Multispectral Scanner
NASA – National Aeronautics and Space Administration
NATO – North Atlantic Treaty Organization

NIR – Near Infra-red
OLI – Operational Land Imager
PCA – Principal Component Analysis
PROBA – Project for On-Board Autonomy
RMSE – Root Mean Squared Error
ROI – Region of Interest
SMA – Spectral Mixture Analysis
SPOT – Satellite Pour l'Observation de la Terre
SVM – Support Vector Machine
SWIR – Short-wave Infra-red
TM – Thematic Mapper
USGS – United States Geological Survey
VALERI – Validation of Land European Remote sensing Instruments
VIPER – Visualization and Image Processing for Environmental Research
VIS – Visible
VISSR – Visible-Infrared Spin-Scan Radiometer

Chapter 1

Introduction

Alberta is a Canadian province rich in oil and gas resources. In response to increasing national and global energy demands, a network of oil and gas operations has resulted across the province. One of the primary methods of extracting these resources is using to drill wells into the oil and gas reserve cavern. This *in-situ* method results in the land surrounding the well becoming disturbed in an effort to meet operational, safety and environmental requirements. At the completion of the oil and gas activities, called abandonment, the wellhead is sheared off below the ground surface, the interior of the well is filled with concrete, the wellhead capped, and attempts are made to establish the pre-vegetated condition on the well-site.

The rate at which wells are being abandoned and the rate of wells being reclaimed are not consistent. In fact, there are significantly more abandoned than reclaimed wells in the province (AESRD, 2013a). With approximately 50,000 un-reclaimed wells in Alberta in 2012, monitoring the status and health of these sites is a monumental task (Figure 1.1). Many studies dealing with reclamation efforts focus on large areas, such as open pit mines (Schmid *et al.*, 2013; Lévesque and Staenz, 2008 and 2004; Lévesque *et al.*, 1999) rather than concentrating on a smaller scale as is the case for oil and gas well-sites. Based on these facts, there is a need for studies focusing on reclamation efforts in Alberta with respect to oil and gas well activities.

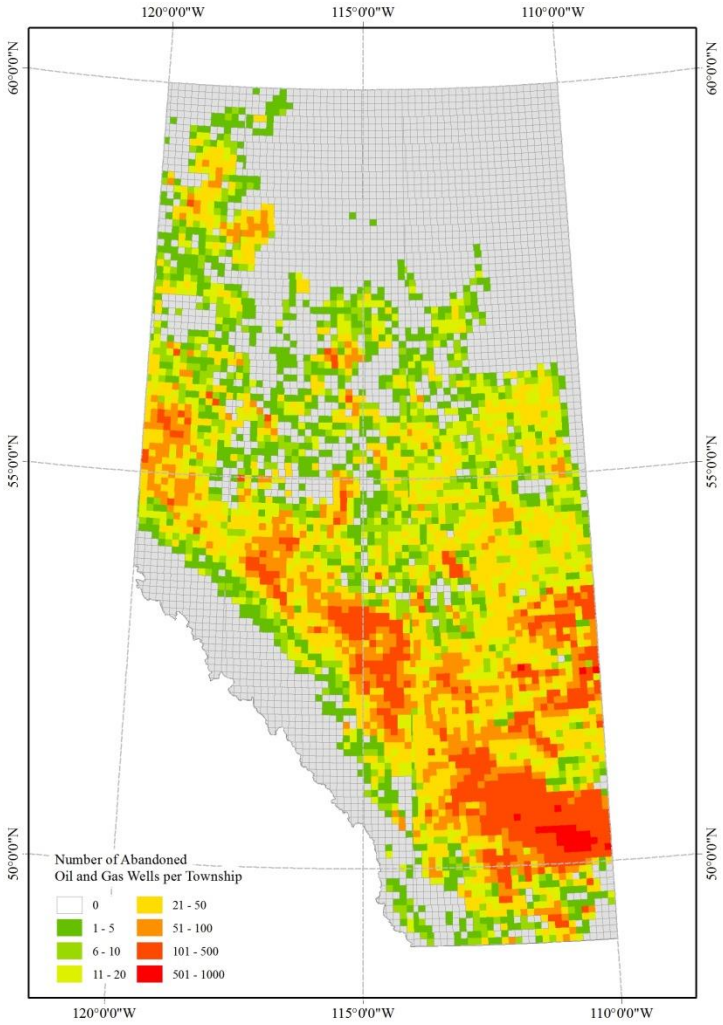

Figure 1.1: Extent of abandoned oil and gas wells in Alberta as of March 2013.

The current form of well-site reclamation assessment in Alberta is conducted by the Alberta Energy Regulator (AER). The responsibility for well-site reclamation formally fell to Alberta Environmental Sustainable Resource Development (AESRD). The assessment consists of an audit conducted by the compliance staff on approximately 10 % to 15 % of the reclamation certification applications submitted by companies to ensure that well-site reclamation meets compliance guidelines (AESRD, 2013a). This is an opportunity to use Remote Sensing, since many wells can be assessed in a single image capture. Unless a well-site is exempt from the reclamation process, all reclamation

activities in the province are governed by Part 6 of Chapter E-12 of the Revised Statutes of Alberta 2000 - Environmental Protection and Enhancement Act (Government of Alberta, 2013). This act states, in part, that an inspector can require further reclamation to be conducted on a site, even if it has been issued a reclamation certification. The process of further reclamation is at the inspector's discretion. Because the current reclamation model has a liability period of 25 years, the land must be monitored during that time to ensure it conforms to the regulator reclamation guidelines.

Reclamation activities attempt to transform the disturbed land back into a productive state that is equivalent to pre-disturbance conditions. However, in order for this to be accomplished, a baseline of natural vegetation around the site needs to be established. One method of accomplishing this is to identify the vegetation species present. This can be completed through field transects and site visits, but this is time consuming and expensive. A much more realistic option is to use remote sensing, which allows for monitoring large areas of land in a cost-effective and timely fashion (Foody and Embashi, 1995; Sellers *et al.*, 1995; Mumby *et al.*, 1999).

Forests cover approximately 60 % of Alberta and specific reclamation guidelines dealing with this land-cover type have been created. This thesis will focus on oil and gas reclamation monitoring in forested areas and will utilize remote sensing technologies in pursuit of tree species mapping.

1.1 Remote Sensing

Remote sensing approaches to forest-cover mapping over large geographical areas have generally focused on multispectral sensors, such as the ones found on the Landsat and *Satellite Pour l'Observation de la Terre* (SPOT) satellites. The use of these two

platforms is well documented in the scientific literature for monitoring land cover and any changes that have occurred (Hansen *et al.*, 2013; Hurd *et al.*, 2001; Masek *et al.*, 2000). Satisfactory results have been achieved using hyperspectral sensors as shown in Table1.1 (Dalponte *et al.*, 2013; Galidaki *et al.*, 2012; Ghiyamat and Shafri., 2010; Plourde *et al.*, 2007; Goodenough *et al.*, 2003a; Martin *et al,*. 1998).

The data capture in contiguous narrow bands from these sensors allows for subtle variations in the electromagnetic spectrum, which permit better discrimination of tree species than with multispectral sensors (Lee et al., 2004; Clark et al., 2005). Additionally, the use of structural information within the vegetation, derived from Light Detection and Ranging (LiDAR) data, can aid in determining tree species when combined with multispectral / hyperspectral sensed imagery (Dalponte et al., 2012; Jones et al., 2010; Anderson et al., 2008; Dalponte et al., 2008; Holmgren et al., 2008).

Table 1.1: Table comparing hyperspectral remote sensing studies.

Sensor Type	**Data Source**	**Classification Methods[1]**	**Overall accuracy**	**Source**
Hyperspectral	Hyperion	SAM	94%	(Galidaki *et al,*. 2012)
Hyperspectral	Hyperion	ML	81%	(Goodenough *et al,*. 2003a)
Hyperspectral	AVIRIS	ML	75%	(Martin *et al.*, 1998)
Hyperspectral	HySpex	SVM, RF	93%	(Dalponte *et al.*, 2013)

[1] Classification Methods: Spectral Angle Mapper (SAM), Maximum Likelihood (ML), Support Vector Machine (SVM), Random Forest (RF),

In 2001, the European Space Agency (ESA) launched the Project for On-Board Autonomy (PROBA) satellite, in part, to test experimental technologies such as the Compact High Resolution Imaging Spectrometer (CHRIS; ESA, 2014). This sensor captures data in 62 bands, covering a wavelength range from 400 nm to 1050 nm, at 34-m ground sampling distance (GSD) in full-swath mode and from 5 different viewing angles

(+55°, +36°, 0°, -36°, -55°) (Table 1.2). Other imaging modes can capture information in either 18 or 37 bands with a 17-m GSD. The band ranges for these other modes were optimized for applications related to land, water and chlorophyll monitoring (ESA, 2014). The different imaging modes vary with respect to both the number and width of spectral bands in order to highlight certain areas of the electromagnetic spectrum during data capture.

Table 1.2: CHRIS Imaging modes.

Mode	Specified Use	Number of Bands	Swath Width (km)	Ground Sampling Distance (m)
Mode 1	Full Capacity	62	~14	34
Mode 2	Land	18	~14	17
Mode 3	Water	18	~14	17
Mode 4	Chlorophyll	18	~14	17
Mode 5	Land	37	~7	17

The hyperspectral aspect of CHRIS means that the sensor captures a wide range of spectral data. However, due to the moderate sized footprint of the sensor (17-m or 34-m), the spectral signature recorded by the sensor is a combination of spectral signatures from all elements present within the pixel. In order to separate the different materials, including different tree species, which may be present within a pixel, sub-pixel spectral unmixing algorithms must be applied. Cataloguing vegetation structural parameters, such as canopy height and cover, return productivity related information about vegetation species and health (Husch *et al.*, 2002), both of which are useful indicators in a reclamation assessment.

Remote sensing systems both passive, observing reflected electromagnetic radiation (EMR) emitted by a source other than the remote sensing system, and active, observing reflected EMR emitted by the remote sensing system itself, provide the means

to acquire these parameters. For example, photogrammetric interpretation of aerial images can provide analysts with canopy height and cover measurements. However, it is time consuming and requires highly specialized individuals to derive accurate information. LiDAR is a more viable tool for reclamation monitoring as this technology can provide point measurements over a large area and is able to penetrate the canopy and provide multiple ranging measurements for a single point. The information contained within each point can be used to derive structural information in a timelier fashion than traditional image interpretation methods.

1.2 Data Fusion

The fusion of different remotely sensed datasets has been used for a long time in remote sensing studies and has proven to be beneficial in that data gaps from one source can filled in by another source (Hall, 1992; Genderen and Pohl, 1994). Structural information, for example, can support spectral information when investigating land-cover and tree-species classification. (Sullivan *et al.*, 2009; Dalponte *et al.*, 2008; Hyde *et al.*, 2006; Gong and Howarth, 1990). Different vegetation types have different structures (e.g., coniferous tree compared to deciduous tree). Accordingly, integrating information about the structure type supports the discrimination of the different vegetation species present.

1.3 Tree Types

Trees can be either evergreen or deciduous. The former means that leaves persist throughout the year, while the latter means that leaves are shed at the end of the growing season. Coniferous refers to trees that are cone bearing and make up the majority of

evergreen trees within the Canadian Boreal region with the exception being Jack Pine and Larch. Determining which type of tree it is and, ultimately, what the species is, can be accomplished by analyzing the structure of the tree as well as the biological parameters. For example, most conifer trees are conical in shape and have needles instead of leaves (Figure 1.2). Deciduous trees, on the other hand, have large broad canopies with leaves (Figure 1.3).

Figure 1.2: White Spruce tree structure (left) and needles (right). Notice the conical shape of the White Spruce trees (darker green).

Figure 1.3: Aspen Poplar tree stand (left) and leaves (right).

Tree types have different spectral characteristics, which can be exploited in order to assist in classifying forest cover in satellite imagery (Nelson *et al.*, 1985; Shen *et al.*, 1985; Hodgson *et al.*, 1988; Lathrop *et al.*, 1994).

Determining what tree species are present around an abandoned well-site as well as species type of vegetation regrowth within the well-site is part of a broader reclamation management plan and is a requirement under Alberta's reclamation guidelines (AESRD, 2013a).

1.4 Objectives and Hypothesis

This thesis documents the work carried out to combine hyperspectral CHRIS data with structural LiDAR data in order to better discriminate tree species in forested regions and abandoned well-sites in the area of Cold Lake, Alberta. The LiDAR data are processed using an open-source computer program to create the vegetation structure products. Selected spectral unmixing and image classification methods are applied to CHRIS data alone and in combination with LiDAR data. The two are compared to each other to determine the method that yields the best accuracy when compared to field-reference data. This, in turn, determines whether the inclusion of LiDAR data are beneficial for oil and gas reclamation monitoring.

In order to evaluate this approach, the research objectives for this study are as follows:

(a) Assess the use of hyperspectral data alone and in combination with LiDAR data for monitoring species composition in the Cold Lake area of Alberta and

(b) Evaluate spectral unmixing methods for tree species discrimination.

The hypothesis of this study is that combining structural and spectral data over a forested reclamation site will yield classification results that are more accurate than those extracted from spectral data alone. Furthermore, the use of spectral unmixing will help to facilitate tree species discrimination better than discrete classifiers such as SVM.

1.5 Thesis Organization

The thesis is organized in a conventional structure where the first chapter provides an introduction to the topic and summarizes the research hypothesis and objectives. The second chapter reviews current and past research on land reclamation using remote sensing as well as including sensor systems and classification technologies. The third chapter outlines the methods, which were used for this study and describes the location of the study area together with characterization of the local and regional environment. The fourth chapter presents the results of the species composition classification and provides an evaluation of using hyperspectral data alone and in combination with LiDAR data. The fifth chapter discusses the results and reveals their implications for reclamation monitoring in the Alberta's oil patch, while the sixth and final chapter concludes the findings of the thesis.

Chapter 2

Literature Review

The following literature review will examine reclamation practices and criteria, related remote sensing fundamentals and data processing techniques, and how remote sensing can be beneficial for monitoring tree species at reclaimed oil and gas well-sites.

2.1 Land Reclamation

The area of land reclamation is diverse in both the methods used to conduct reclamation and the expected outcomes of that reclamation. Taken literally, it can mean the creation of new lands for human occupation from otherwise unsuitable areas (Curtis and Campopiano, 2013). Notable examples include areas in The Netherlands (Hoeksema, 2007), Chek Lap Kok Airport in Hong Kong (Liu *et al.*, 2001), and New Orleans, Louisiana (Olshansky, 2006; Craig *et al.*, 1979).

Another definition of reclamation, used primarily in Canada, refers to returning disturbed lands, anthropogenic or naturally caused, to a state similar to that of the surrounding land or to a state agreed upon by the stakeholders (Bradshaw, 1984). Outside of Alberta, the terms remediation and restoration are often used in place of reclamation. This can cause confusion when describing similar practices. Accordingly, reclamation will be used in this thesis to describe the process outlined by Bradshaw (1984).

The following subsections will provide a brief history of global and national reclamation practices, the methods used to conduct reclamation and its monitoring, and finally, an overview of detailed reclamation practices as they apply to oil and gas activities in Alberta.

2.1.1 History

One of the first instances of regulations, dealing with contaminated lands, came from the United Kingdom in 1936 with the creation of the *Public Health Act* (Parliament of the United Kingdom, 1936). This Act prohibits building on lands where polluting materials have not been cleaned up. Adding onto this act was the *Town and Country Planning Act of 1947* (Parliament of the United Kingdom, 1947) that required a land-use plan in place for any new developments and encouraged any resource extraction activities to have and enact a reclamation plan. Early forms of reclamation regulations have also been in place in Sweden since 1969 with the *Environmental Protection Act* (Swedish Environmental Protection Agency, 1977), Germany since 1972 with the *Federal Waste Disposal Act* (Federal Law Gazette I, 1977) ,the United States since 1976 with the *Resource Conservation and Recovery Act* (EPA, 1992), and the Netherlands since 1984 with the *Soil Protection Act* (Ministry of Infrastructure and the Environment, 1996),

In 1969, the North Atlantic Treaty Organization (NATO) created the Committee on the Challenges of Modern Society (CCMS) tasked with studying, in part, issues affecting the environments of member nations. A pilot study initiated by NATO/CCMS in 1980 assessed available environmental reclamation technologies and methods in order to understand the impact of military activities on the environment. The study identified different reclamation needs and practices in seven nations and sought to provide a means by which reclamation information and techniques can be shared between nations (NATO, 1998).

Reclamation legislation in Canada is the jurisdiction of the provinces and territories for administration/regulation. Each province tailors the regulations to the

industry operating within its borders (Sousa, 2001). Reclamation within Alberta is primarily focused on the oil and gas industry; however, for example, it is also applicable to mines, sand and gravel pits, and commercial development.

Prior to 1963, reclamation of industrial lands was not required and, therefore, not often conducted. The introduction of the *Surface Reclamation Act* (Government of Alberta, 1963) required that minimum standards be met, such as hazard removal and site clean-up, in order for a reclamation certificate to be issued. This act was created in part due to public concerns over the environmental impacts of petroleum exploration and extraction as well as land-owner surface rights compared to mineral rights. This surface/mineral rights discussion focused on the idea that land should be returned to the same condition in which it was found before the disturbance took place (Powter *et al.*, 2012).

In 1973, the *Land Surface Conservation and Reclamation* Act (Government of Alberta, 1973) replaced the *Surface Reclamation Act*. This new act created monitoring, enforcement, and soil salvage guidelines and the establishment of a Land Reclamation Division (LRD) and Land Conservation and Reclamation Council (LCRC). Both entities were tasked with developing policy, monitoring reclamation and issuing reclamation certificates.

Beginning in 1992, the Government of Alberta introduced the *Environmental Protection and Enhancement Act* (Government of Alberta, 1992) and added the *Conservation and Reclamation Regulation Act* a year later (Government of Alberta, 1993). The former consolidated and replaced previous environmental protection acts, while the latter outlined regulatory processes related to the issuance of reclamation

certificates and enforcement of reclamation guidelines. Furthermore, it identified the objective of reclamation as a return of the land to an Equivalent Land Capability (ELC; Government of Alberta, 2013). Reclamation efforts relating to energy extraction are a large part of Alberta's reclamation programs and much effort has been devoted to defining what reclamation entails for the industry involved in natural resource and energy extraction (Powter *et al.*, 2012).

2.1.2 Reclamation Methods

The methods used to reclaim disturbed lands vary depending on the level of severity and the expected condition of the land after the completion of reclamation. Regardless of the desired outcome, all reclamation activities start with an assessment of the disturbed site. The assessment process is in place in order to decide if a specific site poses a risk to human health and/or to the environment (Burger, 2008; DEFRA, 2004). This process is also in place to identify realistic management and technological decisions for reclamation of a specific site (Bardos *et al.*, 2001).

Following an assessment phase, the actual reclamation of the disturbed land takes place. A site can be left to regenerate naturally through vegetation succession, or it can be transformed through changes in topography as well as seeding and planting vegetation to meet specific reclamation goals outlined in the assessment phase. Attention to the biological condition of the site and surrounding lands also needs to be considered.

For some reclamation plans, this can mean simply seeding the site with grass to cover the bare earth and prevent erosion (AESRD, 1995). For other situations, more-involved reclamation activities are required, including an understanding of the soil types

and vegetation species present around the site as well as the environmental controls dictating vegetation growth.

After the disturbed land has been transformed to a reclaimed state, the land needs to be monitored to identify any deviations between the reclamation plan and the current reclamation state. Monitoring is also conducted to ensure that plants and animals on the site are not adversely affected by either the reclamation itself or any contaminants left on the site.

Burger (2008) identified nine steps for reclamation of a disturbed site. The first three steps take the form of a reclamation assessment where the ecosystems of surrounding lands are identified. The remaining six steps address the need for physical manipulation of the site, the selection of appropriate plants and animals, and the creation of a reclamation timeline that mimics the natural succession of vegetation on the site. While these steps are very thorough, they cannot be applied to every site and instead must be tailored to the site in question. Tailoring a site may be due to budgetary constraints or reclamation goals (i.e. different land-use),

Creating a reclamation management plan is crucial for a successful reclamation campaign. Tailoring these plans to specific areas and disturbance types ensures that appropriate reclamation goals are addressed and met.

2.2 Alberta Specific Reclamation Methods

Upstream (i.e., site of resource extraction) oil and gas reclamation can fall into the three land-cover types, cultivated lands, native grasslands, and forested lands, each having different criteria for reclamation (AESRD, 2013 a,b, and c). Within each disturbed site, the following three parameters must be assessed: landscape, soils, and vegetation.

Cultivated lands and native grasslands are beyond the scope of this research, and within the context of forested land reclamation, the focus will be on vegetation. A summary of the vegetation parameters considered for reclamation assessment within forested lands is presented in Table 2.1.

Table 2.1: Summary of the reclamation criteria for vegetation parameters of forested lands (courtesy of AESRD, 2011) .

Assessment Parameter	Forested Lands
Species Composition	Type and mix (woody and herbaceous) should be consistent with native species present and should meet reasonable land management objectives. ***Natural Recovery:*** stem count of 5[1] is required. ***Planted sites:*** stem count of 2[1] is required for the merchantable seedlings; a minimum of 2 growing seasons required after planting prior to conducting the vegetation assessment.
Plant Height Plant Density Head/Pod/Tuber – Length Head/Pod/Tuber – Weight	No criteria requirements.
Litter	No criteria requirements.
Plant Health	Plants should be healthy, signs of stress onsite should be comparable to those observed offsite
Head/Pod/Tuber – Health	No criteria requirements.
Weeds	Weeds present (i.e., composition and distribution) onsite must be comparable to offsite. Weeds must be managed as per the weed control act; restricted weeds destroyed and noxious weeds controlled. Must not impede landowner operability.

[1] A circular plot with a radius of 1.78 -m ($10m^2$) is to be used for the assessment.

From Table 2.1, plant height and density are the only structurally related criteria in the assessment parameters. However, observing reclaimed vegetation structure, such as canopy height and canopy cover, and comparing it to undisturbed vegetation structure,

can help regulators to determine if reclamation is proceeding normally or if further reclamation activities are needed. Additionally, vegetation structure parameters can be useful for enhancing species composition mapping using passive optical remote sensing (Jung, 2011; Puttonen *et al.*, 2010; Dalponte *et al.*, 2008; Hill and Thomson, 2005; Lefsky *et al.*, 2001; Dubayah and Drake, 2000).

2.3 Equivalent Land Capability

The objective of ELC is to ensure that the ability of any land after a disturbance to support different land uses is similar to the ability that existed prior to the land that was disturbed. The above statement is a paraphrase of the definition of ELC in the *Conservation and Reclamation Regulation* (Government of Alberta, 1993). It further states that what defines ELC for a site must be determined by stakeholders.

The capability of the land refers to what the intended land-use of the site will be. Accordingly, in a forested region, the intended land-use might be a return to a forested area or it might be deemed that an ELC is a grassy meadow (OSRIN, 2011). The Alberta regulations regarding reclamation in forested areas are such that in order for ELC to be met, a previously forested area must be returned to a forested state unless an exemption is made (AESRD, 2013a).

If an oil and gas company meets the criteria listed in the reclamation criteria for well-sites and associated facilities (AESRD, 2013a,b, and c) and has achieved an ELC, it can apply for a reclamation certificate, stating that they have performed reclamation on the site in question and that it has been returned to a productive state. Companies are liable for any issues relating to improper reclamation for 25 years after which liability

reverts to the Alberta Government. However, liability remains with the company for lifetime regarding any issues related to contamination (AESRD, 2012b).

Remote sensing science aims to assess what parameters about a target can be assessed from a distance. It can help with long-term monitoring efforts as mapping and monitoring of forests has been ongoing for over 60 years, and the methods used to assess forest parameters have been evolving and improving over that time span (Losee, 1942; Seely, 1949). The technologies, employed when applying remote sensing, are able to cover large geographical areas and can provide information about vegetation health and composition in a timely fashion. Remote sensing technologies have been used to monitor global forest change over time (Hansen *et al.*, 2013), mine growth and reclamation (Townsend *et al.*, 2009) and seasonal variability in vegetation health (Van Leeuwen, 2008; Huete and Didan., 2004).

2.4 Application of Remote Sensing to Reclamation

Vegetation stress at a reclaimed site can be indicative of issues regarding improper reclamation practices or simply a need for further reclamation to be done. Additional reclamation can be needed due to residual effects of the disturbance or out-dated reclamation practices. Gauging the vegetation species development at a reclaimed site can help reclamation managers to assess moisture regimes, topographic controls and soil composition, because certain tree species prefer to grow in specific environmental conditions (Hosie, 1979). Additional information can be obtained about a reclamation site by observing it over a long period of time. These long-term monitoring studies help to account for variability within a site and present a more holistic view of the reclamation area by assessing if a site is following the established reclamation plan or not.

In order to help frame the use of remote sensing in reclamation management, a brief introduction to remote sensing is presented. Following the remote sensing introduction, literature that deals with vegetation assessment will be presented and will describe how these studies relate to reclamation monitoring. Emphasis will be placed on tree species mapping, as that is the stated objective of this thesis. However, attention will also be given to vegetation health monitoring, defined in this thesis as vegetation density, height, and canopy cover. Studies identifying long-term monitoring efforts for reclamation will also be presented.

2.4.1 Remote Sensing Introduction

Reflective optical remote sensing focuses on the interaction of vegetation to EMR in the visible (VIS; 400 – 700 nm), near infrared (NIR; 700 – 1000 nm), and short wave infrared (SWIR; 1000 – 2500 nm) regions. The response of vegetation to the incoming radiation is dependent on the internal structure and chemical composition of this material (Figure 2.1). Within the VIS region, EMR is used by healthy plants to facilitate plant growth by converting EMR into sugars through photosynthesis. Vegetation interactions within the NIR and SWIR regions are dominated by internal plant water content and the physical characteristics of the plant itself. By assessing the different spectral features within the response of vegetation canopies, inferences can be made to the health and species of the plant in question (Myneni *et al.*, 1995).

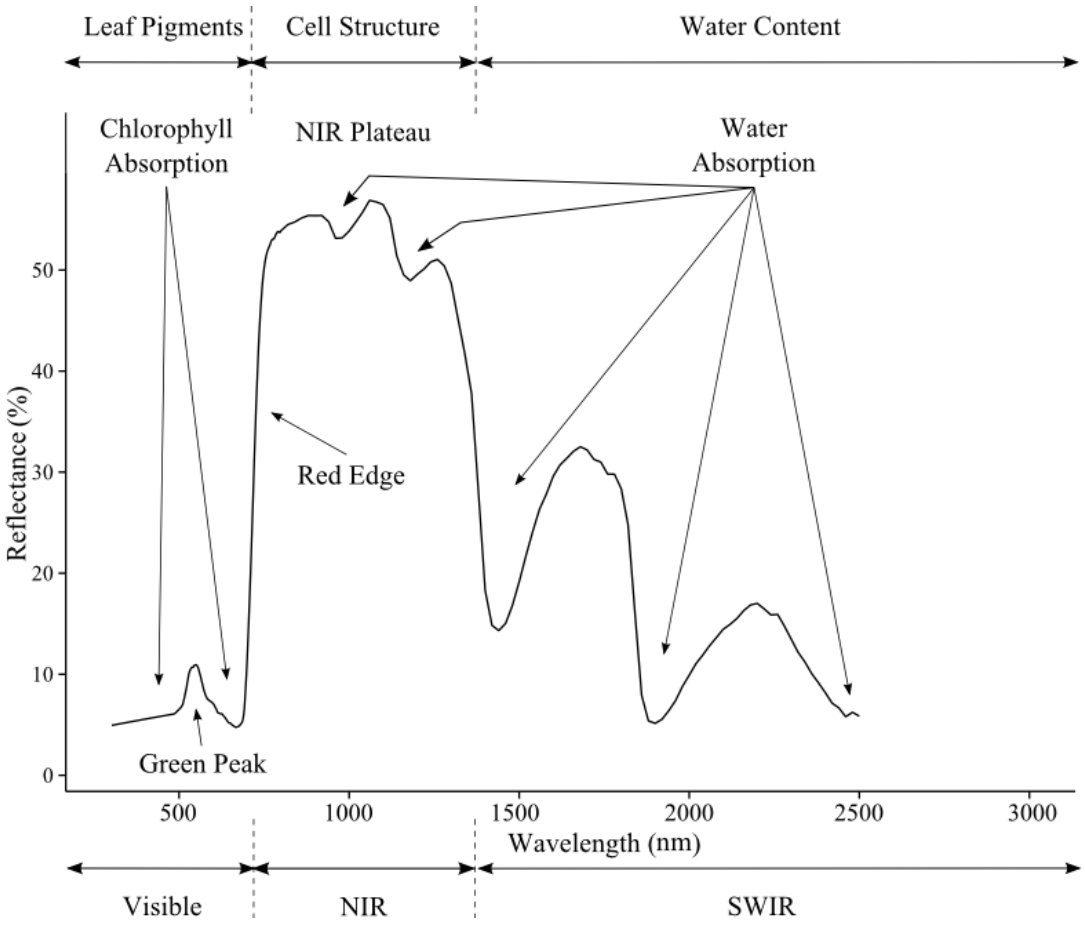

Figure 2.1: Reflectance spectrum of photosynthetically active vegetation highlighting three different EMR regions as well chemical absorption features.

2.4.2 Remote Sensing for the Purpose of Reclamation

Remote sensing studies for reclamation management have focused on pipeline-right-of ways in the oil and gas industry and mining activities (e.g., open pit mining and mine tailings). Um and Wright (1998) conducted research into monitoring vegetation succession along pipeline right-of-ways for the purpose of reclamation monitoring. They determined that using remote sensing as a tool to monitor large areas was a viable method of reclamation management. This approach is well suited for Alberta where many sites, spread out over large areas, need to be monitored. Other studies dealing with reclamation along right-of-ways used remote sensing to assess vegetation health as a proxy for

infrastructure problems (e.g., pipeline leaks; Gauthier *et al.*, 2001) and to determine to what impact spectral resolution influences vegetation anomaly detection for the purpose of identifying pipeline leaks (Booth and Cox, 2009).

Mining activities, in particular open-pit mining, disturb large areas of land and will require reclamation once mining activities have ceased. The ability of remote sensing to cover large tracts of land at medium (approximately 30 m) and high-spatial resolution (approximately 5 m), makes it an ideal monitoring tool for reclamation (Schmid *et al.*, 2013; Townsend *et al.*, 2009; Lévesque and Staenz, 2008; Brown *et al.*, 2005; Irons *et al.*, 1980).

2.4.3 Tree Species

Understanding the distribution of tree species within a forest can help reclamation managers with planning and reclamation following a disturbance (Plourde *et al.*, 2007). Field surveys, aerial photo interpretation and digital satellite remote sensing are all means by which tree species distribution can be estimated. However, the former two are time consuming to plan and execute and cannot cover large areas in a timely manner as quickly compared to satellite remote sensing. It can provide timely data over a large area and is shown to be a good estimator of tree species and their distribution (Dalponte *et al.*, 2012; Galidaki *et al.*, 2012; Heinzel and Koch., 2012; Vaughn *et al.*, 2012; Puttonen *et al.*, 2010; Dalponte *et al.*, 2008; Peddle *et al.*, 2004; Dennison and Roberts, 2003a; Martin *et al.*, 1998; Roberts *et al.*, 1998).

For forest reclamation, identification of the tree species is important not only as a measurement tool to check reclamation success, but also as a way to help guide the entire reclamation strategy, including the ELC for a site. There are many remote sensing

techniques by which tree species can be assessed, but they generally fall into the two categories, discrete classification methods, whereby a pixel is assigned only to one class, and sub-pixel unmixing methods, whereby a pixel is assumed to contain a proportion of different spectral signatures, including but not limited to different tree species. Some examples of different approaches are presented below.

The spatial resolution of a remotely sensed image can be a factor in determining the remote sensing techniques that will be used for reclamation monitoring. A lower spatial resolution, which is the smallest object on the ground that can be detected in an image, coupled with a discrete classification method can fail to observe tree species by only classifying the dominant spectral signature within a pixel. Conversely, a high-spatial resolution image coupled with a sub-pixel unmixing method can overwhelm an unmixing algorithm with too many different spectral signatures, thus, rendering the classifier unable to reliably determine tree species within the pixel. While these issues are not unique to tree species identification using remote sensing nor are they unique to reclamation management, they highlight a need for further studies investigating tree species mapping within the land reclamation framework.

Nevertheless, tree species mapping using remote sensing technologies is a way for reclamation managers to track the progress of vegetation growth on a reclamation site. Understanding the type of tree species present at a site ensures that one of the criteria towards ELC is met. For a more in-depth assessment of the vegetation health at a reclamation site, a vegetation health-monitoring program is needed.

2.4.4 Vegetation Health Monitoring

Vegetation can be affected by natural stressors such as competition, fire, and insects (Goodwin *et al.*, 2008; Tømmervik *et al.*, 1995; Carter, 1993) or anthropogenic stressors, such as vegetation isolation through harvesting, oil and gas activities, and climate change (Bachelet *et al.*, 2001; Pezeshki *et al.*, 2000; Halpern and Spies, 1995). Remote sensing provides the tools by which quantitative measures of vegetation health affect the spectrum and can be assessed.

In a study on monitoring vegetation stress after a petrochemical spill, Li *et al.* (2005) investigated polynomial fitting and Spectral Mixture Analysis (SMA) as means of detecting vegetation stress. The authors used two hyperspectral Airborne Visible / Infrared Imaging Spectrometer (AVIRIS) datasets, one prior to the spill and one post spill. They found that while the fitting methods were inconsistent in detecting a spectral change, SMA was able to show a decrease between 10 % and 30 % in vegetation abundance. The authors attributed this decrease to vegetation stress due to a change in vegetation cover.

The spectral profile of vegetation can be examined to derive vegetation stress indicators, such as chlorophyll content, Leaf Area Index (LAI), and water content (Wang and Li. 2011; Delegido *et al.*, 2008; Zarco-Tejada *et al.*, 2001; Serrano *et al.*, 2000; Peddle *et al.*, 1999).

Remote sensing for vegetation health assessment has proven to be a useful tool for large and small vegetation monitoring projects and will continue to be. This is evident by the recent satellite launches from the Landsat and Sentinel remote sensing programs of the National Aeronautics and Space Administration (NASA) in collaboration with the

United States Geological Survey (USGS) and European Space Agency (ESA), respectively. An understanding of the vegetation health in and around a reclamation site can inform managers about the efficiency of their reclamation practices. However, in order to understand vegetation health trends and how they can be implemented in reclamation management, a series of measurements over time is required.

2.4.5 Temporal Studies

Time series analysis of remote sensing data is able to provide a history of changes within an area and can be useful for reclamation managers in charting site development. A multi-temporal approach was used by Townsend *et al.* (2009) to monitor and map land-cover as it changed from an active mine site to reclaimed land. Emphasis was placed on the use of good ancillary Geographic Information System (GIS) data such as land-cover polygon classes in order to help identify how land cover changed over time and to chart the trajectory of reclamation management practices. The authors found a reduction in surface mine areas using the remotely sensed images. Furthermore, the authors were able to quantify the reduction in surface mine areas between image acquisition dates. A similar approach was used by Champagne *et al.,* (2004) using the Hyperion hyperspectral sensor and Landsat 5 multispectral sensor to monitor vegetation health over reclaimed mine areas in Sudbury, Ontario. The authors were interested in the ability of remote sensing to provide reliable land-cover information over long time periods. Their results showed that remote sensing produces information on land cover that is consistent with field based campaigns.

Remote sensing is well suited for use in forest reclamation management and monitoring projects as it provides the ability to monitor environmental receptors, such as

tree species and vegetation health. Both of which are able to provide indications to the status of previous reclamation practices and highlight the need, if any exists, for further reclamation management. Furthermore, by capturing remotely sensing images of a long-time period, trends in progression of reclamation can be evaluated.

2.5 Remote Sensing and Remote Sensing Methods

Remote sensing systems can capture radiance data that needs to be processed before reliable information can be retrieved. This involves the calibration of the radiance data and correcting for atmospheric attenuations, geometric alignment due to topographic effects, and sensor artifacts, such as smile and keystone, if any are present. The processing procedure converts the data received by the sensor into physically meaningful values such as reflectance. These steps are known as pre-processing, because they create data that will be further processed using dedicated algorithms to extract information related to the surface condition.

The electromagnetic energy that is reflected and emitted by the Earth's surface that reaches a remote sensing system is affected by some atmospheric gases (e.g., H_2O, CO_2) and scattering effects by particulates in the atmosphere. The signal reaching the sensor is recorded as a digital number (DN). It is consistent within the image and between spectral bands with respect to illumination angles and corrected for sensor artifacts such as keystone and smile. However, it is not consistent between different acquisition dates nor is it consistent between different sensors, (Franklin, 2001).

Converting the DN to at-sensor radiance helps to compare images from different sensors and different dates. This is accomplished by using calibration coefficients (published by the sensor manufacturer) to calculate the at-sensor radiance as follows (Teillet, 1986) :

$$L_s = a_0 + a_1 DN, \quad (2.1)$$

where L_s is the at-sensor radiance in (W m^{-2} μm^{-1} sr^{-1}), a_0 (offset) and a_1 (gain) are the sensor calibration coefficient and offset, and DN is the digital number.

To remove atmospheric scattering and absorption effects, correction algorithms and programs, such as the Moderate resolution atmospheric Transmission) (MODTRAN) code, are applied to the at-sensor radiance data. Through these algorithms, surface reflectance can be computed (Gao *et al.*, 2009; Liang *et al.*, 2001; Berk *et al.*, 1999; Richter, 1997). The conversion to reflectance produces a physically meaningful measure in order to compare images from different periods in time and under different atmospheric conditions. Once a physical measure for a pixel has been computed, remote sensing products that depict information about land condition, such as change over time or concentrations of chemical components, can be created.

Geometric alignment of pixels within an image is performed so that users can be sure that the data recorded by a sensor at a specific location is the same as if a user physically visited that location (Jensen, 2009). Geometric correction is also used to correct for topographical variation within a scene (Congalton, 1991a). Inherent in most sensors are artifacts, which affect the recorded signal. Examples include pixel dropout and image striping (Gómez-Chova *et al.*, 2008). The latter is found in Landsat

Multispectral Scanner (MSS), Landsat Thematic Mapper (TM), GOES Visible-Infrared Spin-Scan Radiometer (VISSR), SPOT pan, and Hyperion data (Sun *et al.*, 2008). Correcting for these artifacts requires algorithms, which look for patterns within the data and attempt to correct it by interpolating new pixel values from surrounding unaffected pixels (Gao *et al.*, 2009). The new pixel values then replace the noise-affected pixels. If the image is left uncorrected, improper and incorrect data can be used for image classification and for the retrieval of biophysical and biochemical parameters.

2.5.1 Hyperspectral Remote Sensing

Hyperspectral sensors, like CHRIS, typically have more than 40 spectral bands that are able to measure EMR in contiguous narrow spectral regions. Data from these sensors highlight subtle variations in the spectral response curves of a given material that would otherwise be hidden when using multispectral sensors (Lee *et al.*, 2004).

Due to the large number of spectral bands present in hyperspectral images and the fact that hyperspectral sensors often, but not exclusively, capture high-spatial resolution data, attention must be given to the processing steps involved. A high degree of sophistication and engineering goes into designing of hyperspectral sensors and with that a higher chance of sensor artifacts such as smile and keystone occurs. Smile is a property whereby the center wavelength of a band is shifted for each column in the image. Thus, a specific band in an image will not have the same wavelength and the image will be spectrally distorted (Neville *et al.*, 2008). Keystone is a property whereby a pixel within an individual band is spatially shifted and does not align properly with pixels in other bands, thus, spatially distorting the image (Yokoya *et al.*, 2010; Neville *et al.*, 2004).

In addition to the sensor properties that can cause problems in hyperspectral image processing, the higher spectral dimensional nature of hyperspectral imagery lends itself to issues, whereby the increased number of spectral bands for a target decreases the resolving power of a classifier (Chi *et al.*, 2008; Canty, 2006). Because there are more data available to a classifier and that data can be highly correlated, the amount of data needed to reliably determine a class for a pixel grows.

However, the subtle spectral variations present in hyperspectral imagery increase accuracies when compared to multispectral imagery as shown in the literature in studies investigating species composition (Dalponte *et al.*, 2013; Galidaki *et al.*, 2012; Goodenough *et al.*, 2003b). Furthermore, the subtle spectral variations present in hyperspectral imagery can help to identify vegetation health and stress.

2.5.2 Image Classification

Classification of remote sensing imagery is performed in order to place pixels with similar spectral characteristics into classes that are meaningful and of use to a user. Land-cover and land-use maps are remote sensing products, which are created using classification procedures (Foody, 2002). These products can be used by users in determining land-use policies and for decision making, such as wetland classification (Ozesmi and Bauer., 2002), Mountain Pine Beetle infestation (Wulder *et al.*, 2006), and grassland inventory (Xu *et al.*, 2014).

Image classification falls into three categories, supervised, unsupervised, and a hybrid that combines the former two. In a supervised classification system, the user with *a priori* knowledge of the area under investigation selects regions of interests (training samples) in the image that the analyst is confident belongs to a certain class (Jensen,

2009). The selected regions, or training areas, are then fed to the classifier to assign pixels in the rest of the image to the appropriate classes based on factors such as similarity between the pixel's DN value and the average class DN value.

An unsupervised classification system attempts to spectrally separate pixels and group them into classes without any user guidance (Lillesand *et al.*, 2004). This process is often run when there is not enough knowledge about an area and is sometimes used in order to investigate potential training areas for supervised classifiers. However, this creates a conundrum, because an unsupervised classifier may define two or more classes that are, from a user's perspective, in fact one class, but simply have a large degree of spectral heterogeneity. The bottleneck of unsupervised classification systems is the labelling of the classes. Because unsupervised classifications do not label classes, it is up to the user to apply meaningful labels either through the use of spectral libraries or through ground verification.

Iterative Self-Organizing Data (ISODATA) and K-means clustering are examples of unsupervised classifiers. Both of these classifiers are iterative processes whereby pixels are assigned to a class based on the closest class mean within the feature space. In ISODATA, a class mean will change at each iteration until a maximum number of iterations is run or a tolerance threshold, the smallest distance a pixel will move to a class mean, is reached. K-means requires clusters to be defined prior to running the algorithm (Duda and Hart, 1973), while ISODATA does not. ISODATA is a modification of K-means, but allows for classes to be split and merged at each step based on thresholds. For example, using ISODATA, a class can be split into two different classes if the standard deviation of the mean associated with the original class exceeds a defined threshold.

Alternatively, a class can be merged if the number of pixels within the class is below a predefined threshold (Ozesmi and Bauer., 2002). For example, a Balsam Poplar tree and an Aspen Poplar tree class could be merged to a Poplar tree class.

Classifying an image, as accurately as possible, is an objective of remote sensing projects. An accurately classified image allows for comparisons over time by observing how a pixel's class assignment changes. Accordingly, patterns in the distribution of a class can be realized. Classifications are also useful in determining the areal coverage of a class within a scene by summing the area of the pixels assigned to a certain class.

2.5.2.1 Support Vector Machine (SVM)

SVM is a supervised classification method that attempts to find the boundary, called the hyperplane, between classes with the least error possible (Huang *et al.*, 2002). Data points that contribute to the definition of the hyperplane are called support vectors and these are often the most difficult data points to classify. In remote sensing, this is accomplished by using training pixels that are close to the boundaries of a class in order to define that class. SVM is well suited to data that do not have a normal distribution and to reference data with a limited number of training pixels (Foody and Mathur, 2004 and 2006).

The caveat of using supervised and unsupervised classification methods is an assumption that a pixel only contains one tree species. In the case of medium to low-spatial resolution imagery, this is not always the case. In order to reliably determine the tree species at a reclamation site and to conform to any reclamation guidelines, other methods of image classification, such as spectral unmixing, may need to be explored.

2.5.2.2 Spectral Unmixing

Determining the proportions of spectral signatures that contribute to the spectral response of a pixel is often accomplished through unmixing algorithms, such as SMA (Adams et al.et al., 1993) and Multiple Endmember Spectral Mixture Analysis (MESMA; Roberts *et al.*, 1998). These two unmixing methods derive the relative abundance of features within a pixel based on the linear combination of spectral returns from pure targets, called endmembers (Figure 2.2).

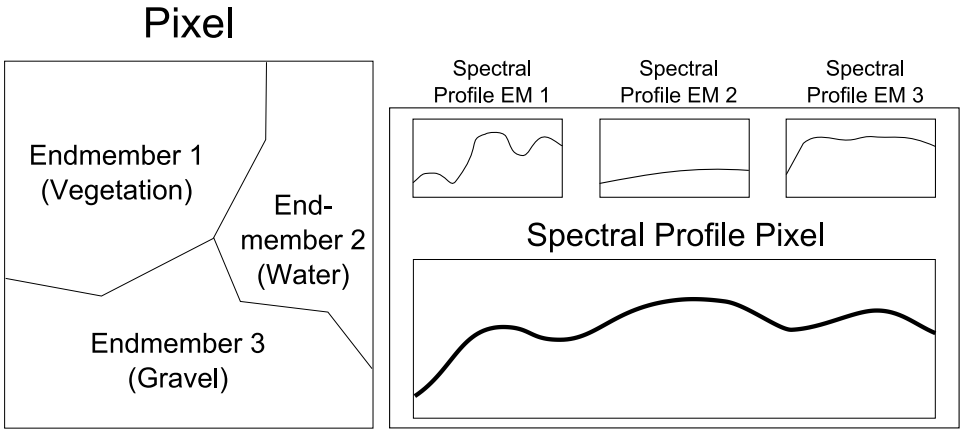

Figure 2.2: Representation of the physical constituents and spectral profile of a pixel with three contributing endmembers (EM).

SMA is a technique used to identify the proportion of a pixel that is occupied by a set of endmembers (Adams *et al.*, 1993). It assumes that each pixel to be unmixed within the image contains either some proportion of the endmembers selected by a user and/or an unknown constituent (Lu and Weng, 2004).

The SMA model is defined as follows:

$$R_i = \sum_{k=1}^{N} f_k R_{ik} + e_i \qquad (2.2)$$

where R_i is the reflectance of a spectral mixture for a given pixel in band *i*. It can be modelled as the sum of the reflectance R_{ik} of the k^{th} endmember in band *i* of *N* image endmembers. A weighting factor, f_k, represents the fraction of endmember *k*, and e_i is the band residual error.

The output from the unmixing approach is an image for each endmember that contains its abundances (fractions) for each pixel. The fractions that represent the endmembers can vary from an unconstrained model

$$-\infty < f < \infty \tag{2.3}$$

to a fully constrained model.

$$0 \leq f \leq 1 \tag{2.4}$$

where f is the fraction of an endmember. An unconstrained model allows for the sum of the fractions of endmembers that make up a pixel to represent more than 100 % of that pixel, while in a fully constrained model the fractions of endmembers that make up a pixel sum to 100 %.

An assumption when using SMA is that all image endmembers are accounted for and that the spectral homogeneity of the endmembers does not change across the spatial extent of the image (Roberts *et al.*, 1998). Problems arise when one or more of the endmembers are absent from a pixel or there is spectral variation within an endmember. The former can occur when utilizing a remotely sensed image with large ground sampling distances, such as the 30-m Landsat Operational Land Imager (OLI) or the 17-m CHRIS data. The latter can occur when selecting endmembers for discrete groups, such as a single-tree species. Individual trees can yield different spectral signatures even though they are of the same species. The different spectral signatures can be attributed to

differences in local environmental conditions, growing rate, topography and viewing/illumination geometries. Each of which will impact the interaction with incoming EMR determine the outgoing radiation in different ways.

An improvement has been made to the SMA algorithm through the development of Multiple Endmember Spectral Mixture Analysis (MESMA), which accounts for spectral differences within endmembers across the spatial extent of the image (Roberts *et al.*, 1998). Instead of using one model for the entire image, MESMA allows for multiple models to be used that contain different numbers of endmembers. For example, in a three-endmember model, one of the endmembers will consist of photometric shade, one endmember may consist of water and the third endmember may consist of vegetation. Within the water endmember, there will be sub-types (clear water, water with sediment, and water with algae) that are used to unmix the water endmember. And within the vegetation endmember the subtypes may be tree, grass, and shrub. Thus, a general outline is applied (shade, water, vegetation), but the endmembers within each category can vary. Using this approach, the spectral heterogeneity of the endmembers is accounted for. Multiple different models (combination of endmember sub-types) can be created to unmix the image, hence, the name MESMA. This method of spectral unmixing has been applied to mapping Chaparral in California (Roberts *et al.*, 1998), shrubs and grasses in New Mexico (Thorp *et al.*, 2013), and trees and shrubs in Ontario (Sonnentag *et al.*, 2007). However, its use in mapping boreal tree species is limited and, thus, further investigations are required (van der Sluijs, 2012).

Utilizing MESMA for reclamation monitoring requires that spectral signatures for each tree species be acquired. However, due to the confusing nature of spectral signatures

between similar tree species (Price, 1994), the use of a hyperspectral sensor is beneficial in order to satisfactorily discriminate tree species that may be present within a pixel (Dennison and Roberts, 2003a; Roberts *et al.*, 1998). This is accomplished by identifying subtle variations in the spectrum of tree species, which are masked when using a multispectral sensor with wider band widths.

2.5.3 Accuracy Assessment

Assessing the accuracy of classifications is an important step in the image classification chain. This can be accomplished through the use of an error or confusion matrix (Congalton, 1991a). In this step, pixels from the classified image are compared to known reference pixels. Accuracy measures are calculated including: overall accuracy, commission and omission errors, and kappa coefficient (Congalton, 1991b). These measures will be expanded on in Sections 3.4.4 and 3.5.4

2.5.4 LiDAR

LiDAR remote sensing systems are active systems that rely on emitted and, subsequently, reflected EMR in order to obtain height and in some cases, intensity data about a target (Dubayah and Drake, 2000). Intensity is the return strength of the LiDAR beam and is based, in part, on the reflectivity of a target. The information that LiDAR sensors acquire allows for the determination of the distance from the sensor to the target. LiDAR systems on-board an aircraft carry an Inertial Measurement Unit (IMU) and a Global Positioning System (GPS) in order to accurately measure the sensor's location (Lim *et al.*, 2003). The IMU is used to record airspeed, roll, pitch, and yaw, while the GPS is used to determine the latitude, longitude, and flying height of the aircraft.

Combining the IMU and GPS data with LiDAR measurements allow for the accurate 3-dimensional location of a LiDAR data point (Figure 2.3). A LiDAR system samples the ground by using a laser beam pulsed onto a mirror to scan the ground from side to side. The emitted light pulses interact with the ground and the number of light pulses that are returned to the sensor over a given time period and over a given unit area are referred to as the point density of the LiDAR system (Wulder *et al.*, 2008a). The point density is a function of the aircraft speed, elevation and instrument scan angle.

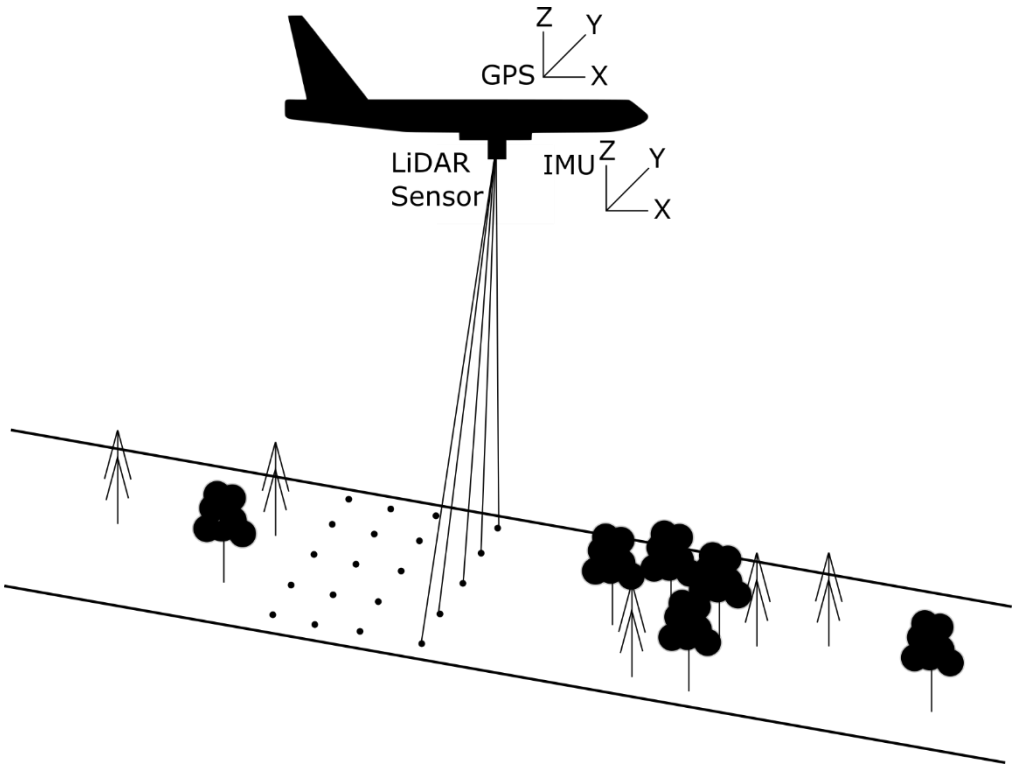

Figure 2.3: LiDAR sensor set up and data acquisition.

Two types of LiDAR systems can be employed, a large footprint waveform LiDAR or a small footprint discrete form LiDAR (Figure 2.4). Waveform LiDAR typically has a footprint between 15 m and 25 m and continuously records the returning laser (Mallet and Bretar, 2009). When the laser pulse interacts with structural components of the vegetation (leaves, branches, or gaps), a peak or trough is recorded on the

continuous pulse. This allows for a more complete picture of the structure of the vegetation, because the subtle changes as well as abrupt changes are recorded. Discrete form LiDAR, typically with a footprint size of 20 – 80 cm, only records the peaks of the returning light pulse (Evans *et al.*, 2009). The light pulse emitted from the sensor interacts with the vegetation at least once, depending on the complexity of the vegetation canopy. This creates differing return times for the same light pulse. The number of returns that are captured can be useful in determining the structure of vegetation. The first return is useful in creating surface models that depict canopy distribution (Dubayah and Drake, 2000). Subsequent returns can show the internal structure of vegetation and can help reveal vegetation density and canopy fractional cover (Hopkinson and Chasmer, 2009). The last return is useful in creating Digital Terrain Models (DTM) and Digital Elevation Models (DEM) (Maune, 2007).

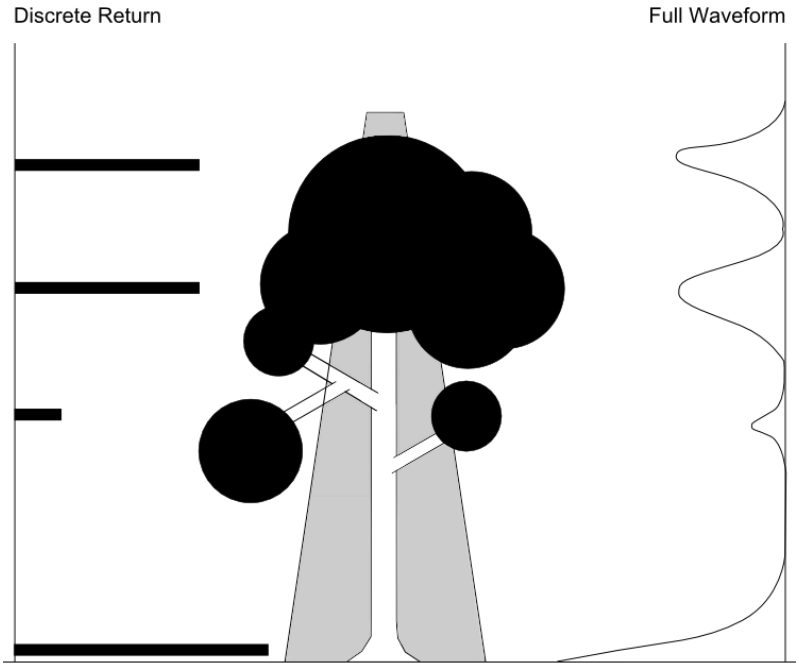

Figure 2.4: Discrete and waveform LiDAR pulses in a forest. The high peaks are captured as discrete points and the peaks and troughs represent the intensity of the return waveform LiDAR pulse from the vegetation.

Processing of LiDAR data incurs a high computational cost as the number of data points is often in the millions. For this reason, specialized algorithms are used to create accurate LiDAR data products (Brandtberg *et al.*, 2003; Kraus and Pfeifer, 2001). Two of the more commonly created products for forestry and environmental monitoring include canopy height and fractional cover.

2.5.5 Canopy Fractional Cover

Fractional cover is an important measure of forest structure that relates to tree photosynthetic productivity and health (Jennings *et al.*, 1999). Assessing fractional cover at a forest stand scale can help to show trends in health that can be used alongside other reclamation parameters to provide a more complete understanding of an oil or gas site's reclamation condition. Fractional cover, F_c, estimates are usually created by dividing the total number of LiDAR returns from the canopy above a specified height, R_{canopy}, by the total number of all LiDAR returns, R_{total} (Hopkinson and Chasmer, 2009). This can be summarized as follows:

$$F_c = \frac{\sum R_{canopy}}{\sum R_{total}}. \tag{2.5}$$

2.5.6 Canopy Height

Height measurements of forest structure can occur at either the individual tree level or at the stand level and are an important variable for species classification studies and forest productivity models because different tree species can have different average heights (Anderson *et al.*, 2008; Dalponte *et al.*, 2008; Plourde *et al.*, 2007; Popescu and Wynne., 2004). The direct measurement of height from LiDAR data occurs when a laser

pulse interacts with the tree canopy and returns to the LiDAR sensor. The height is derived by subtracting the ground elevation above a reference point, usually the ground surface (DEM), from the canopy height above a reference point (DSM), as follows:

$$H_c = DSM - DEM \ . \tag{2.6}$$

However, most LiDAR sensors do not acquire measurements with high enough point densities required to accurately measure the tree height and, subsequently, underestimate the true height of trees (Glenn *et al.*, 2011; Zimble *et al.*, 2003; Næsset and Økland, 2002; Popescu *et al.*, 2002). This is due to the pulses interacting with the tree without hitting its apex (Figure 2.5).

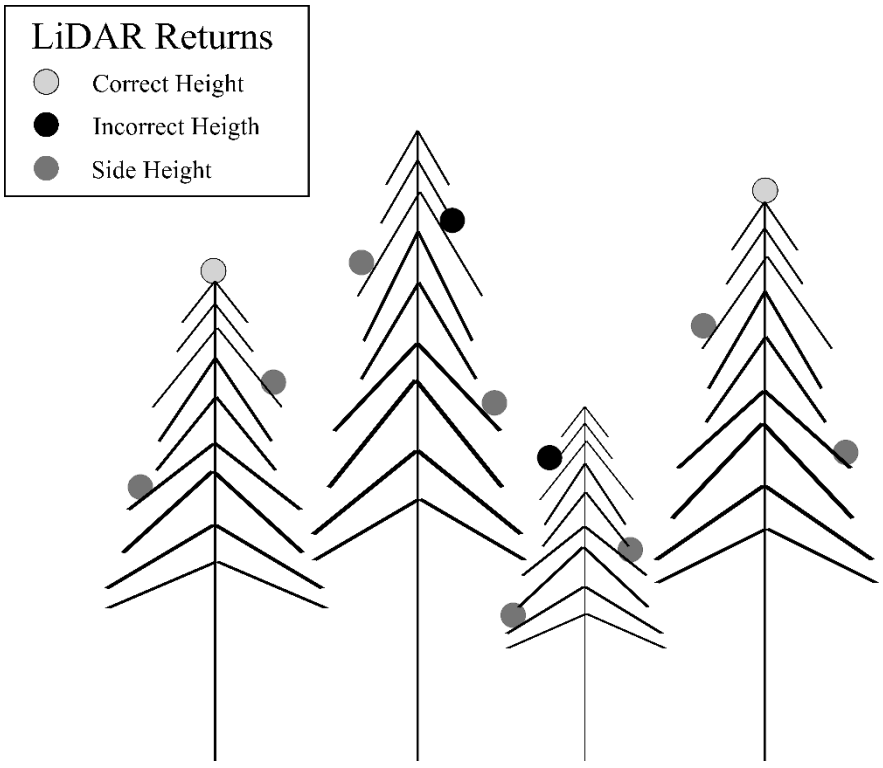

Figure 2.5: Interaction of LiDAR pulses and tree structure.

Holmgren (2004) conducted a study in which the ability of medium sized footprint LiDAR systems in estimating forest height was investigated. Specifically tree heights at

the plot level for a boreal forest in southwestern Sweden was explored. Using a basal-area-weighted model, the study was able to correlate LiDAR-derived mean tree heights to field- measured tree heights with a coefficient of determination, R^2, of 0.995. Another study, Næsset and Økland (2002), used a similar method and were able to achieve an R^2 between 0.47 and 0.91 for plot sizes ranging from 100 m^2 to 200 m^2. The large difference in the R^2 was attributed to different site locations and field measurement techniques used in the comparison with the LiDAR-derived height measurements.

In relation to reclamation monitoring, the tree height can be an important variable in attempting to distinguish tree species (Puttonen *et al.*, 2010) and in assessing whether tree growth over time meets expectations.

Remote-sensing-based information about forest structure can reveal information about the health of forests through direct and indirect measures and constitutes an important variable in forest productivity models (Wulder, 1998; Running *et al.*, 1994). LiDAR remote sensing in particular is better able to provide structural information than passive optical remote sensing as it collects information in three dimensions. The information it provides has shown to be a good indicator of forest health (Solberg *et al.*, 2004; Lim *et al.*, 2003).

2.6 Summary

Understanding the history of reclamation and how it is applied to anthropogenic disturbances today, helps to guide monitoring efforts that deal with the long-term effects of a disturbance. A successful reclamation plan is the end goal of a cumulative environmental monitoring program, but how the reclamation plan is carried out varies depending on local customs and requirements. Nevertheless, the general steps, such as

assessing a reclamation site, performing the reclamation and finally monitoring the reclamation, are universal. Because Alberta is a province rich in natural resources, reclamation activities within the province are geared towards the natural resource extraction industry, specifically oil and gas.

Remote sensing is uniquely positioned to be of benefit to the reclamation industry as it is able to provide large-area coverage in a short time period and has the ability to revisit a site over and over again. The literature regarding remote sensing applications for environmental monitoring is vast and its outcomes can be utilized for reclamation monitoring. The processing steps required to produce meaningful information from the remotely sensed data are well documented, but have not been readily applied to reclamation monitoring in Alberta's oil and gas sector. Applying techniques like MESMA and SVM to tree species classification will assist reclamation efforts by documenting the occurrence of particular tree species and will help reclamation managers in determining the corrective course of action if necessary.

Chapter 3

Methods

3.1 Introduction

This chapter will provide the methods by which species mapping can be accomplished in the Cold Lake area in support of reclamation monitoring. Hyperspectral remote sensing will be assessed alone and in conjunction with LiDAR remote sensing to meet this objective.

3.2 Study area

The study area is located west of Cold Lake, Alberta in the Cold Lake – Beaver River Basin (Figure 3.1). The area lies within the Boreal Forest Natural Region and is characterized by upland forests with wetlands in lower elevation areas. Elevation ranges from 500 m – 700 m above sea level with the topography generally flat with a slight undulation. The surficial geology is composed of glacial derived lacustrine and glacial till on top of cretaceous shale bedrock. The soils in the region are comprised of Luvisolic soils in forested areas and Gleysolic soils in wetland areas (Natural Regions Committee, 2006). The climate is humid continental, averaging approximately 322 mm per year of precipitation (Environment Canada, 2013).

The distribution of tree species within the area follows topographical, environmental, and anthropogenic controls, such as drainage areas, soil conditions, and reclamation areas. Mixed forests of Trembling Aspen (*Populus tremuloides Michx.*), Balsam Poplar (*Populus balsamifera L.*), and White Spruce (*Picea glauca (Moench) Voss*) comprised the upland forested areas, while Black Spruce (*Picea mariana (Mill.)*

BSP) and Tamarack (*Larix laricina (Du Roi) K. Koch*) populate the low-lying wetland areas. Jack Pine (*Pinus banksiana Lamb.)* and Lodge Pole Pine (*Pinus contorta Loudon*) can also be found, which are common in well-drained sandy soils (Natural Regions Committee, 2006).

The study area sits on top of the Cold Lake oil sands and as a result, ongoing oil and gas operations (well pads, pipeline right-of-ways, lease roads) are within the study area boundaries.

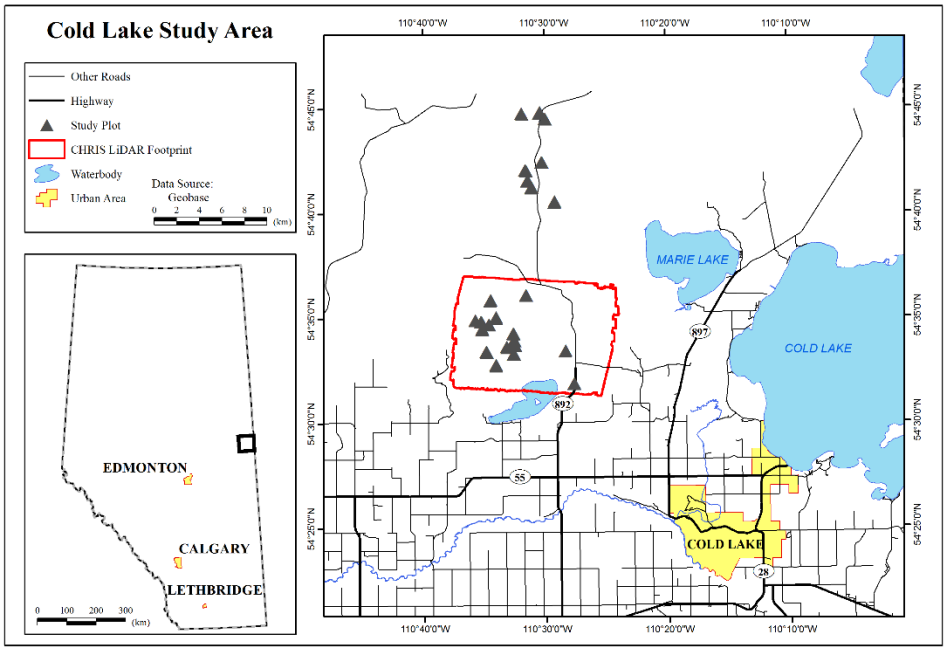

Figure 3.1: Cold Lake study site. Geographic center coordinates of the test site are 54°38'8.86" N, -110°31'35.186" W.

3.2.1 Test Plot Selection

The selection of test plots was directed by the location of certified reclaimed well-sites within the Cold Lake area. Two test plots were set up for each reclaimed well-site, one on the well-site and one at a minimum of 60 m (reference site) away from the well-

site boundary. An average well-site is approximately 100 m x100 m and has compacted gravel as a base (Figure 3.2).

Figure 3.2: Example of an average well-site with a small facility (in red) to the left of the wellhead (in black). A representation of the well-site is to the left of the image with an example of a well-site plot (bottom black square) and reference plot (top black square) (not to scale).

A 60-m distance was chosen so that the reference sites were at least one pixel away from the edge of well-sites when viewed using CHRIS satellite imagery. The Alberta Vegetation Inventory (AVI; AESRD, 2012a) and well reclamation certification dates were used to ensure that different forest structures and reclamation ages were sampled.

3.2.2 Field Sampling Method

The field protocol adopted for this study is based on guidelines established by Canada's National Forest Inventory (CFS, 2008). Test plots of 20 m x-20 m were established, and the geographic corners and centers of the plots were surveyed using a handheld GPS device. The vegetation characteristics for both over-storey and under-storey were recorded (Figure 3.3). When the test plot was located in a mature tree stand, a

10-m x 10-m subplot was established for tree characterization. Tree height, Diameter at Breast Height (DBH), and species were recorded for every tree within the subplot and the number of trees (density) within the plot was tallied. For young regenerating tree (below 1.3-m), mostly located in reclaimed well-sites, a 7-m x 7-m subplot was established whereby the tree structure was characterised using the same methods as for the 10-m x 10-m plot. A total of 24 certified reclaimed well-sites and 24 reference sites were characterized during the field campaign. However, only 32 sites in total were within the CHRIS image and airborne LiDAR data coverage. The original intent was to use airborne imagery covering all 48 sites. However, due to image delivery delays, the CHRIS data were used exclusively and only covered 32 sites. Field data were acquired in order to create accurate training and validation samples for the remote sensing applications like spectral unmixing and image classification techniques used in this study.

Figure 3.3: Land-cover, species, and structural characteristics of a reclaimed well-site (left) and reference site (right). The well and reference sites are dominated by Trembling Aspen with a Balsam Poplar mix with the reference site having considerable undergrowth.

Hemispherical photographs were taken at all sample sites following the Validation of Land European Remote sensing Instruments (VALERI) Protocol (VALERI, 2014) in order to determine canopy fractional cover at each site. For forested areas, two sets of 13

images were collected with the lens facing downward and upward to characterize the understorey and overstorey, respectively (Garrigues *et al.*, 2002). The images were spread around the field site as demonstrated in Figure 3.4.

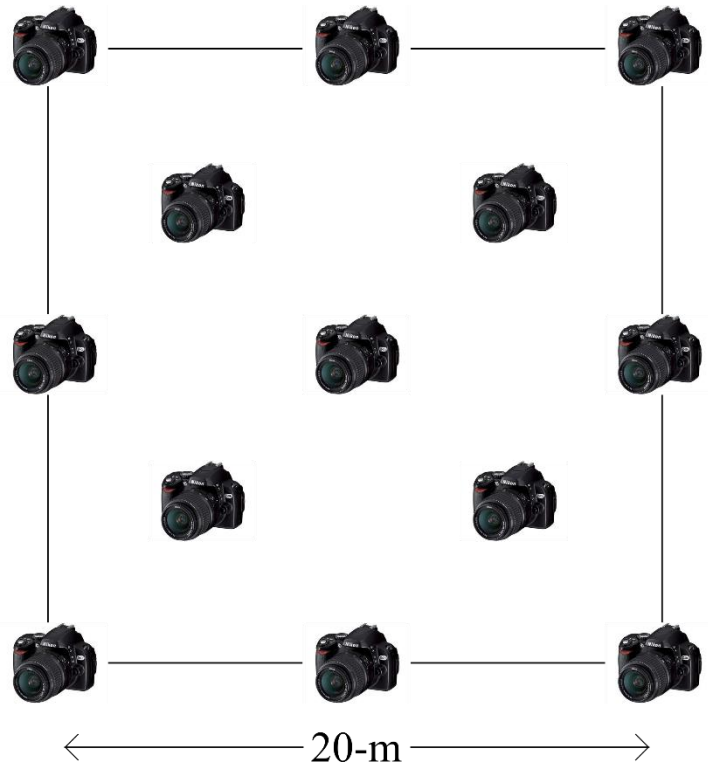

Figure 3.4: Hemispherical photograph locations acquired in a 20-m plot using the VALERI protocol. Images were captured in a North/South and East/West orientation.

The hemispherical photographs were processed using the CAN-EYE software (INRA, 2013), following guidelines established in the user manual (Weiss and Baret, 2010). Photographs were initially masked when operators or field equipment were visible in the image. Training pixels were used to grow regions that corresponded to vegetation, non-vegetation, and sky (or soil), which were then visually assessed to ensure they fell into the correct class (Figure 3.5).

Figure 3.5: Hemispherical photograph (left) and classified image using the CAN-EYE software (right); black corresponds to tree, white to sky or soil (depending on the orientation of the image), and grey is mixed.

The gap fraction was calculated within a solid angle defined by the viewing zenith and azimuth angle intervals for each image. The gap fractions around the nadir direction derived from the 13 processed images were then averaged for the entire plot.

The fractional cover, *fCover*, was calculated as follows:

$$fCover = 1 - P(0)\,, \tag{3.1}$$

where P(0) is the gap fraction at nadir. In the case of canopy-cover estimations, the gap fraction for a given viewing zenith and azimuth angle is the probability that a beam of light will penetrate the canopy without interaction with vegetation elements (leaves, branches). It is calculated using a Poisson model as follows:

$$P(\theta) = \left(\frac{-G(\theta,\alpha)PAI_{eff}}{cos\theta}\right), \quad (3.2)$$

where $P(\theta)$ is the gap fraction at the viewing zenith angle θ. In this case, the zenith angle varies between nadir (0°) and the edge of the processed hemispherical image (57°). $G(\theta,\alpha)$ is the leaf angle projection and θ and α are the average leaf angle and zenith angle respectively. PAI_{eff} is the effective plant area index (Cook *et al.*, 2009). The CAN-EYE software uses the 0-10 zenith angle interval to calculate *fCover* (Weiss and Baret., 2010).

A simple triangulation using an inclinometer and measuring tape was used to estimate the height of each tree above approximately 2-m (above the height of an extended hand). Below 2-m, the height was measured using a tape measure, whereby one end was held against the ground and the other measured to the tallest woody part of the tree. The species composition within each plot was determined by counting the number of trees per species within the plot boundaries.

3.3 Remotely Sensed Data

Remote sensing data, including airborne and satellite hyperspectral imagery as well as airborne LiDAR, were captured for this study and used to identify tree species distribution. Techniques used to derive information from the LiDAR and hyperspectral data will be discussed including the steps required to process the data into a useable form prior to information extraction.

3.3.1 Ancillary Data

Airborne hyperspectral imagery was captured at the same time as the LiDAR data using the Airborne Imaging Spectrometer for Applications (AISA) sensor. However due to late delivery, the hyperspectral data were only used for visual assessment in this study. Data preprocessing, georeferencing and atmospheric correction, were performed on the data by the data provider, Centre for Applied Remote Sensing, Modelling, and Simulation (CARMS)/University of Victoria. LiDAR data captured in 2006 and licensed through Airborne Imaging[1] by AESRD was used to investigate and highlight structural differences between 2006 and 2012. The LiDAR processing method discussed in the next section was applied for both 2006 and 2012 LiDAR data.

3.3.2 Airborne LiDAR Data

Airborne LiDAR data were captured on August 11th and 12th, 2012 over Cold Lake, Alberta at an altitude of 1500 m above ground. An area of approximately 270 km^2 was imaged with a relative horizontal accuracy of 15 cm. The average point density of the LiDAR data was approximately 2.27 pts/m^2. The data were delivered in tiles from CARMS/University of Victoria with geometric corrections and outlier reduction procedures already applied. Thus, no preprocessing was performed on the LiDAR tiles. From the captured LiDAR data, a bare earth model, a canopy height model, and a canopy fractional cover model were created.

[1]Airborne Imaging Inc. <http://airborneimaginginc.com/>

3.3.2.1 Bare Earth Model

A bare earth model was created from the processed LiDAR point cloud data using the FUSION/LDV software (McGaughey, 2007). The ground-filter tool, an algorithm developed by Kraus and Pfeifer (1998), was used to estimate the bare earth by means of a linear prediction.

$$P_i = \begin{cases} 1 & v_i \leq g \\ \frac{1}{1+(a(v_i-g)^b)} & g < v_i \leq g + w \\ 0 & g + 2 < v_i \end{cases}, \quad (3.2)$$

where P_i is the weight to be applied, v_i are the residuals from the terrain and vegetation points, *g* is a shift value that applies a weighting to a point in order to pull it to a surface, *w* is a threshold value that determines whether a point will have an influence on the surface, and *a* and *b* determine the steepness of the weighting function.

It first creates a surface that is between the last recorded point and the first recorded point. Applying user defined parameters, the algorithm iteratively computes a new surface by applying weights to the data points that pull the surface down if it is above a certain height and raise the surface if it is below a certain height.

Default values of 1.0 and 4.0 were used for the *a* and *b* parameters. The implementation of the algorithm within FUSION has default parameters that have been shown to work well for most surfaces (McGaughey, 2007). However, for this study the *g* parameter and the *w* parameter were changed to -0.1 and 0.5 respectively (numbers are unitless). These values were determined empirically in order to produce a surface that matched the 2006 LiDAR data.

The 2006 bare earth LiDAR data were assumed to be correct as the data were provided, already processed, by AESRD. The algorithm was iterated for 5 turns and a 3 x

3 smoothing filter was applied on each run to produce the final bare earth model. The model was then transformed to a raster with 2-m spacing, because the average point density was 1.6 pts/m^2.

3.3.2.2 Canopy Height Model (CHM)

The CHM was created by subtracting the bare earth elevation points from the first return of the raw LiDAR point cloud, normalizing the canopy heights to a common baseline. This allows the model to report vegetation heights as if they were measured on the ground instead of reporting them in metres above sea level. For this experiment, a 17-m pixel was used rather than the 34 m pixel of the Canopy Surface (CS) and CHRIS datasets in order to better capture the maximum vegetation height present on the ground. A similar approach was used by Næsset (1997) to determine tree stand heights in Norwegian forests using grid sizes of 15 m – 30 m. A 17-m pixel was used, because it can easily be resampled to 34-m in order to be integrated with the CHRIS and CS datasets. This was accomplished by aggregating four 17-m pixels and averaging their cell values to create the new 34-m cell value. The resulting height values were stretched between 0 and 1 using the minimum and maximum tree heights in the entire dataset. This was done in order to match the scale of the CHRIS dataset and the CS dataset.

3.3.2.3 Canopy Surface (CS)

The CS creates a surface from the height normalized LiDAR point cloud produced from the CHM. It was used to estimate the canopy closure of a surface that varies between 0 % and 100 %. The model calculates canopy closure by dividing the total number of LiDAR returns above a height threshold by the total number of returns in an

area. For this study, a height break of 1.3 m was used during processing in order to filter out returns related to low grasses and shrubs. A 34-m raster with was created in order to align the canopy closure map with the GSD of the CHRIS imagery. The *canopymodel* function in the FUSION/LDV software was used to create this product.

3.3.3 Spaceborne Hyperspectral Data

ESA's experimental sensor, CHRIS, was used to collect hyperspectral data for this study. Image data capture coincided with field-data acquisition in order to have similar environmental conditions in both datasets. CHRIS image data were collected in mode 1 on July 20th, 2012 at 15:36 UTC. The top left image coordinates are 54°38'52.86" N, 110°35'56.92" W with the bottom right coordinates being 54°29'54.81" N, 110°26'10.63" W (Figure 3.6). The specific CHRIS mode 1 sensor characteristics are listed in Table 3.1.

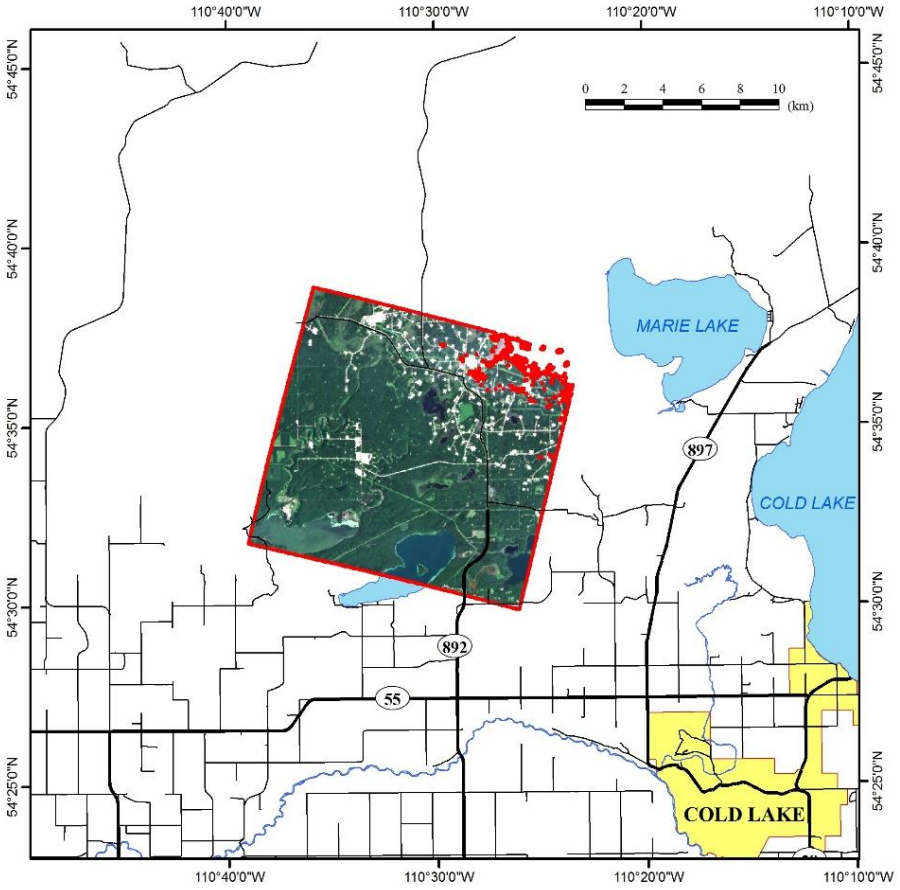

Figure 3.6: Extent of CHRIS imagery in the study area.

Table 3.1: Mode 1 sensor characteristics for the CHRIS sensor onboard the PROBA satellite.

Bands	62
Spectral Range (nm)	410-1050
Spectral Resolution (nm)	1.3 – 12.0
Spatial Resolution (m)	34
Temporal Resolution (days)	7
Swath Width (km)	14
View Angles	5 (-55°, -36°, 0°, 36°, 55°)
Altitude (km)	598

Due to time constraints and study scope goals, only the nadir (0°) viewing angle imagery was selected for further processing after noise removal. Images were processed using the ESA BEAM toolbox and developing platform (Brockmann-Consult, 2013).

3.3.3.1 Noise Removal

The CHRIS image was first processed to remove vertical striping inherent in the CHRIS data (Figure 3.7). This was accomplished through a noise reduction tool in BEAM that identifies pixel dropout and vertical striping. The dropout pixels were corrected using a weighted average of neighbouring pixels, where the weighting factor is the inverse of the spectral Euclidean distance between the neighbor pixels and the pixel to be corrected. CHRIS is known to have striping issues present in the imagery (Gómez-Chova *et al.*, 2008). All five viewing angles were used in the noise reduction process as the multiple viewing angles allow the de-striping algorithm to better estimate the contribution of actual surface values to the overall image. Because each viewing angle captures a different image at a slightly different time, any striping present manifests as an artifact at the same pixel position and not as a feature on the ground. The algorithm identifies these artifacts as noise thereby removing them and de-striping the image (Brockmann-Consult, 2010).

Figure 3.7: CHRIS radiance RGB (bands 24, 7 and 12) imagery captured in mode 1 over Cold Lake, Alberta. Major striping is noted by the white arrow.

A probabilistic cloud mask was then applied to the de-striped image. The cloud-screening tool was run with a target output of 20 clusters iterated 50 times. From the clustered image, one class was identified by the user as being cloud, and a cloud probability mask was then created for each pixel in the image. Additionally, a fully constrained linear unmixing was performed on the image that assigns each pixel a probability value that it is cloud. The probability and the fractional value were then multiplied on a pixel-by-pixel basis to form a cloud product. A threshold of 0.05 was applied to the cloud product to create a cloud mask that was used for subsequent processing. This means that any pixel with a cloud product value greater than or equal to 5 % cloud would be removed.

3.3.3.2 Atmospheric Correction

Following the cloud mask, an atmospheric correction was applied to the image data using the default parameters in the CHRIS toolbox provided in BEAM. This

software uses the MODTRAN 4 radiative transfer code to perform the atmospheric correction assuming a constant Aerosol Optical Thickness (AOT) across the image. When using the default parameters present within BEAM, a look-up table (LUT) is used to approximate the AOT at 550-nm. A value of 0.0 was used as an input within BEAM, resulting in a look-up table AOT value of 0.05 (Guanter *et al.*, 2005). The Columnar Water Vapour (CWV) retrieval uses a band-fitting technique around the atmospheric water absorption feature at 940-nm to estimate the CWV for each pixel. A value of 1.0 g/cm^2 was used to estimate the CWV (Guanter *et al.*, 2006).

Once the AOT and CWV are retrieved, a surface reflectance image was created using the values present in the LUT. Correction of adjacency effects and spectral polishing are completed on the image at this time as well. Adjacency correction is performed in order to reduce image blurring caused by scattering of photons by the atmosphere from adjacent or neighbouring pixels. Spectral polishing using a third-order polynomial transformation is applied to atmospherically corrected data to remove the residuals left behind from atmospheric correction as well as sensor artifacts present within the CHRIS instrument. For example, these artifacts can be caused by random noise effects due to changes in sensor temperatures (Gómez-Chova *et al.*, 2008). Spectral polishing attempts to remove errors in the apparent reflectance data by smoothing out the spectral profile.

After these corrections were performed using the BEAM software, it was noticed that reflectance values in the blue region were negative; subsequently, five spectral bands with negative values were removed from further processing.

3.3.3.3 Geometric Correction

A geometric correction was applied to the CHRIS imagery using an image-to-image registration technique. This is to ensure that locations identified in the CHRIS image are the same location identified during field work. The base image used was an airborne hyperspectral image with a 2-m GSD captured at the same time as the LiDAR data. Due to the fact that the AISA data were delivered geometrically corrected, it was decided to use it as the base image to correct the CHRIS data. Ground Control Points (GCP) were selected from the AISA data using bright targets such as well pads and facilities. The CHRIS image was registered to the AISA image using 52 ground control points spread across the image with a Root Mean Squared Error (RMSE) of 0.48 pixel or ~17-m. A first-order polynomial transformation was then applied to the CHRIS imagery in order to correct it. The result of the noise removal, atmospheric and geometric corrections are presented in Figure 3.8.

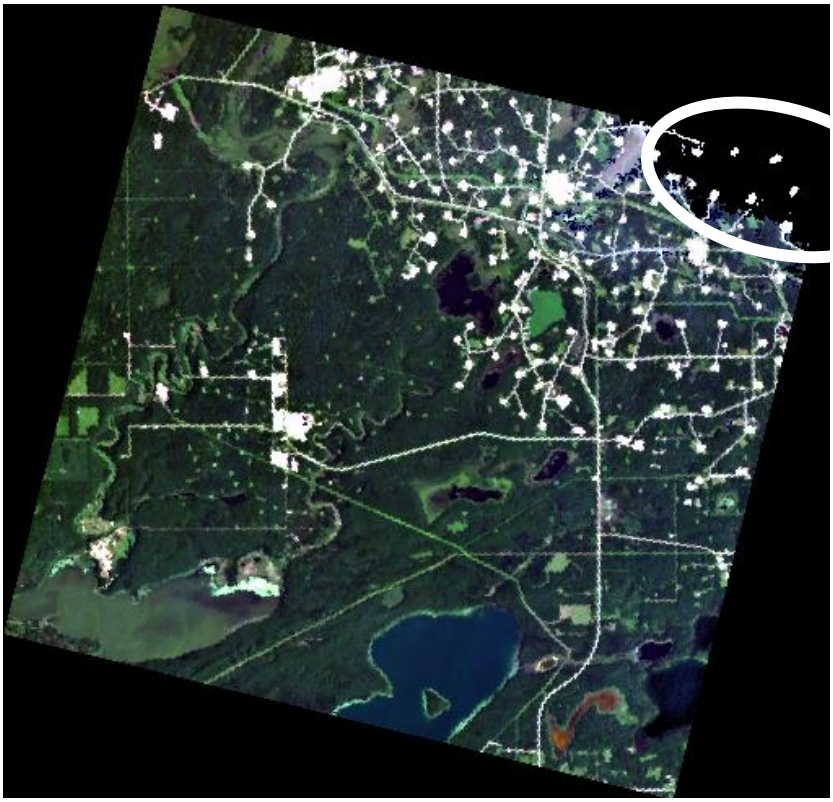

Figure 3.8: Processed CHRIS reflectance imagery (bands used, R:17 G:8 B:1). Note the cloud removal in the upper right of the image, highlighted by the white oval.

3.3.3.4 LiDAR/Hyperspectral Fusion

Because the CHRIS and LiDAR datasets cover different spatial extents, a polygon was created that encapsulated the extents of both to ensure that any area under investigation in this study would have both spectral and structural information. This polygon was then used to mask the CHRIS and LiDAR images before stacking them together to create a LiDAR/hyperspectral fused image. The resulting image was a 59-band product at 34-m ground resolution.

The canopy cover and height products were appended onto the CHRIS image and assigned wavelengths of 1002-nm and 1003-nm, respectively. These values were chosen

to ensure that the LiDAR structural data would not break up the spectral profiles of the CHRIS imagery.

3.3.4 Classification Hierarchy

A classification hierarchy was employed for land-cover and tree-species mapping. The hierarchy was divided into the three levels, land cover, tree type, and tree species (Figure 3.9). The initial land-cover level was created in order to produce a mask of non-forested areas so that subsequent classifications would only be conducted on the forested areas. This classification level was loosely based on the Earth Observation for Sustainable Development (EOSD) harmonized land-cover product (Wulder *et al.*, 2008b), where guidelines for classes such as wetland and herb were used to distinguish land-cover types in the image. The un–masked forested areas, level 2, were split into coniferous, deciduous, or mixed tree type classes based on field and AVI polygon data. From the level 2 classes, six tree species classes were created for use in the SVM classification and MESMA procedure.

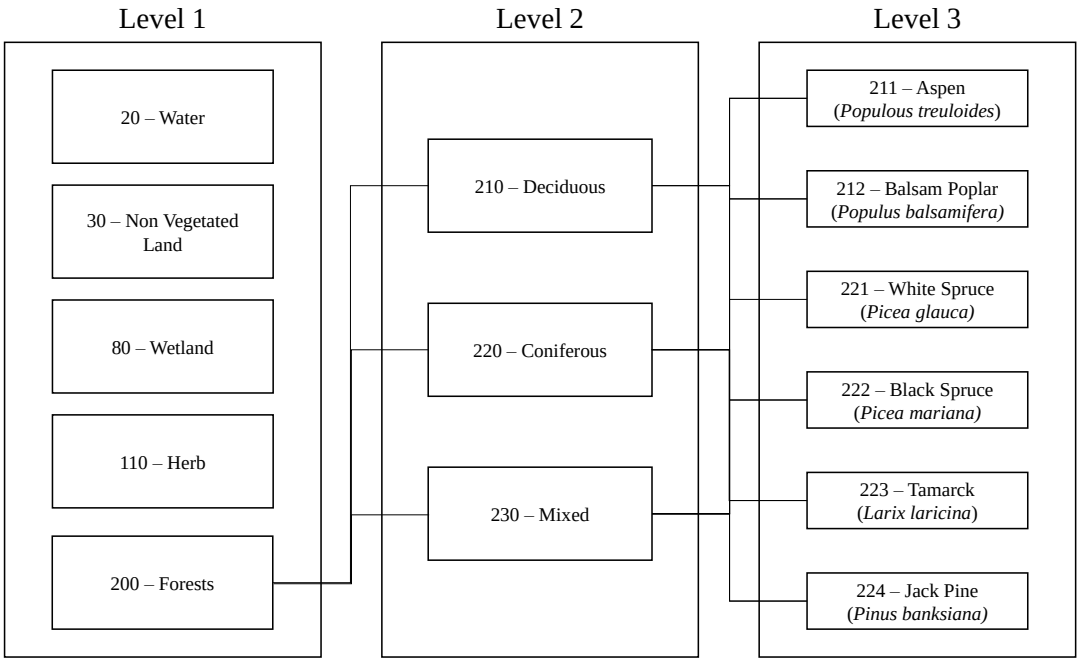

Figure 3.9: Classification hierarchy used. Level 1 is based on EOSD classes.

Classification was initially performed using the classes defined in level 1 on CHRIS imagery data alone and a forest mask created for subsequent processing using the SVM classifier. Level 2 classifications were performed using SVM wherein the forestry masked data was classified into coniferous, deciduous, and mixed tree types for the CHRIS data alone and the CHRIS/LiDAR fused data. Coniferous, deciduous, and mixed tree masks were then created and applied to the data.

Level 3 classifications were conducted using SVM and MESMA on tree-type masked data. Four species classes were used for the coniferous data, while two species classes for deciduous data and all six species classes were used for the mixed data.

3.3.4.1 ROI Selection

Regions of Interest (ROI) for classification levels 1 to 3 were created by selecting pixels based on the greatest level of detail required (tree species) and then up-scaled to

levels with lesser detail, such as tree-type and land cover. The selection of the pixels within the ROI's was accomplished by using AVI polygons to help select land cover and was refined using the captured AISA imagery to capture homogeneous areas in the CHRIS imagery for the coniferous and deciduous classes and heterogeneous areas for the mixed class.

AVI polygons with a reported species abundance equal to or greater than 80 % were used for the coniferous and deciduous classes, while polygons with less than 80 % were used for the mixed class. A threshold of 80 % of a certain tree species was deemed an acceptable accuracy and is the threshold used for determining an acceptable accuracy of photo interpretation used for AVI classification (AESRD, 2012a). AVI species abundances for polygons were determined by an aerial-photo interpreter, estimating species composition based on canopy closure (AESRD, 2005). In order to ensure that ROIs represented the tree canopies and not the ground cover, only AVI polygons with a canopy cover greater than 80 % were used to create the classes. Data gathered from field plots were also used to assist in the selection of pixels for the ROIs. Training pixels were defined by randomly selecting 70 % of the total ROI for each level. The remaining 30 % of the pixels were used for validation purposes.

ROIs followed the labelling schema of L1_L2_L3, which correspond to the EOSD land-cover thematic classes, tree-cover type and tree-species type, respectively. For example, F_D_Pb would equate to forest land-cover type, deciduous tree type and Balsam Poplar tree species. Lower levels of classification were aggregated to higher levels, depending on the level of classification being performed (e.g., for level 1 , 2, and 3 processing, F_D_Pb would be classified as F, F_D and F_D_Pb, respectively).

3.4 SVM

SVM was chosen for this study, because it has been shown to have superior accuracy when compared to unsupervised classification methods, does not require an assumption of data normality, and can use a small training dataset (Mountrakis *et al.*, 2011; Bartzen *et al.*, 2010; Foody and Mathur., 2004). This classification approach was used in all levels of processing. For each classification level a grid search was employed using the ENMAP-Box image SVM software (Rabe *et al.*, 2014) to find the optimal parameters with which to run the searches. A Gaussian radial basis function kernel was used to perform the classification.

3.4.1 Level 1 Classification

A minimum of 250 pixels were selected across the image to represent each class (Table 3.2 and Figure 3.10). The pixels were selected using criteria identified above (Section 3.3.4.1. ROI Selection). A mask was created that concealed anything not classified as forest from further classifications.

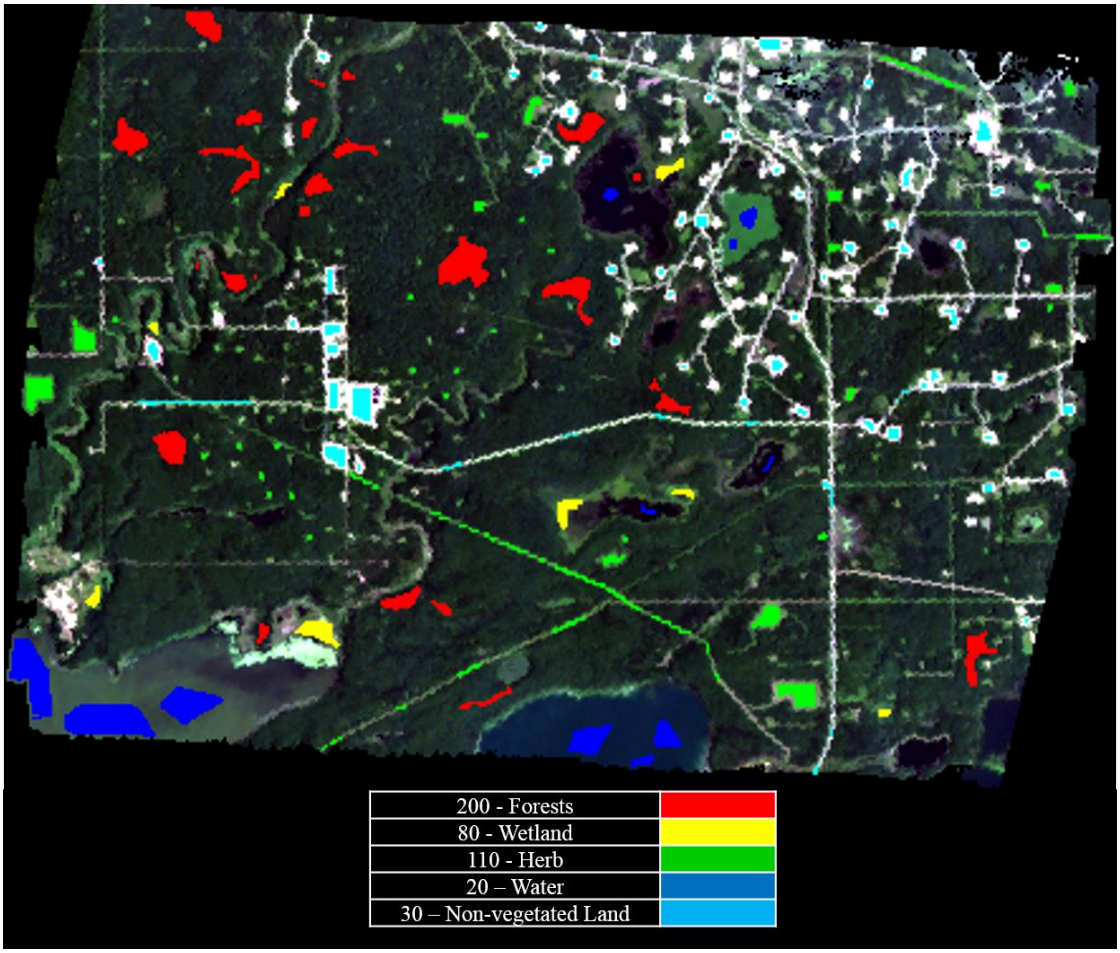

Figure 3.10: Distribution of Level 1 ROIs across the CHRIS image. A linear 2-% stretch has been applied (bands used: R:17 G:8 B:1.)

Table 3.2: Land-cover classes and SVM training and validation pixels.

EOSD Class Name	Description	Training Pixels	Validation Pixels
200 - Forests	Coniferous, Deciduous, Mixed	1123	482
80 - Wetland	Marsh, Low shrub	179	78
110 - Herb	Grassland, Agricultural, Pasture	598	257
20 – Water	Lakes rivers, streams, ponds	857	368
30 – Non-vegetated Land	Gravel, Roads, Buildings	476	205

3.4.2 Level 2 Classification

Tree-type classification was performed using the SVM classifier to separate the coniferous, deciduous, and mixed trees in the forest-masked CHRIS and CHRIS/LiDAR images. Following the ROI selection established in the previous section, 70 % of the total pixels were used for classification, with the remaining 30 % used for validation (Figure 3.11 and Table 3.3). The same pixels for training and validation were used for the CHRIS and CHRIS/LiDAR data. After classifications were conducted, masks were created to conceal coniferous, deciduous, and mixed tree areas for performing species classifications on these individual areas separately.

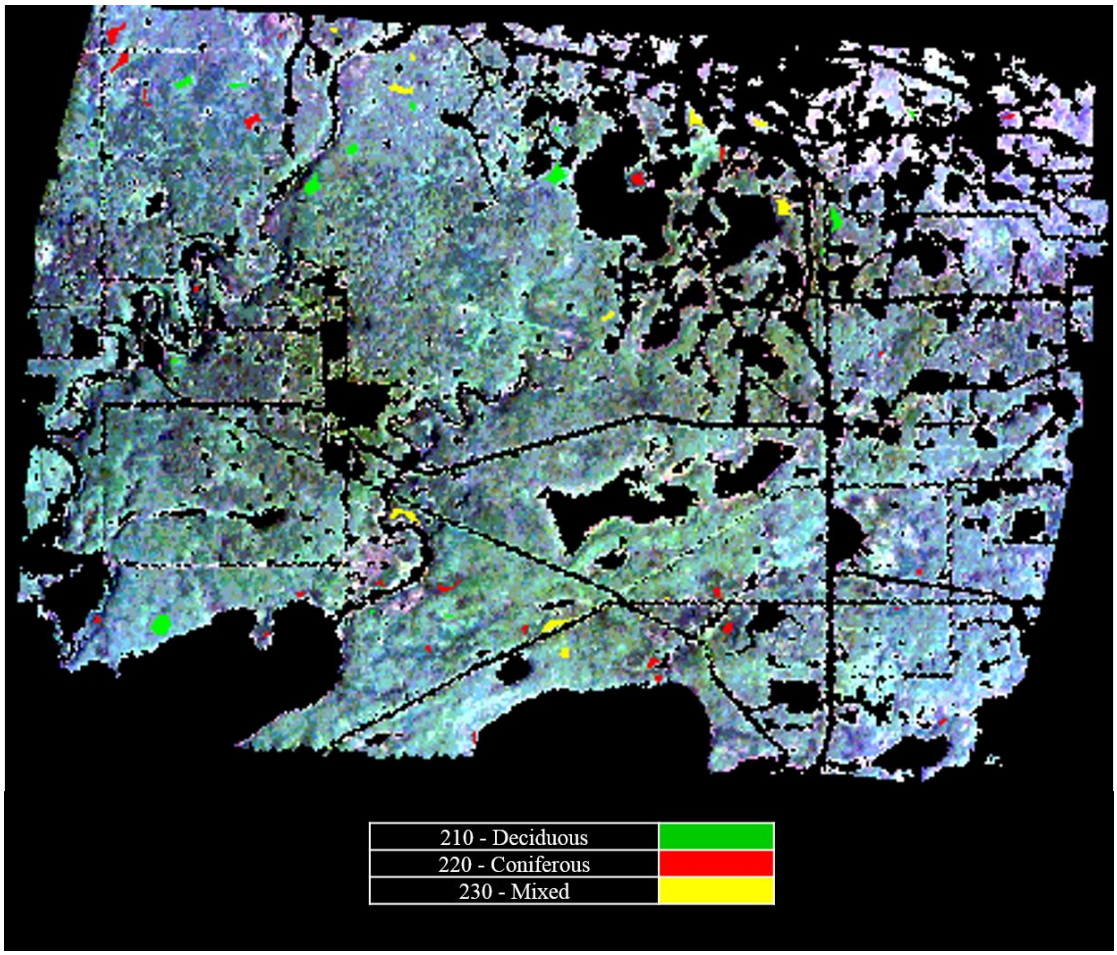

Figure 3.11: Distribution of Level 2 ROIs across the forest-masked CHRIS image. A linear 2-% stretch was applied, resulting in the spectral variations within the forested areas becoming more pronounced (bands used: R:17 G:8 B:1).

Table 3.3: Tree-type classes and number of SVM training validation pixels.

Class Name	Description	Training Pixels	Validation Pixels
210 – Deciduous	Coniferous tree species	908	390
220 – Coniferous	Deciduous tree species	794	341
230 – Mixed	Mixed tree species	234	103

3.4.3 Level 3 Classification

Species classification was performed for each tree-type masked image, for both data sets (Figure 3.12 Figure 3.13). This resulted in a total of six classifications being performed. Again a 70 % / 30 % split for training and validation pixels were used and the

same ROIs were applied for the CHRIS and the CHRIS/LiDAR data (Table 3.4 andTable 3.5).

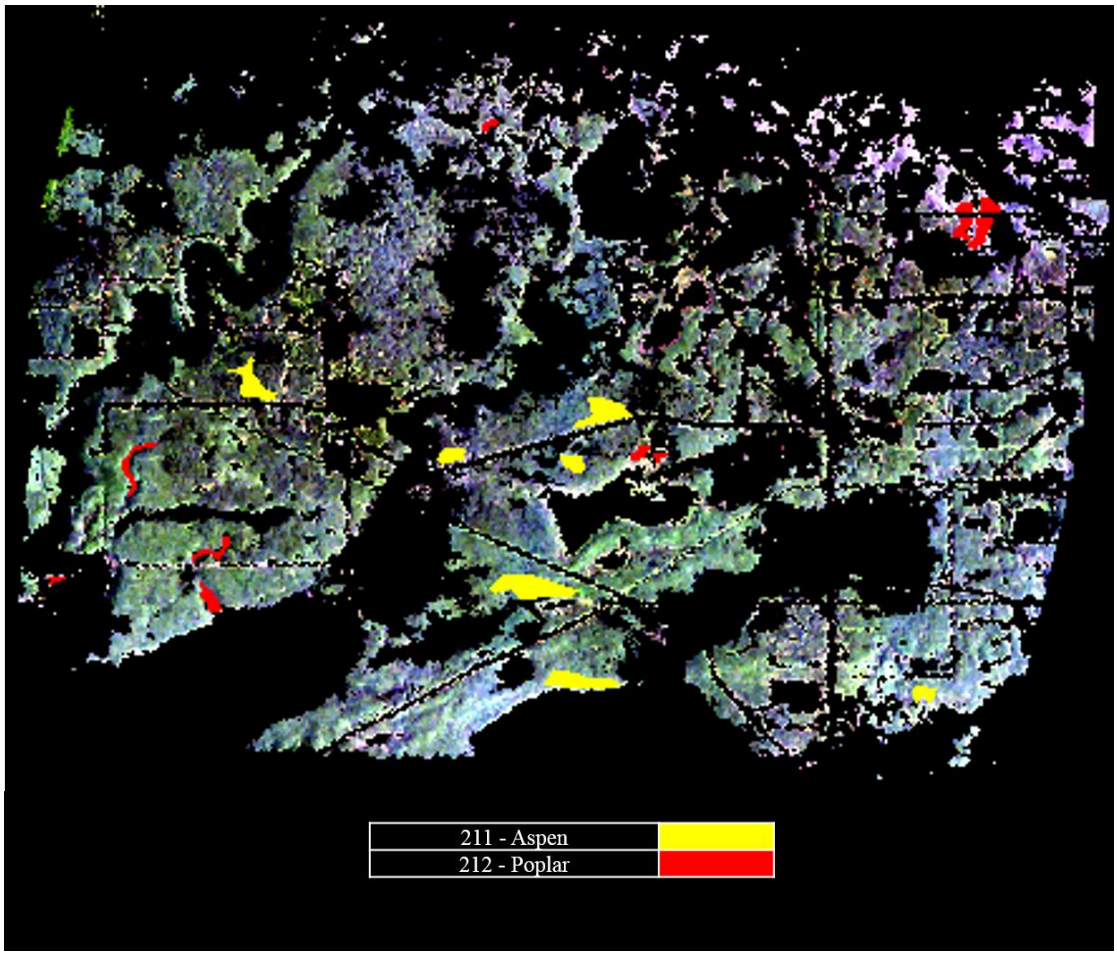

Figure 3.12: Distribution of Level 3 ROIs across the deciduous-masked CHRIS image. A linear 2-% image enhancement has been applied, resulting in the spectral variations within the forested areas becoming more pronounced (bands used: R:17 G:8 B:1).

Table 3.4: Deciduous species classes and number of SVM training validation pixels.

Class Name	**Description**	**Training Pixels**	**Validation Pixels**
211 – Aspen	*Populus tremuloides*	254	109
212 – Poplar	*Populus balsamifera*	28	13

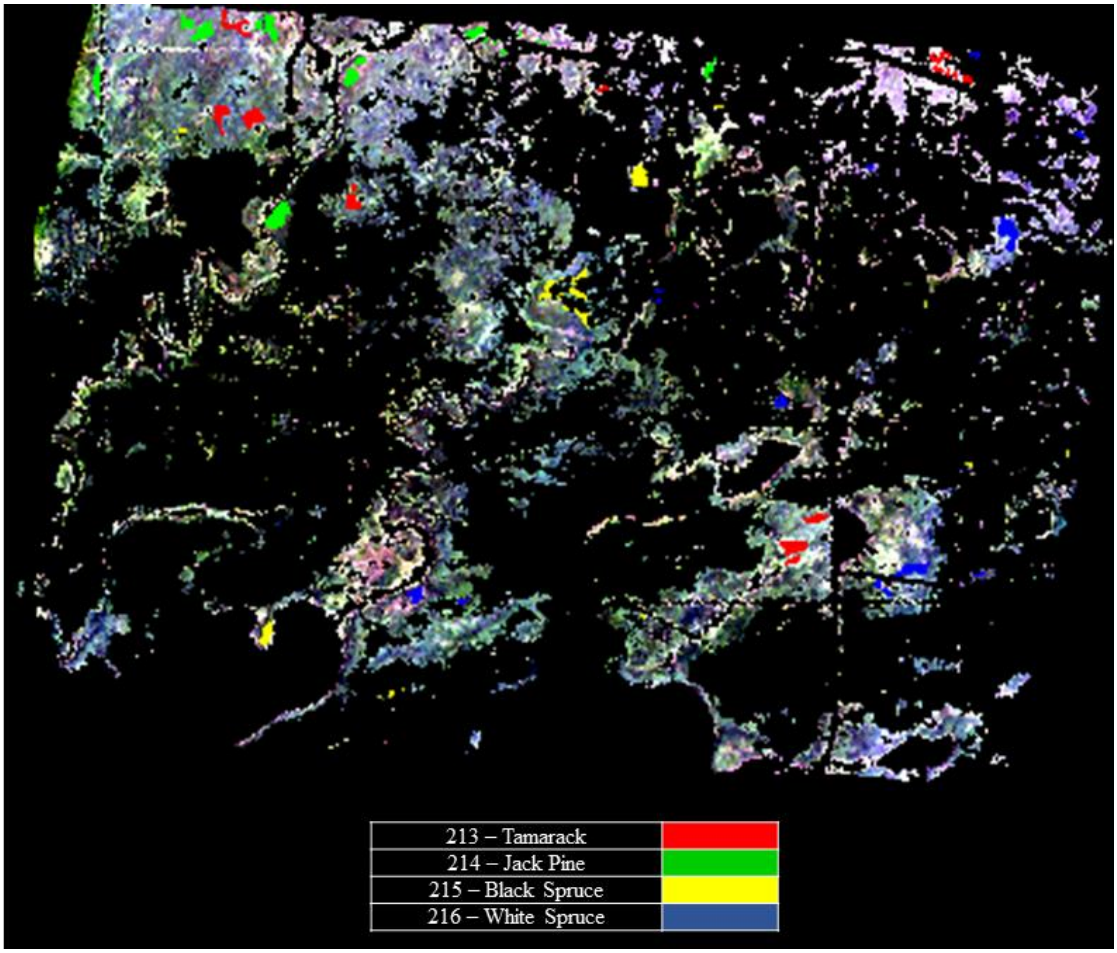

Figure 3.13: Distribution of Level 3 ROIs across the coniferous-masked CHRIS image. A linear 2-% image enhancement was applied, resulting in the spectral variations within the forested areas becoming more pronounced (bands used, R:17 G:8 B:1).

Table 3.5: Coniferous species classes and number of SVM training validation pixels.

Class Name	Description	Training Pixels	Validation Pixels
213 – Tamarack	*Larix laricina*	32	14
214 – Jack Pine	*Pinus banksiana*	29	13
215 – Black Spruce	*Picea mariana*	38	17
216 – White Spruce	*Picea glauca*	77	33

3.4.4 Accuracy Assessment

Classification accuracy was assessed using 30 % of the total number of ROI pixels in each class. The pixels used in the accuracy assessment were not included to train the

classifiers. The producer's and user's accuracies were assessed through the confusion matrix and give an indication that the probability of a land-cover type on the ground is correctly identified in a pixel and that a land-cover type in the classified imagery actually exists on the ground, respectively. Both of these accuracy measures give an indication with respect to the accuracy of individual classes, whereas the overall accuracy reports the accuracy of the image as a whole and can be written as follows:

$$Producer's\ Accuracy\ (\%) = \frac{X_{ij}}{\sum x_j} * 100 \tag{3.4}$$

$$User's\ Accuracy\ (\%) = \frac{X_{ij}}{\sum x_i} * 100\ , \tag{3.5}$$

and

$$Overall\ Accuracy\ (\%) = \frac{D}{N} * 100\ , \tag{3.6}$$

where X_{ij} is the number of correctly classified pixels in column *i* and row *j* of a class within an error matrix, $x_i$is the number of correctly classified pixels in column *i*, and X_j is the number of correctly classified pixels in row *j*, *D* is the total number of correct pixels within each class, and *N* is the total number of pixels in the confusion matrix.

The Kappa coefficient is used to determine if the overall accuracy is significantly different than using a random classification. Because the Kappa coefficient is the overall accuracy against chance, comparisons can be made against other classification methods to determine if the overall accuracy assessment from one confusion matrix is statistically different from another one (Foody, 1992).

Using the hierarchical classification approach, the species classified images from the coniferous, deciduous, and mixed classifications were merged together to create a

final species classified image. Validation was then performed on the final classified image.

A pairwise comparison of the classified images was accomplished using McNemar's test with a continuity correction (Foody, 2006 and 2004). This test is used when the same sets of training and validation samples are used in thematic map generation. Because the samples are not independent of one another, the use of the standard Z-test is not appropriate (Cohen, 1968). McNemar's test is a non parametric and can be described as follows:

$$Z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}}, \tag{3.7}$$

where f_{12} is the frequency of correctly classified pixels in both classifications and f_{21} is the frequency of incorrectly classified pixels in both classifications. These numbers are derived from the classification images themselves and not the error matrices.

A continuity correction is employed here due to the relatively small sample sizes (< 30) used for validation as follows (Rozenstein and Karnieli., 2011; Foody, 2004):

$$Z = \frac{|f_{12} - f_{21}| - 1}{\sqrt{f_{12} + f_{21}}}. \tag{3.8}$$

3.5 MESMA

The spectral profile of pixels present in remote sensing imagery is made up of a combination of different materials. In order to determine what materials and how much of a material is present within a given pixel, spectral unmixing methods like MESMA may be used.

MESMA was applied to spectrally unmix the CHRIS and CHRIS/LiDAR data based on the classification hierarchy established in the previous sections. In order to produce meaningful results from MESMA, the selection of spectra for the creation of a spectral library is an important step. Endmembers that are selected must be representative of the class to be unmixed.

3.5.1 Spectral Library

In order to select endmembers, MESMA requires a spectral library from the training data. For this study, the spectral library was created from the tree species ROIs delineated using the CHRIS image data. The Visualization and Image Processing for Environmental Research (VIPER) tools add-on in ENVI creates spectra for every pixel within an ROI to ensure that spectral diversity is captured rather than averaging the spectra over the entire ROI. From the spectral library, endmembers were selected from the total number of spectra of each class to represent the classes as described in the next section.

3.5.2 Endmember Selection

Different methods exist for endmember selection. The endmembers can either be selected through a semi-automated manner such as the Pixel Purity Index (PPI; Boardman *et al.*, 1995), N-FINDR (Winter, 1999), Principal Component Analysis (PCA; Deng *et al.*, 2008), Vertex component analysis (Nascimento and Bioucas Dias, 2005) or through a manual process. The latter relies on user guidance to ensure that a pixel spectrum is representative of the material to be unmixed. For this study, the manual approach was selected and facilitated by the ENVI plug-in VIPER tools. Within these tools, the

following different metrics of endmember selection can be employed: Count-based Endmember Selection (CoB), Endmember Average Root Mean Square Error (EAR), and the Minimum Average Spectral Angle (MASA). The former is a measure that allows a user to determine how well the spectrum from a selected endmember models spectra within the same endmember class, and how well it models spectra outside of the class (Roberts *et al.*, 2003). For this metric, two parameters are produced, an In-CoB and an Out-CoB. An endmember that is well suited to represent its spectral class will have a high In-CoB value and a low Out-CoB value. It models other spectra within the class well and does not model spectra outside of the class well.

EAR is a measure of the average Root Mean Squared Error (RMSE) when a spectra is compared to all other spectra within its class (Dennison and Roberts, 2003b). A low EAR value for a spectrum means that it is a good fit when compared with all other spectra in its class and is a good candidate for endmember selection. MASA compares the spectral angle of a pixel to all other pixels in the class and is similar to EAR in that they both compare the fit of a spectrum to others within its class. A low spectral angle represents a good endmember (Dennison *et al.*, 2004).

Franke *et al.* (2009) suggested selecting unique representative endmembers each time MESMA is run. However, in order to compare whether the addition of LiDAR data has an impact on the spectral unmixing of forested areas, endmembers selected for the CHRIS imagery alone were also used for the CHRIS/LiDAR imagery. The endmembers were selected based on a high In-CoB and low Out-CoB value as well as low EAR and MASA values.

3.5.3 Level 3 Modelling

A three-endmember model was used to separate tree species in the deciduous-masked dataset, while a four-endmember model was used to unmix tree species in the coniferous and mixed datasets. An additional endmember model was added to the coniferous and mixed classes due to the larger number of tree species present within those classes compared to the deciduous ones.

The deciduous class had two tree species to unmix, the coniferous class had four, and the mixed class had six. Creating an additional model allowed for another layer of distinction to be added to the unmixing procedure. Coniferous trees with similar growing environments (Black Spruce and Tamarack in wetter environments and White Spruce and Jack Pine in dryer environments) were grouped into their own model. In the three-endmember model, one model consisted of upper canopy vegetation (tree species), one of understory (shrubs, grasses) and one of photometric shade.

In the four-endmember model, one model consisted of understory (shrubs, grasses) and another model consisted of photometric shade. The remaining two models consisted of two tree species each, one composed of Tamarack and Black Spruce tree species, while the other was composed of Jack Pine and White Spruce. The four tree species were separated from each other and then grouped into two separate models, since these tree species are typically found growing in the same environmental regimes. Tamarack and Black Spruce are located in low lying, wetter areas, while Jack Pine and White spruce are typically found in elevated, drier areas. All pixels were unmixed using a different combination of endmembers generating numerous three and four-endmember models. Three endmembers were selected for each species and six endmembers selected

for the understory. When multiple pixels had the same In-CoB value, the pixel with the lowest EAR value was selected. An example of the endmembers selected is provided in Figure 3.14.

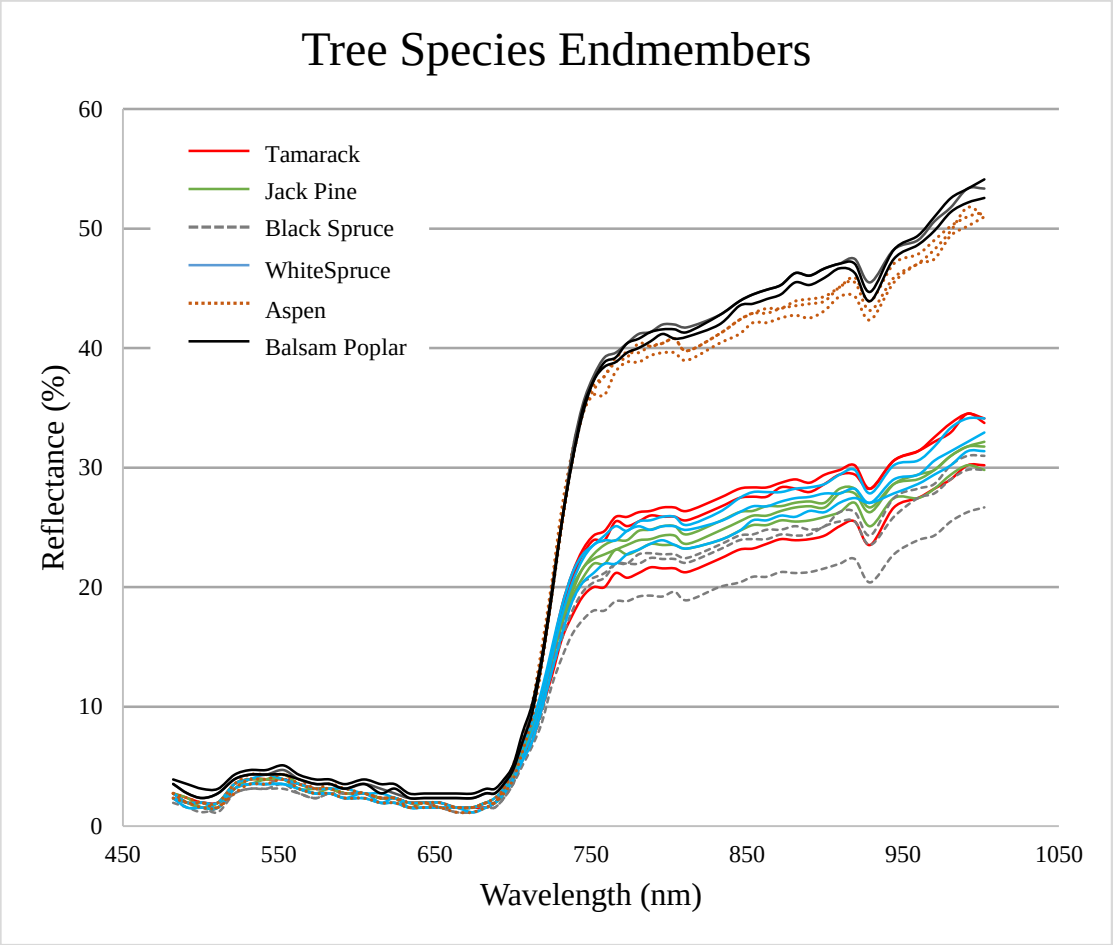

Figure 3.14: CHRIS endmember spectra for tree species showing the spectra used in the models.

MESMA was run using minimum and maximum endmember fractions of -0.10 and +1.10, respectively, a maximum shade fraction of 0.80, and a maximum allowable RMSE of 1.0. The values selected were empirically determined to yield optimum results.

A total of 36 models were created to unmix each pixel within the deciduous-masked image, while 72 models were created to unmix each pixel within the coniferous-masked image. The larger number of models used to unmix the coniferous image is because there are four tree species that need to be represented. A total of 126 models were created to unmix each pixel within the mixed masked image. Again, the increase in the number of models required to unmix an image is due to the number of endmembers used, i.e., six tree species.

An unmixing model was created for each unique combination of endmembers in each spectral library. Figure 3.15 demonstrates how 36 models can be created in a three-endmember unmixing procedure (one of the endmembers is shade and included in every model). Each of the tree-species endmembers on the left side of the figure is matched to the understory endmembers in a one-to-many relationship. The procedure is repeated for each tree-species endmember. Accordingly, Balsam Poplar (Pb) 1 is matched to Understorey (Un) 1-6, Pb 2 is matched to Un 1-6, and Pb 3 is matched to Un1 1-6. The procedure is the same for Aspen Poplar (Aw) 1-3.

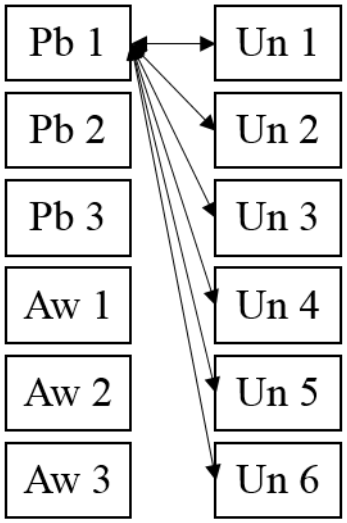

Figure 3.15: Example of one endmember (Pb 1) creating 6 models by matching with each of the understory (Un #) endmembers. By combining each endmember on the left side of the figure with each endmember on the right side of the figure, 36 unmixing models are created.

Shade normalization was performed on the fractional images produced by MESMA. This procedure effectively removes the shade component from abundance estimates for each pixel. A classification of an image is accomplished by assigning a class to a range of models produced by MESMA. For example, in the three-endmember models used to unmix deciduous tree pixels, 18 out of the 36 models identified Aspen as the dominant spectra. Thus, any pixel that was assigned a model number between 1 and 18 can be classified as Aspen.

3.5.4 Accuracy Assessment

Accuracy was assessed using validation pixels created from the original ROIs to derive a confusion matrix. Like the SVM, accuracy assessments, user's and producer's accuracies as well as the kappa coefficient were derived. Since the same training and validation data sets used to create the classified maps for the CHRIS and CHRIS/LiDAR

products, a pairwise classification was facilitated through McNemar's test (Section 3.4.4). Again the continuity correction was applied due to the low sample size. This test does not determine whether one classification is better than the other. However, it reveals whether the classifications are different from one another.

3.6 Summary

This chapter presented the methods used to map tree species in an area dominated by oil and gas operations and reclamation activities in the Cold Lake area, Alberta. Hyperspectral and LiDAR data were used separately and fused together to facilitate the tree-species composition mapping. A hierarchical approach was used to classify the tree species from broad land-cover classes to tree-type classes and finally tree species. A supervised classification method, SVM, and a spectral unmixing technique, MESMA, were used in this study. The results from the classification and unmixing were assessed using field data captured near the same time as image data acquisition and compared to areas in the imagery that were examined based on visual interpretation. The next chapter will present the results achieved followed by a discussion of the results.

Chapter 4

Results

4.1 Introduction

The aim of this thesis was to facilitate tree discrimination at the species level in and around abandoned and reclaimed oil and gas well-sites in the Cold Lake area, Alberta using remote sensing technologies. Two objectives of this study were to assess the use of hyperspectral data alone and combined with LiDAR data to determine tree species and to evaluate spectral unmixing methods in pursuit of tree-species discrimination. The following paragraphs present the results obtained in an attempt to address the objectives.

CHRIS data were processed to correct for atmospheric and sensor noise effects and then processed alone and in conjunction with airborne LiDAR data in order to investigate tree-species composition around Cold Lake, Alberta. The latter processing step involved the combination of the hyperspectral CHRIS data with the LiDAR data to create a fused CHRIS/LiDAR product. Two methods of cataloguing tree species were investigated: SVM classification and MESMA spectral unmixing.

A three-tiered hierarchical approach was used to select areas for classification (Figure 3.9). The first level involved a separation of dominant land-cover types (Forest, Wetlands, Grass, Water and, Non-Vegetated). From this first level of classification, a mask was created over any areas not identified as forest. The second level used the forest mask in order to separate coniferous, deciduous, and mixed tree types. SVM classification was used for the first two levels of the classification hierarchy. The separated coniferous, deciduous, and mixed tree datasets were then used for the final level of classification in which SVM and MESMA were used to map tree species.

Evaluation of the SVM classification and MESMA specified classes was accomplished through the use of confusion matrices, whereby pixels in the classified images were compared to validation pixels specified through field validation and visual assessment.

The results obtained using the methods above are presented in six sections. The first section describes the steps involved processing the CHRIS and LiDAR datasets from raw data to processed information. An analysis of data gathered during the field campaign, including a comparison of field measured parameters on and off well-sites is presented, followed by a comparison of the field gathered data to remotely interpreted data of the same well-sites. The next four sections show the results of the classification hierarchy, Levels 1, 2, and 3, and present a comparison of the SVM and MESMA classification methods used.

4.2 Pre-Processing

4.2.1 CHRIS

The CHRIS hyperspectral data were pre-processed in order to remove noise present in the raw data. This was accomplished through the use of ESA's BEAM toolbox (Brockmann-Consult, 2010) and involved the use of weighted averaging of neighbouring pixels to correct for pixel dropouts and utilizing all five look angles of the CHRIS data to correct for striping (Figure 4.1).

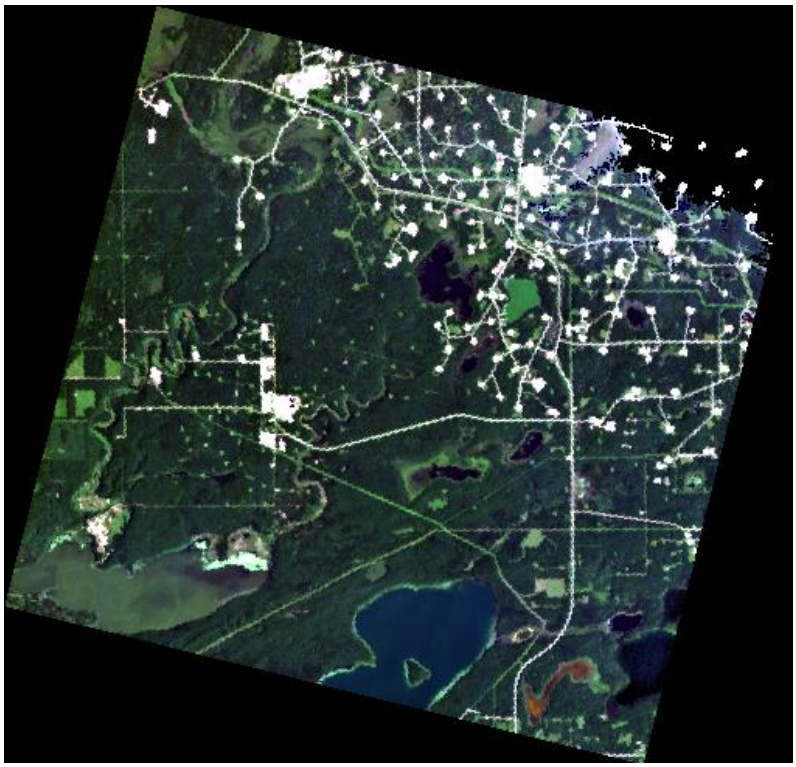

Figure 4.1: De-striped CHRIS imagery (bands used, R:17 G:8 B:1). See Figure 3.7 for the raw CHRIS imagery.

Atmospheric and geometric corrections were then applied to the noise corrected CHRIS data. The correction to surface reflectance was facilitated through an implementation of the MODTRAN 4 algorithm present in the BEAM toolbox. Geometric correction was accomplished using an image-to-image registration technique on geometrically rectified airborne hyperspectral data.

After the CHRIS image correction was accomplished, bands with values that were identified as being too low (reflectance values below 0) or affected by residual errors were removed from processing or smoothed using the Empirical Flat Field Optimal Reflectance Transformation (EFFORT) tool present in ENVI 5.0, respectively (Figure 4.2).

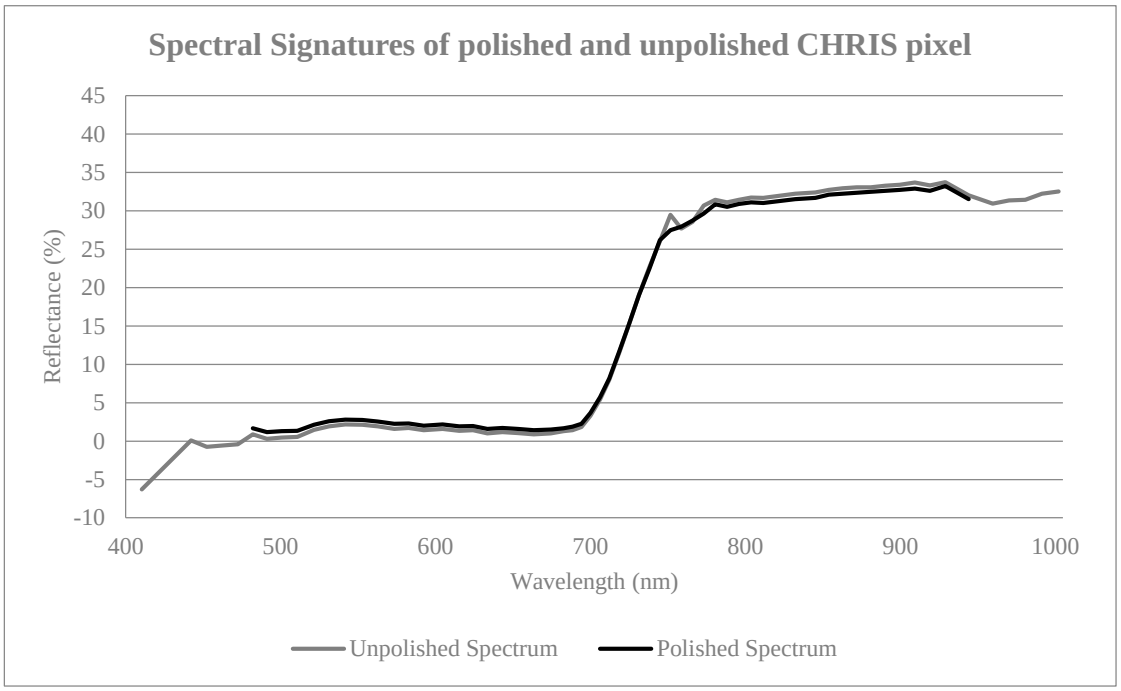

Figure 4.2: Forest pixel spectrum of atmospherically corrected CHRIS data before and after spectral polishing. Notice the peak in the spectrum of unpolished and the smoothness of the polished data at approximately 720 nm.

4.2.2 LiDAR

LiDAR data were delivered geometrically corrected from the data provider. The delivery format was in tiles, which were merged together. From the merged tiles, three LiDAR data products were created: a bare Earth surface model, a canopy surface dataset, and a canopy height model (Figure 4.3). All three LiDAR products were created using tools present in the FUSION/LDV software (McGaughey, 2007).

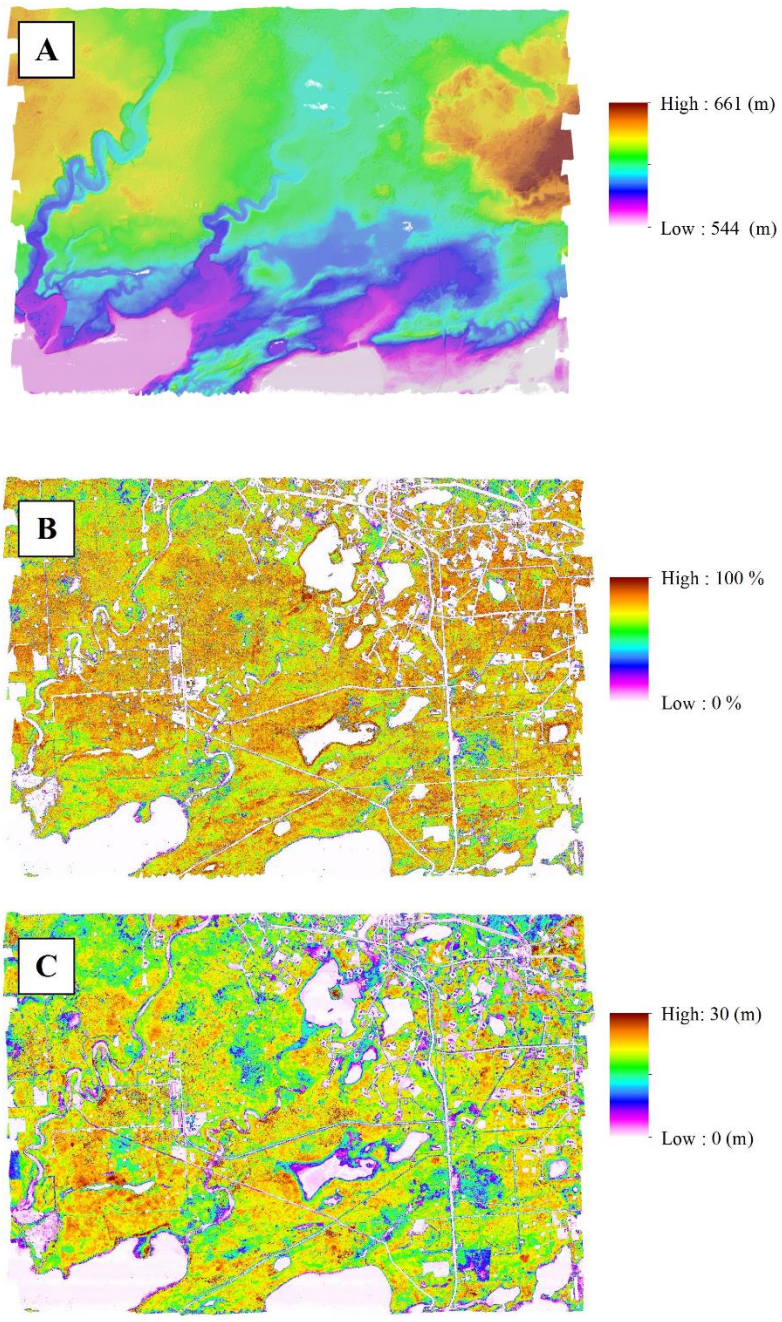

Figure 4.3: LiDAR products from data captured in August 2012 and cropped to the extent of the CHRIS imagery: *A*. Bare Earth DEM (masl), *B*. Canopy fractional cover, and *C*. Canopy Height.

CHRIS data were merged with the canopy cover and canopy height datasets to create a unified CHRIS\LiDAR dataset. The LiDAR products were appended to the

processed CHRIS data. The resulting spectral profile showed the CHRIS reflectance values and the LiDAR product data values (canopy cover percentage and height) for each pixel (Figure 4.4). Analyses were conducted on this merged data product as well as the processed CHRIS data product.

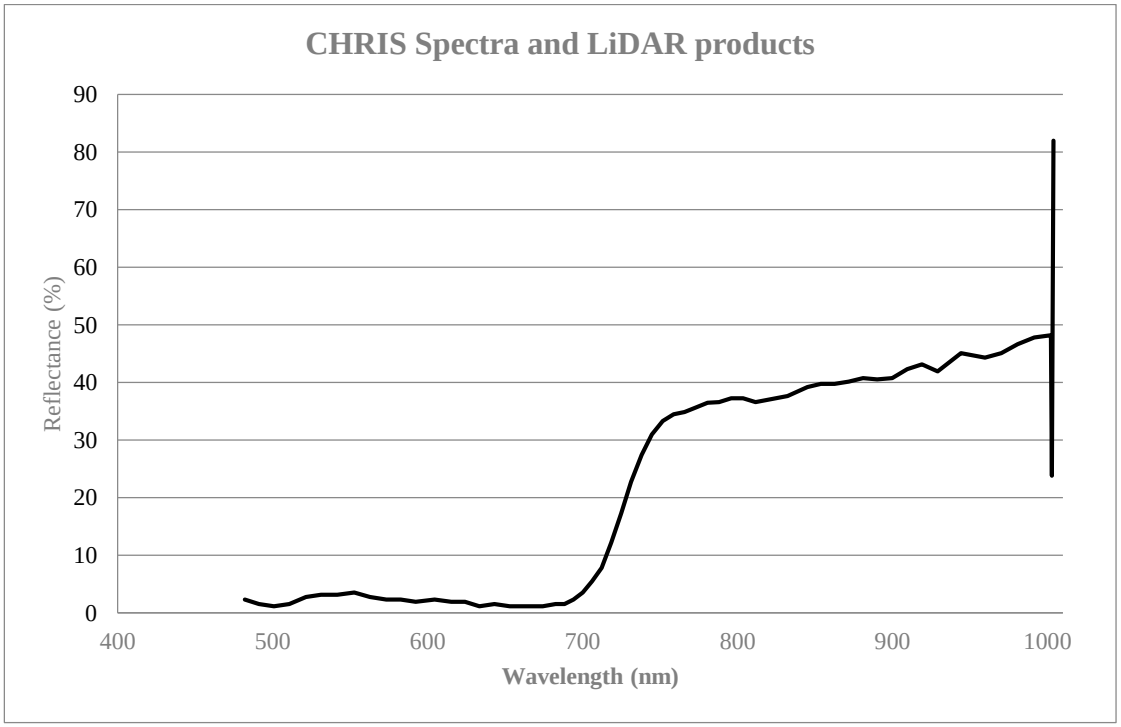

Figure 4.4: CHRIS spectra and LiDAR canopy height and canopy cover products appended to the end of the spectrum and stretched between 0 % and 100 %.

4.3 Data Collection

Data captured in the field from reclaimed well-sites and reference sites were compared in order to determine the progress of reclamation at the well-sites. All 48 plots (24 well-sites and 24 reference sites) were examined. However, only 32 plots (16 well-sites and 16 reference sites) were within the footprints of the CHRIS and LiDAR datasets. Well abandonment dates and reclamation certification dates were provided by AESRD while dominant species, average DBH, average canopy height, and average canopy

fractional cover were determined for each site. The full table of information for each site can be found in Appendix 1.

A large number of well-sites lack trees as evident by the data points with a value of 0 for DBH, canopy height, and canopy fractional cover. The average measurements as well as standard errors are presented in Figures 4.5 - 4.7.

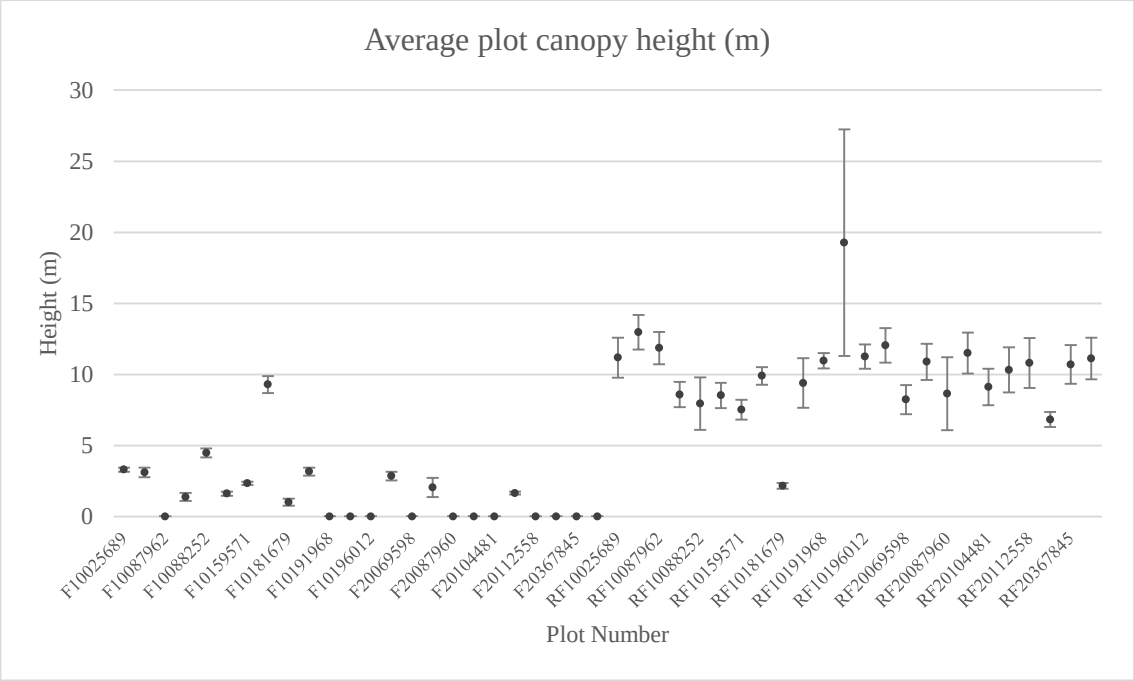

Figure 4.5: Average canopy height in m with standard error bars shown. Site RF10195902 had one very large tree within the site boundary which creates a large standard error. The average canopy height at the reclaimed sites is lower than the one at corresponding reference sites. Information in the figure is from ground measurements.

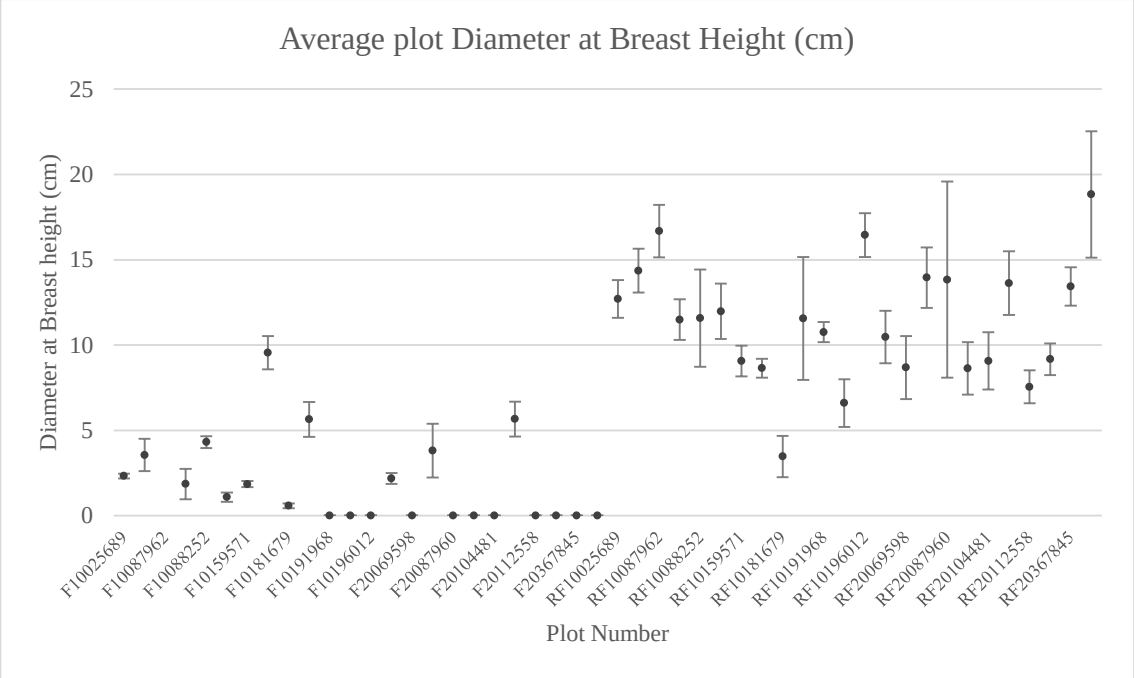

Figure 4.6: Average DBH in cm with standard error bars showing the deviation of DBH values within the study area. The average DBH at the reclaimed sites is lower than the average DBH at corresponding reference sites.

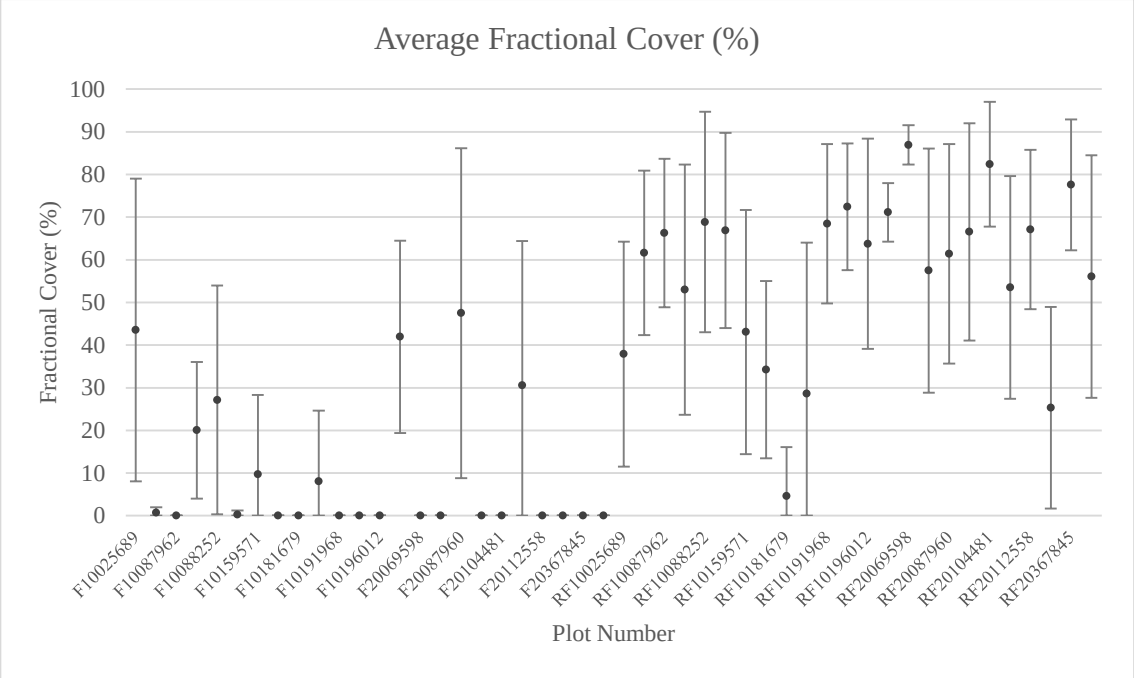

Figure 47: Average fractional cover with standard deviation bars showing the deviation of fractional cover values from the hemispherical photographs. The large deviation is due to clumping of trees within the study site and the tree species at that site.

The average canopy height at the reclaimed sites are lower than the average canopy height at the corresponding reference sites. This is to be expected since the reclaimed plots that were measured were seeded to grass initially. The average DBH shows lower values for reclaimed sites compared to their corresponding reference sites. Again, this is to be expected as the trees in the reclaimed sites are younger than the older, larger trees in the reference sites. A large standard deviation at all sites (reclaimed and reference) is apparent in the calculated average fractional cover. This is due to the types of trees present at a site as well as their distribution within a site. Coniferous trees will have a narrow canopy when compared to deciduous trees and the distribution of both tree species within a site is not uniform. Hemispherical photographs taken in a clump of trees will have a higher fractional cover than a hemispherical photo taken in a clearing.

4.4 Field Parameters as Estimated by Remote Sensing

Similar parameters to those captured in the field were retrieved using the processed remote sensing data for each reclaimed well-site and reference site. Canopy height, fractional canopy cover, and species were extracted from the combined LiDAR and CHRIS data. For those sites outside of the CHRIS extent, only canopy height and canopy fractional cover were reported (Figure 4.8 and Figure 4.9).

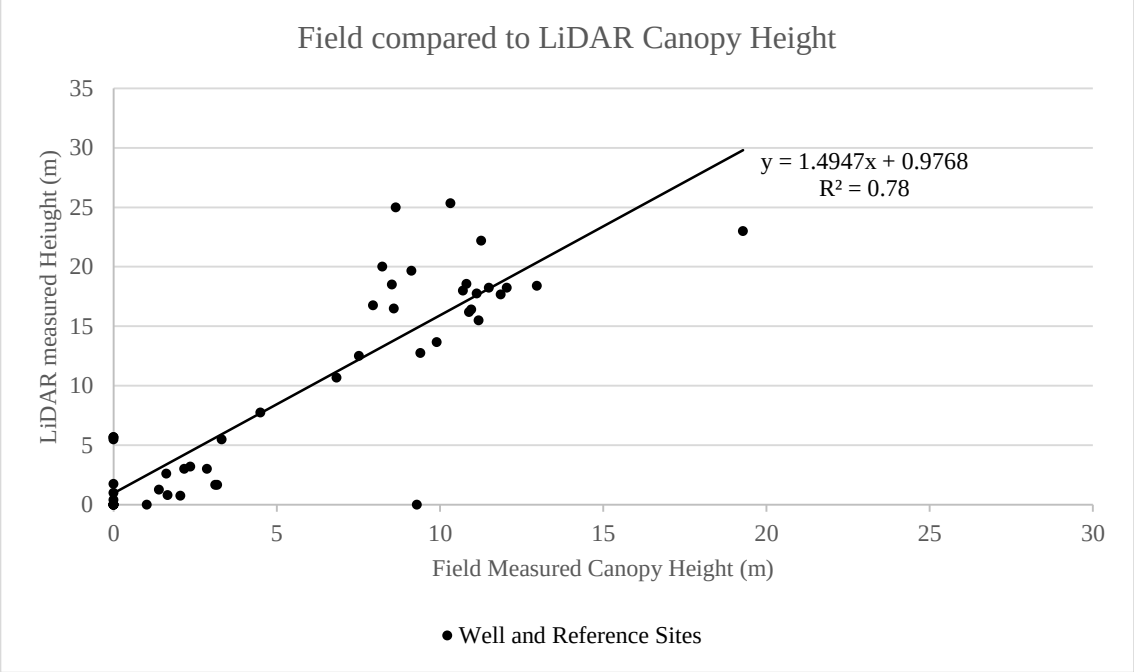

Figure 4.8: Average field canopy height compared to average LiDAR derived canopy height at a sample site. Each point represents a field measured location. A moderately strong relationship exists between field measured canopy height and LiDAR measured canopy height.

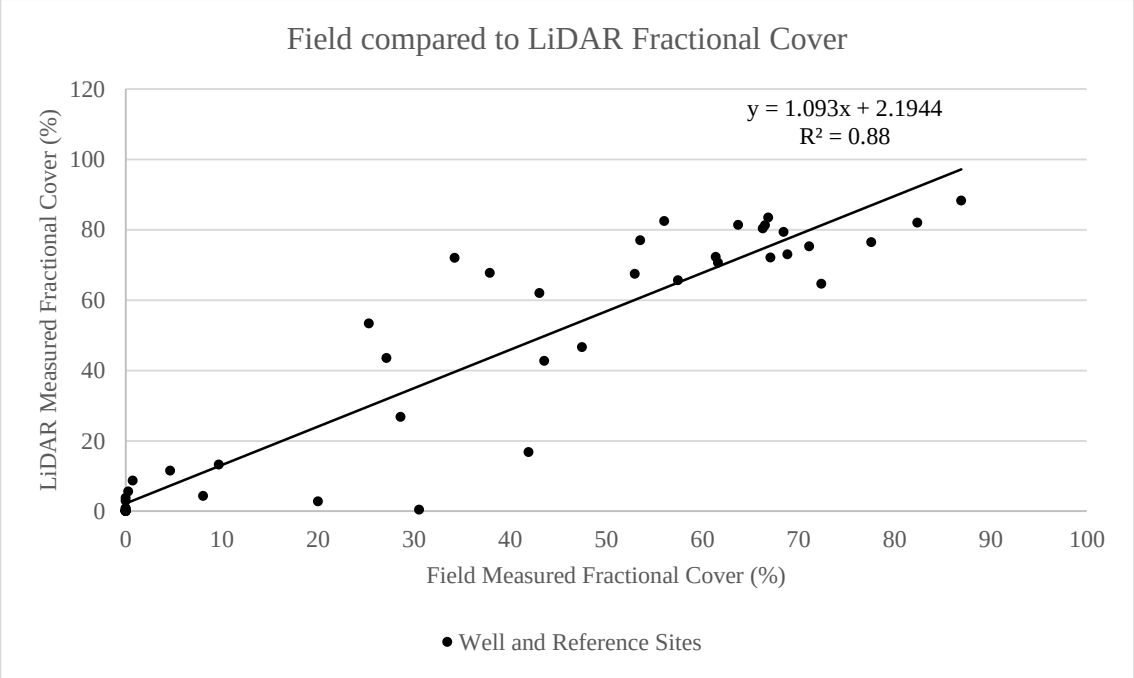

Figure 4.9: Average field fractional cover compared to average LiDAR derived fractional cover at a sample site. Each point represents a field measured location. A relatively strong relationship exists between field and LiDAR measured fractional cover.

Comparing the field measured parameters to LiDAR measured parameters, it can be seen that canopy height showed a moderately strong relationship between the two data capture methods with an R^2 value of 0.78, while the fractional indicates a stronger relationship with R^2 value of 0.88. The well-sites that were lacking trees when measured in the field were found to have a very low if not 0-% fractional cover in the derived LiDAR product; 0 % means that the site was completely devoid of any tree canopies and the hemispherical images only captured the sky.

A comparison of the species measured at each site using the different methods is presented in Table 4.1. Only two of the sites identified in the field as having no trees present were correctly identified by all remote sensing methods.

Table 4.1: Comparison of species by different measurement methods.

Plots within CHRIS image extent	Majority Field Measured Tree Species[12]	Species From CHRIS SVM	Species From CHRIS LiDAR SVM	Species From CHRIS MESMA	Species From CHRIS LiDAR MESMA
F10027735	Pb	Aw	Aw	Unknown	Pj
F10087962	No Trees Present	Aw	Aw	Aw	Aw
F10088251	Aw	Pb	Pb	Aw	Pb
F10088252	Pb	Aw	Aw	Aw	Aw
F10135076	Pb	Aw	Aw	Pj	Sw
F10191968	No Trees Present	Aw	Aw	Aw	Pj
F20051406	Pb	Aw	Aw	Pj	Pb
F20069598	No Trees Present	Aw	Aw	Aw	Unknown
F20087929	Pb	Aw	Aw	Lt	Pj
F20087960	No Trees Present	Pb	Sb	Aw	Sb
F20088253	No Trees Present	Unknown	Unknown	Unknown	Unknown
F20104481	No Trees Present	Aw	Aw	Pb	Aw
F20105627	Pb	Pb	Pb	Unknown	Pb
F20112558	No Trees Present	Unknown	Unknown	Unknown	Unknown
F20367845	No Trees Present	Aw	Aw	Unknown	Aw
F20368906	No Trees Present	Aw	Aw	Unknown	Aw
RF10027735	Aw	Pb	Pb	Aw	Aw
RF10087962	Aw	Aw	Aw	Aw	Aw
RF10088251	Aw	Aw	Aw	Aw	Aw
RF10088252	Aw	Pb	Pb	Aw	Lt
RF10135076	Aw	Pj	Lt	Aw	Pb
RF10191968	Lt	Sw	Sb	Pj	Sb
RF20051406	Pb/Aw	Aw	Aw	Pb	Sw
RF20069598	Aw	Aw	Aw	Unknown	Aw
RF20087929	Pb	Sw	Sw	Sw	Sw
RF20087960	Pb	Aw	Pb	Aw	Pj
RF20088253	Aw	Aw	Pb	Aw	Pb
RF20104481	Aw	Aw	Aw	Pb	Aw
RF20105627	Pb	Aw	Aw	Unknown	Sw
RF20112558	Aw	Aw	Aw	Pb	Pb
RF20367845	Aw	Aw	Aw	Aw	Aw
RF20368906	Aw	Aw	Pb	Aw	Aw

Aspen Poplar (Aw), Balsam Poplar (Pb), White Spruce (Sw), Black Spruce (Sb), Jack Pine (Pj), Tamarack (Lt).

[1] Dominant tree species are identified using AVI methods (i.e., 80 % of total species).

[2] Most of the coniferous sites captured in the field campaign lie in the north end of the study site and outside of the imagery bounds.

4.5 Level 1 Results

The first level of image classification involved separating broad land-cover classes adopted from the EOSD into more general classes to suit the needs of this study. A supervised SVM classification was performed on the CHRIS data alone. A minimum of 250 pixels were selected for each land-cover class with 70 % of the pixels in a class used to train the classifier and the other 30 % used to validate the classification. Level 1 classification land-cover maps are presented in Figure 4.10 and the confusion matrix is shown in Table 4.2.

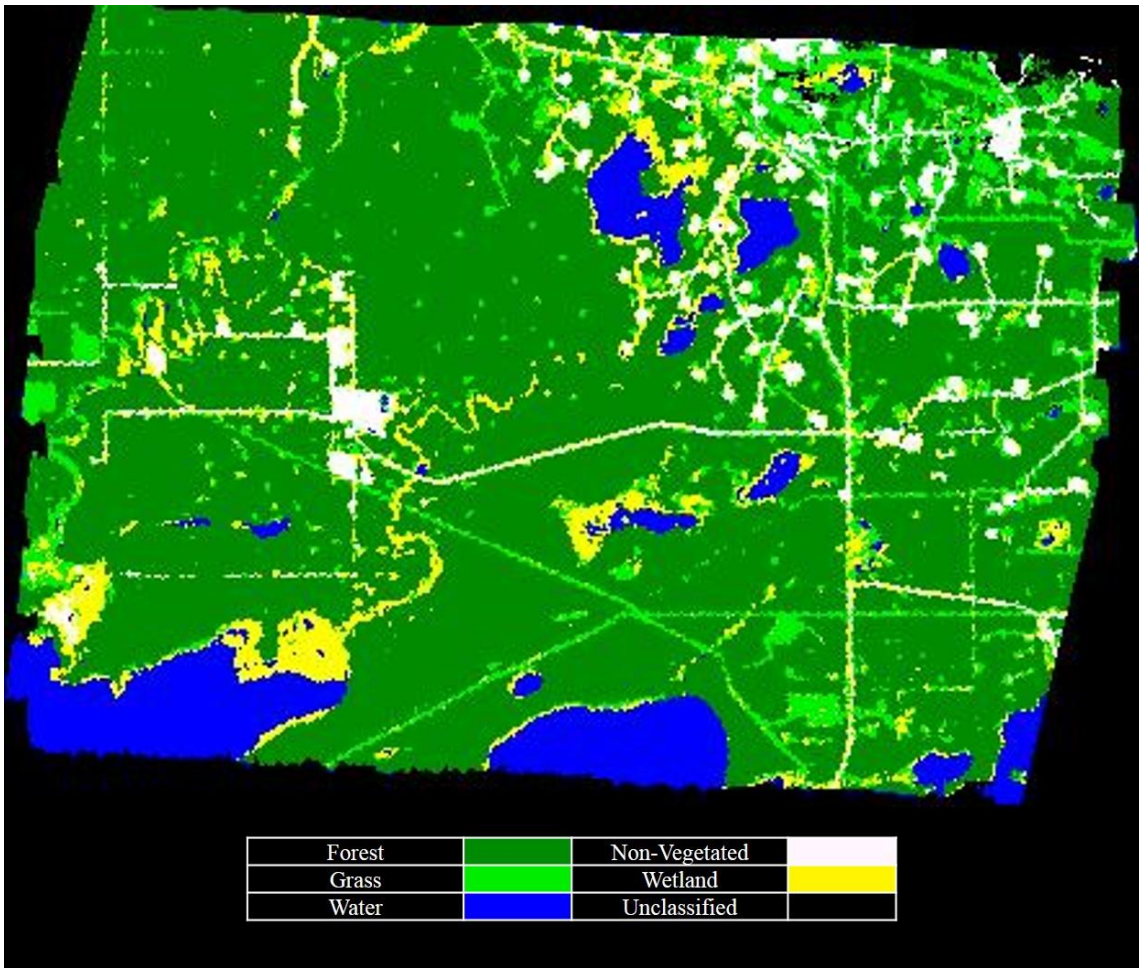

Figure 4.10: Land-cover map of Level 1 classification created from CHRIS data using the SVM classifier.

Table 4.2: CHRIS Level 1 SVM classification accuracies.

Map Prediction (pixels)	Reference (pixels)						User's Accuracy (%)
	Forest	Grass	Water	Non-Vegetated	Wetland	Total	
Unclassified	0	0	0	0	0	0	
Forest	479	5	0	0	0	484	98.9
Grass	2	235	0	4	5	246	95.5
Water	0	0	368	0	0	368	100
Non-Vegetated	0	0	0	201	0	201	100
Wetland	1	17	0	0	73	91	80.0
Total	482	257	368	205	78	1390	
Producer's Accuracy (%)	99.4	91.4	100	98.1	93.6		Overall Accuracy 97.5
Kappa							0.97

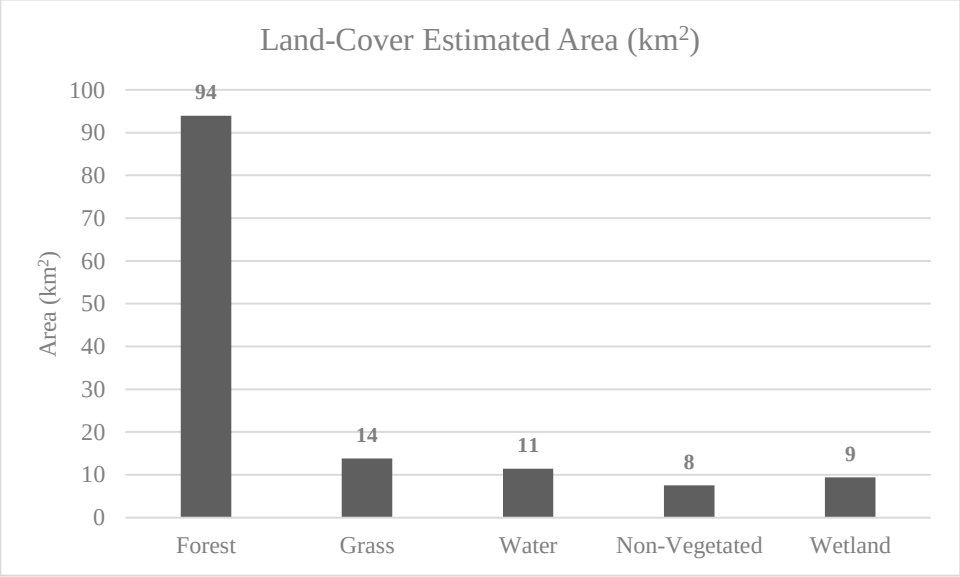

Figure 4.11: Estimated areas of land cover from the results of Level 1 classification. Area values are presented above data bars.

The overall accuracy for level 1 classification is 97.5 % with a Kappa value of 0.97. The forest class with 94 km^2 is the largest land-cover type by area, followed by grass with 14 km^2. Reported class areal coverages are presented in Figure 4.11. The producer's accuracies, the probability that a land-cover type is properly identified in an

image, are 99.4 % and 100 % for these classes, respectively, while the user's accuracy, the probability that a land-cover type identified by a classifier actually exists on the ground, are 98.9% and 95.5% for these classes, respectively.

4.6 Level 2 Results

Supervised SVM classifications were performed on the CHRIS data alone and on the CHRIS/LiDAR dataset after applying the forest mask produced in level 1 classification. A minimum of 700 pixels were used in the classification process with 70 % and 30 % of the pixels being used respectively to train and validate the classifier. Level 2 land-cover maps and classification accuracies for CHRIS data alone are presented in Figure 4.12 and Table 4.3, while those accuracies for the combined CHRIS/LiDAR data are presented in Figure 4.13 and Table 4.4.

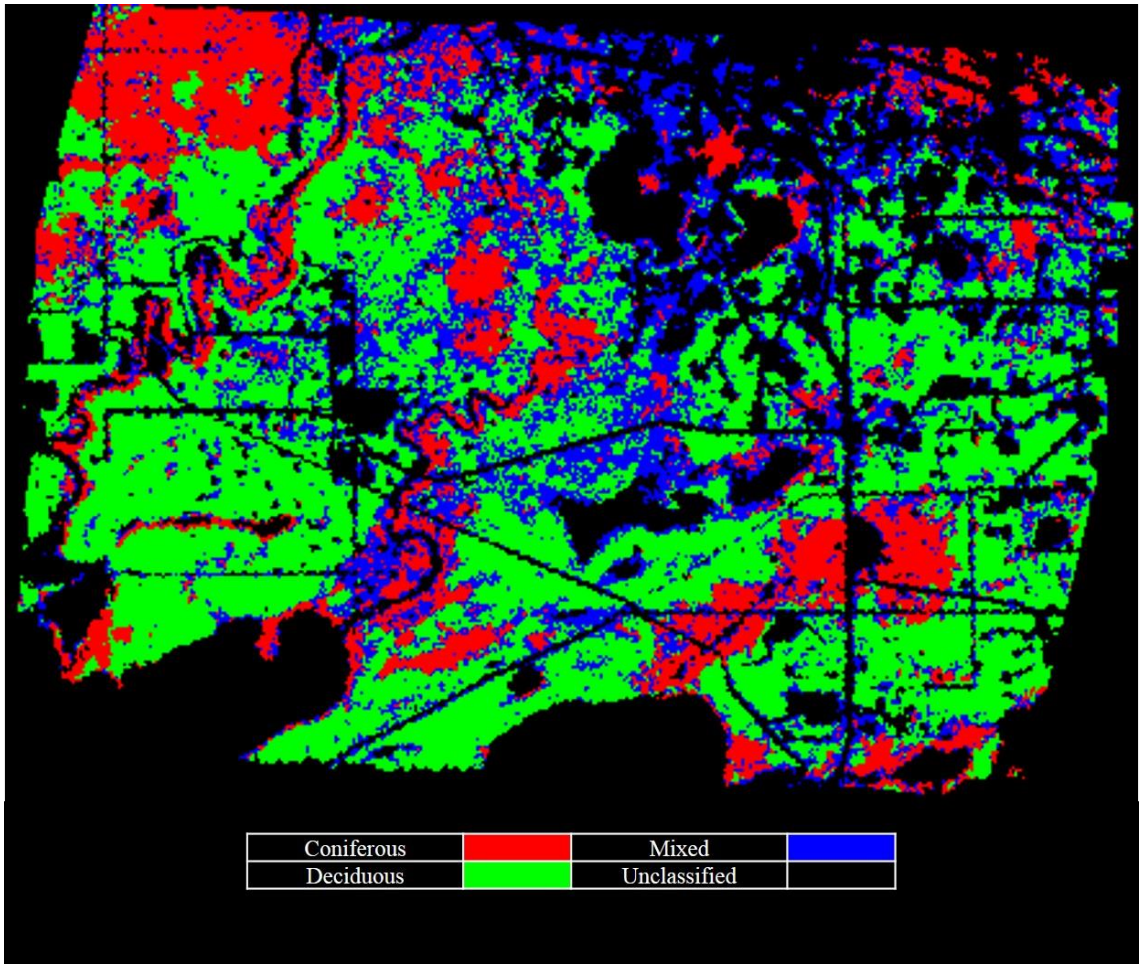

Figure 4.12: Forest-cover map of Level 2 classification created from CHRIS data using the SVM classifier.

Table 4.3: CHRIS Level 2 SVM classification accuracies.

Map Prediction (Pixels)	Reference (Pixels)				User's Accuracy (%)
	Coniferous	Deciduous	Mixed	Total	
Unclassified	0	0	3	3	
Coniferous	127	10	35	172	73.4
Deciduous	11	102	29	143	71.8
Mixed	15	7	36	58	62.0
Total	153	119	103	375	
Producer's Accuracy (%)	83.0	85.7	34.9		Overall Accuracy 70.7
Kappa					0.55

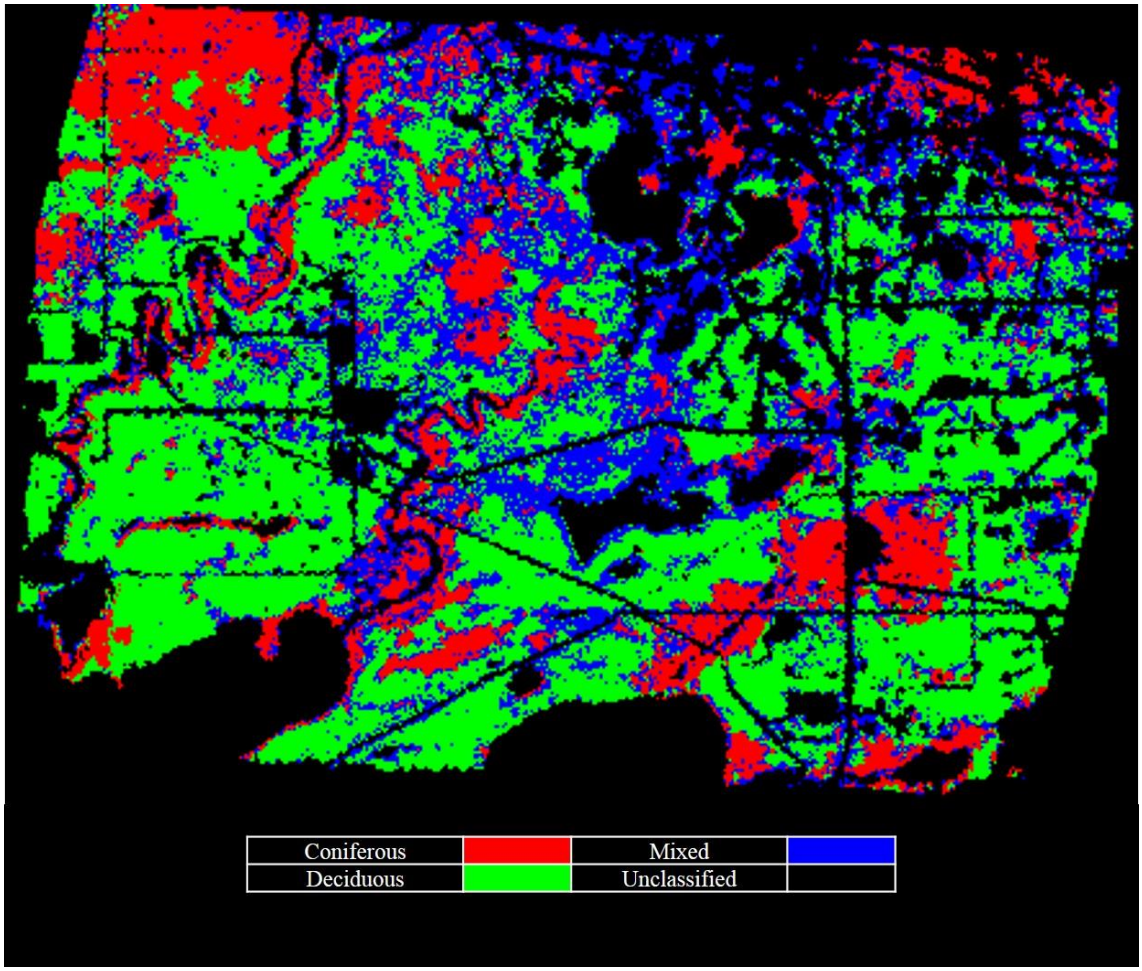

Figure 4.13: Forest-cover map of Level 2 classification created from CHRIS/LiDAR data using the SVM classifier.

Table 4.4: CHRIS/LiDAR Level 2 SVM classification accuracies.

Map Prediction (Pixels)	Reference (Pixels)				User's Accuracy (%)
	Coniferous	Deciduous	Mixed	Total	
Unclassified	0	0	3	3	
Coniferous	127	9	36	172	73.4
Deciduous	7	105	17	129	81.4
Mixed	19	5	47	71	66.2
Total	153	119	103	375	
Producer's Accuracy (%)	83.0	88.0	45.6		Overall Accuracy 74.4
Kappa					0.61

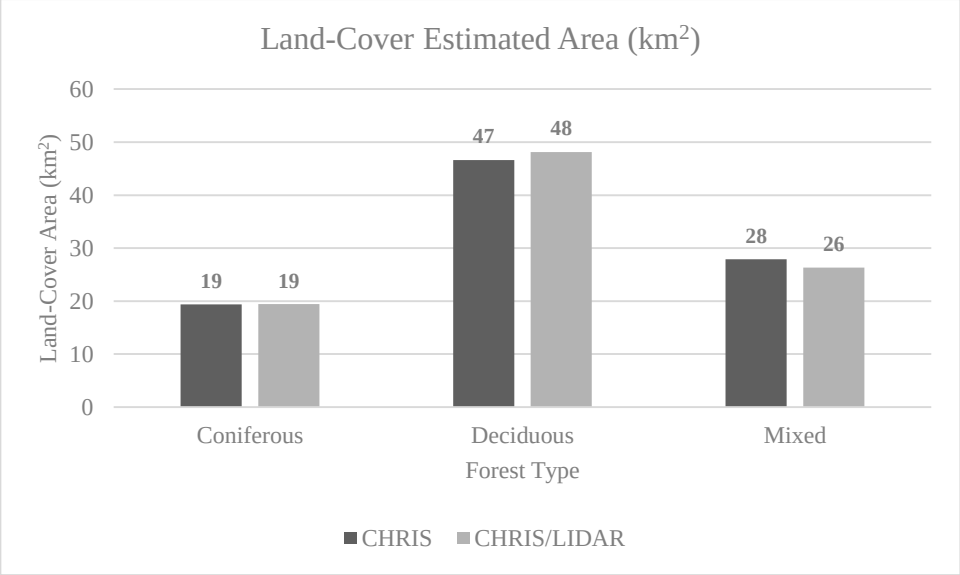

Figure 4.14: Estimated areas of land cover from the results of Level 2 classification. Area values are presented above data bars.

The CHRIS/LiDAR dataset had a higher overall accuracy of 74.7 % compared to the CHRIS dataset alone at 70.7 %. Results from both classifications show that deciduous trees were the largest forest-cover type by area, followed by mixed trees (Figure 4.14). The producer's accuracies for the CHRIS dataset for these classes were 85.7 % and 83.0 %, respectively. These values are the same or lower than the accuracies achieved for the same classes in the CHRIS/LiDAR dataset with 88.0 % and 83.0 %, respectively. Producer's accuracies of 34.9 % and 45.6 % for the CHRIS and CHRIS/LiDAR datasets, respectively, were achieved for the mixed class. The user's accuracies for the CHRIS dataset for these classes were 71.8 % and 73.4 %, respectively. These values are the same or lower than the accuracies achieved for the same classes in the CHRIS/LiDAR dataset with 81.4 % and 73.4 %, respectively. User's accuracies of 62.0% and 66.2% for the CHRIS and CHRIS/LiDAR datasets, respectively, were achieved for the mixed class.

A pairwise comparison of the CHRIS and CHRIS/LiDAR classification maps using McNemar's Test Z-statistics yielded a Z value of 2.77 (Table 4.5).

Table 4.5: Comparison of the classifications derived from the CHRIS and CHRIS/LiDAR datasets in pixels.

		CHRIS		
		Correct	Incorrect	Total
CHRIS/LiDAR	Correct	261	18	279
	Incorrect	4	92	96
	Total	265	110	

Comparing the CHRIS and CHRIS/LiDAR classification results, it was determined based on the Z-statistic ($p < 0.05$) that the difference between the two datasets is statistically significant.

4.7 Level 3 Results

The third-level classification attempted to separate tree species using SVM and MESMA. Masks were created from the CHRIS/LiDAR level 2 classification for each tree type, and the classification and unmixing procedures were performed on each masked dataset. Supervised SVM classifications were applied to each tree type for the CHRIS and CHRIS/LiDAR datasets. Unmixing models were created using MESMA on the masked CHRIS and CHRIS/LiDAR datasets as well for the SVM classifications. Again 70 % of the pixels were used to train and 30 % were used to validate the classifier.

Level 3 SVM land-cover maps and classification accuracies for the CHRIS data alone are presented in Figure 4.15 and Table 4.6, respectively, while MESMA land-cover maps and classification accuracies for CHRIS data alone are presented in Figure 4.16 and

Table 4.7. The corresponding results for the CHRIS/LiDAR data are presented in Figure 4.17 and Table 4.8 and Figure 4.18 and Table 4.9, respectively.

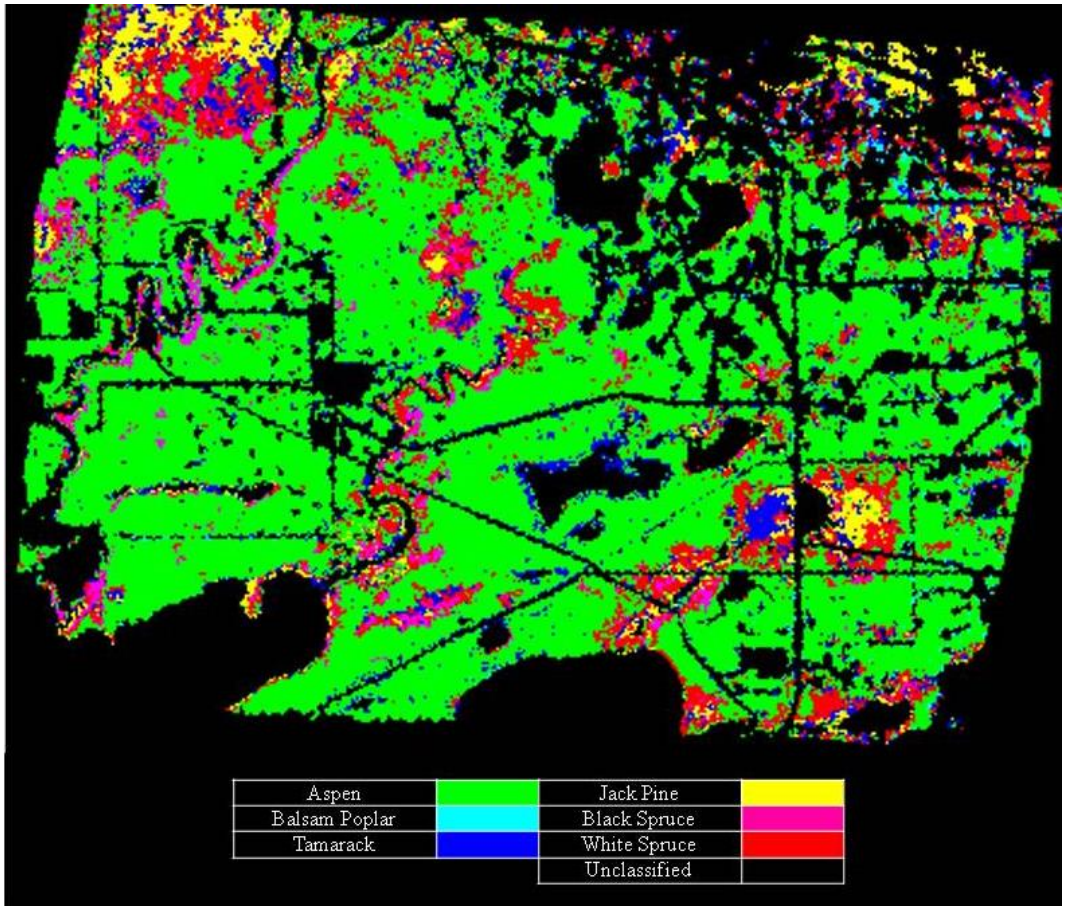

Figure 4.15: Species-cover map of Level 3 classification created from CHRIS data using the SVM classifier.

Table 4.6: CHRIS Level 3 SVM classification accuracies.

Map Prediction (Pixels)	Reference (Pixels)							User's Accuracy (%)
	Aspen	Poplar	Tamarack	Jack Pine	Black Spruce	White Spruce	Total	
Unclassified	1	0	1	0	0	0	2	
Aspen	102	4	0	0	0	6	112	91.1
Balsam Poplar	0	6	0	0	0	0	6	100.0
Tamarack	4	0	6	2	0	2	14	42.9
Jack Pine	0	0	1	4	0	0	5	80.0
Black Spruce	0	0	0	0	6	1	7	85.7
White Spruce	2	3	6	7	11	24	53	45.3
Total	109	13	14	13	17	33	199	
Producer's Accuracy (%)	93.6	46.2	42.9	30.8	35.3	72.7		Overall Accuracy 74.4
Kappa								0.60

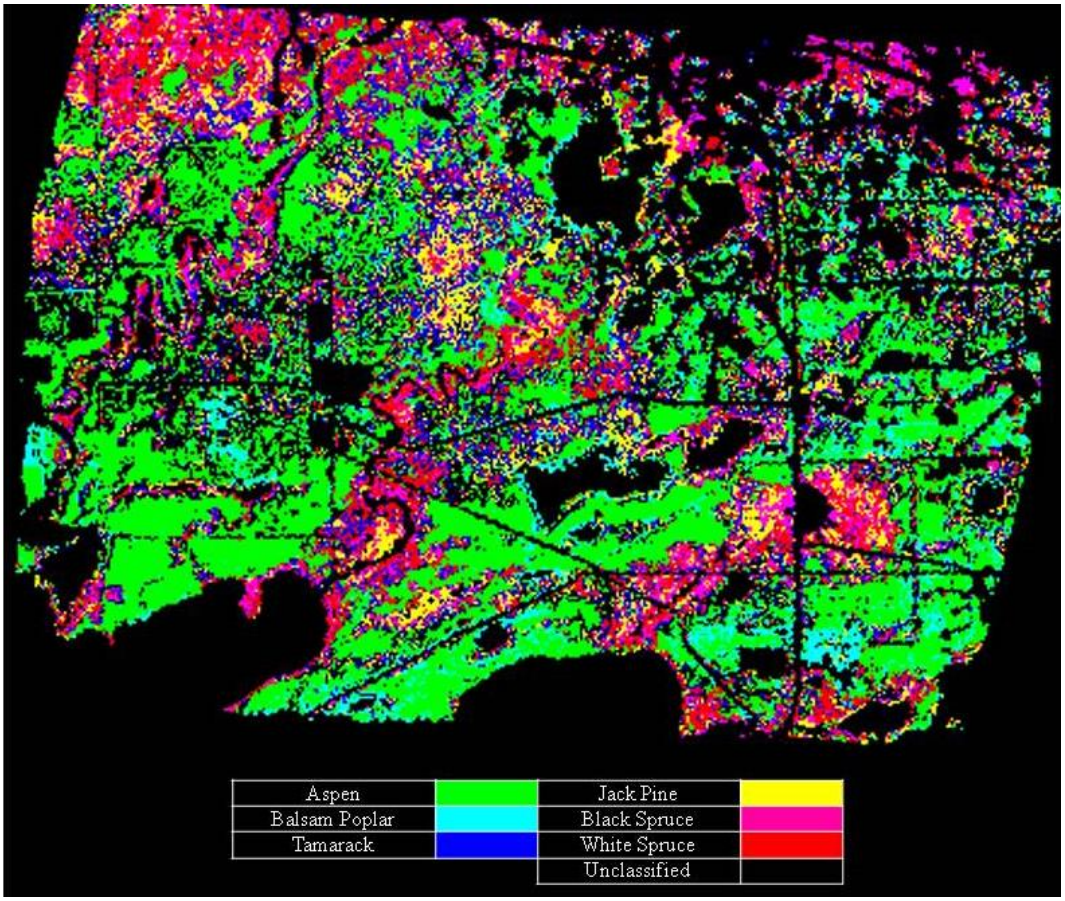

Figure 4.16: Species-cover map of Level 3 classification created from CHRIS data using MESMA unmixing.

Table 4.7: CHRIS Level 3 MESMA classification accuracies.

Map Prediction (Pixels)	Reference (Pixels) Aspen	Poplar	Tamarack	Jack Pine	Black Spruce	White Spruce	Total	User's Accuracy (%)
Unclassified	18	1	2	5	2	7	35	
Aspen	61	3	0	0	0	3	67	91.0
Balsam Poplar	26	6	0	0	0	1	33	18.2
Tamarack	1	2	6	5	1	7	22	27.3
Jack Pine	0	0	0	2	0	3	5	40.0
Black Spruce	3	0	6	1	4	5	19	21.1
White Spruce	0	1	0	0	10	7	18	38.9
Total	109	13	14	13	17	33	199	
Producer's Accuracy (%)	56.0	46.2	42.9	15.4	23.5	21.2		Overall Accuracy 43.2
Kappa								0.27

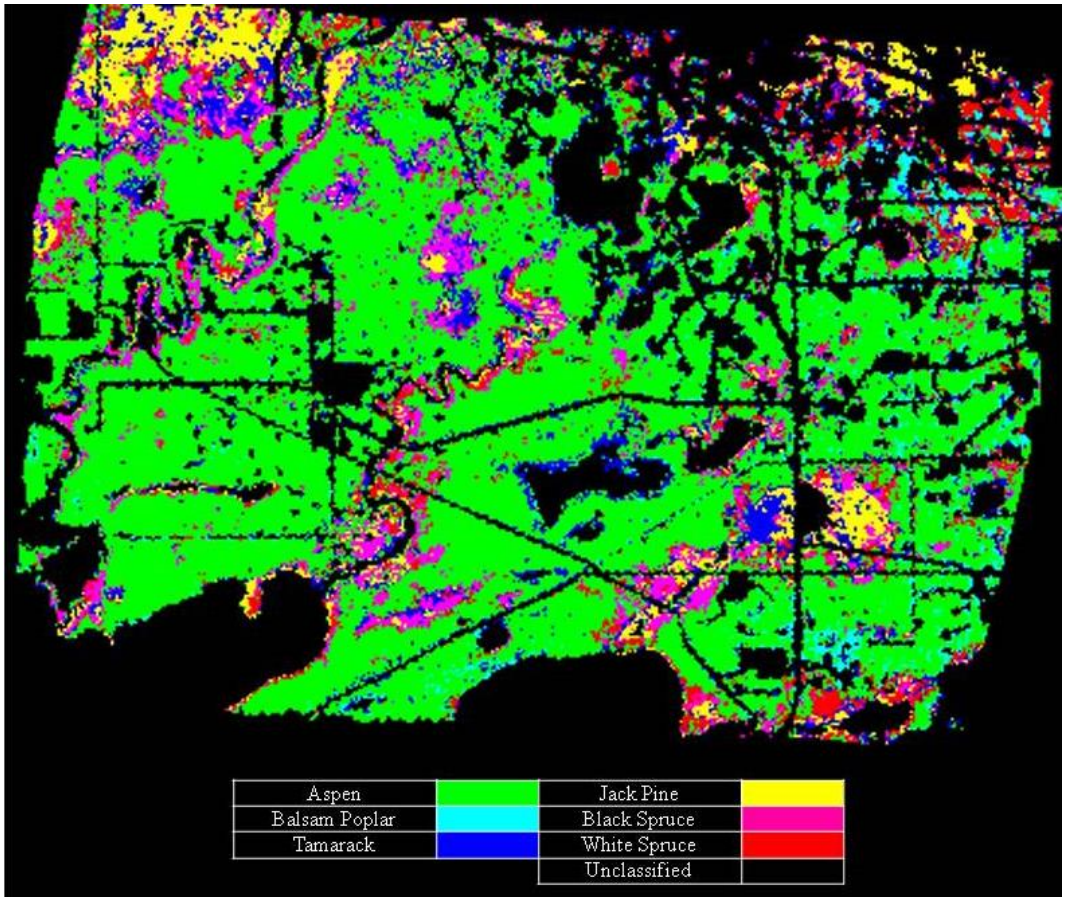

Figure 4.17: Species-cover map of Level 3 classification created from CHRIS/LiDAR data using the SVM classifier.

Table 4.8: CHRIS/LiDAR Level 3 SVM classification accuracies.

Map Prediction (Pixels)	Reference (Pixels)							User Accuracy (%)
	Aspen	Poplar	Tamarack	Jack Pine	Black Spruce	White Spruce	Total	
Unclassified	1	0	1	0	0	0	2	
Aspen	102	4	0	0	0	5	111	91.9
Balsam Poplar	0	6	0	0	0	0	6	100.0
Tamarack	6	0	9	2	0	1	18	50.0
Jack Pine	0	0	1	10	1	0	12	83.3
Black Spruce	0	2	3	1	14	2	22	63.6
White Spruce	0	1	0	0	2	25	28	89.3
Total	109	13	14	13	17	33	199	
Producer's Accuracy (%)	93.6	46.2	64.3	76.9	82.4	75.8		Overall Accuracy 83.4
Kappa								0.75

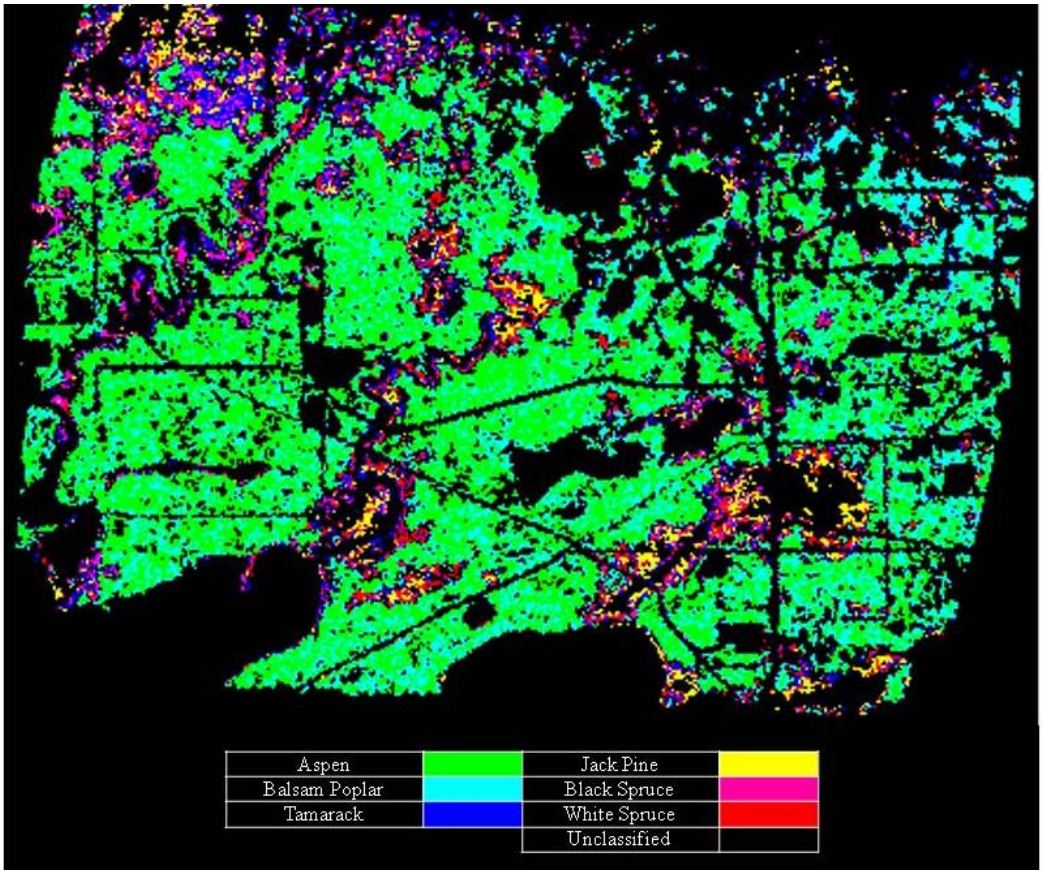

Figure 4.18: Species-cover map of Level 3 classification created from CHRIS/LiDAR data using MESMA unmixing.

Table 4.9: CHRIS/LiDAR Level 3 MESMA classification accuracies.

Map Prediction (Pixels)	Reference (Pixels) Aspen	Poplar	Tamarack	Jack Pine	Black Spruce	White Spruce	Total	User Accuracy (%)
Unclassified	13	4	1	0	0	1	19	
Aspen	86	0	0	0	0	4	90	95.6
Balsam Poplar	3	6	0	0	0	0	9	66.7
Tamarack	4	0	12	1	1	8	26	46.2
Jack Pine	0	0	1	2	2	0	5	40.0
Black Spruce	2	2	0	4	8	7	23	34.8
White Spruce	1	1	0	6	6	13	27	48.1
Total	109	13	14	13	17	33	199	
Producer's Accuracy (%)	78.9	46.2	85.7	15.4	47.1	39.4		Overall Accuracy 63.8
Kappa								0.49

4.8 Comparison of Classification Techniques

For the CHRIS dataset alone, SVM had a higher overall accuracy of 74.4 %, whereas the overall accuracy for the MESMA classification was 43.2 %. In both classification methods, Aspen has a producer accuracy of 94.4 % and 56.0 % using SVM and MESMA, respectively. For the CHRIS/LiDAR dataset, SVM revealed a higher overall accuracy of 83.4 % compared to the overall accuracy of 63.8 % for the MESMA classification. In all classification methods, Aspen was identified as the largest tree species by area (Figure 4.19).

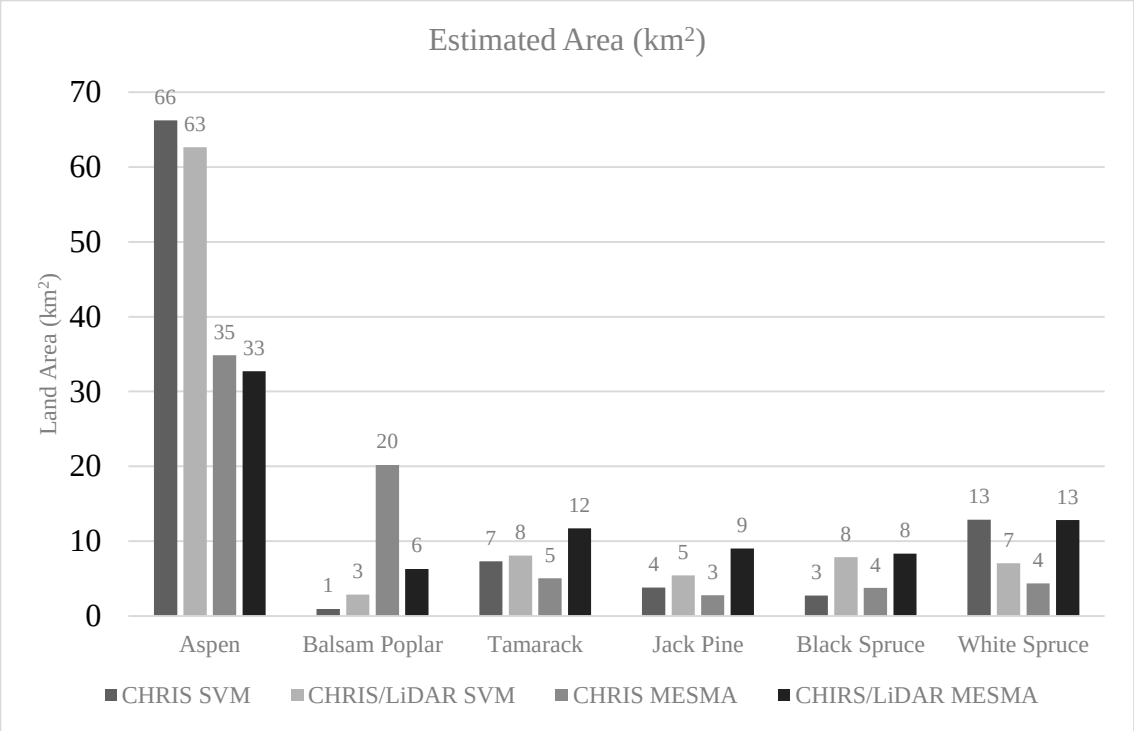

Figure 4.19: Comparison of the estimated land cover for each tree species for each classification method.

A pairwise comparison of the Z-statistics, calculated using McNemar's continuity corrected test, between the CHRIS and CHRIS/LiDAR SVM and between the CHRIS and

CHRIS/LiDAR MESMA Kappa analysis yielded a Z value ($p < 0.05$) of 3.47 and 4.74, respectively (Table 4.10 and 4.11).

Table 4.10: Comparison of the classifications derived from the Level 3 CHRIS and CHRIS/LiDAR datasets using SVM.

		CHRIS (SVM)		
		Correct	Incorrect	Total
CHRIS/LiDAR (SVM)	Correct	145	21	166
	Incorrect	3	30	33
	Total	148	51	

Table 4.11: Comparison of the classifications derived from the Level 3 CHRIS and CHRIS/LiDAR datasets using MESMA.

		CHRIS (MESMA)		
		Correct	Incorrect	Total
CHRIS/LiDAR (MESMA)	Correct	71	56	127
	Incorrect	15	57	72
	Total	86	113	

In comparing the SVM and MESMA classification approaches applied to the two datasets, it was determined that there are significant differences between the classifications results of the two approaches. Thus, the classifications are statistically different from each other.

Chapter 5

Discussion

5.1 Introduction

A comparison of MESMA and supervised SVM classification as a means to determine tree species around reclaimed well-sites is the topic of this thesis. A hierarchical approach was used to refine the classifications from broad land-cover classes down to tree species classes. The difference is that one is a general grouping of land-cover types, while the other is classification of a tree species.

5.2 Pre-Processing

Processing of imagery is a necessity prior to image classification so that any atmospheric effects and sensor artifacts can be minimized or corrected. The aim of this is to correct the image such that it is as if the data were captured on the ground. This helps with classifications at a later date as the only changes to a pixel will be due to spectral changes of the target and not caused by atmospheric attenuation or sensor noise.

De-striping of the CHRIS imagery was accomplished using ESA's BEAM program. The imagery was corrected to surface reflectance using MODTRAN 4 present within BEAM. This produced satisfactory results when evaluated qualitatively. Following atmospheric correction, an abnormally high reflectance value (near 100 %) was present in a couple pixels within the scene. The image was normalized by applying a contrast enhancement, ignoring the anomalous pixel values. Because these anomalies were confined to non-vegetated areas, they were masked for any further processing in the level

1 classification. Upon closer inspection, the anomalies were situated on active well-sites and caused by specular reflection from the facility roofs. The negative reflectance values that were found in the blue region of the spectrum (410 – 472 nm) were caused by atmospheric scattering effects, especially Rayleigh Scattering, when calculating the path radiance that was used to correct the raw signal. An over estimation of the path radiance was used when subtracting from the signal resulting in negative reflectance values.

5.3 Field Parameters

In this study, field parameters were compared to parameters derived from remote sensing with the aim of addressing individual well site reclamation. The parameters included canopy height, canopy fractional cover, and tree species. Nine of the 16 well sites with a reclamation certificate within the image extents did not have trees present and, thus, field data relevant to forest structure were not captured for those sites. However, the remaining seven sites were characterized and compared to reference sites nearby.

The results of comparing the tree species at a well site to tree species measured by remote sensing indicated that using CHRIS data alone with the MESMA classifier was the best predictor of tree species. The correct tree species was identified in nine out of the 16 sites. CHRIS with LiDAR and CHRIS alone using the SVM classifier correctly identified tree species half of the time while CHRIS/ LiDAR with MESMA identified tree species 6 out of the 16 times.

One of the difficulties in comparing tree species between remote sensing and field methods in this study is the fact that only three of the field measured sites contained coniferous species. The other 13 sites were dominated by either Balsam Poplar or Aspen.

Due to the difficulty in separating these species, not only at a spectral level but also in the field, combining these two species into a single class would improve the results.

One of the major challenges in using remote sensing to address reclamation goals in oil and gas well-sites is the size of the well site and the spatial resolution of the remote sensing data used to assess the well site. In this study, the imagery used had a spatial resolution of 34 m, which is a third of the size of a typical well pad (assuming 100 m). The number of spectral signatures contributing to the overall pixel spectral signature is too broad to reliably determine tree species at a particular well site. This is the mixed pixel issue identified in Section 2.5.2.2 and is a function of both the purity of endmembers selected to represent a class as well as the spectral purity of the pixel being unmixed. However, determining the broad tree types (coniferous or deciduous) using medium-spatial resolution imagery is feasible as a broader range of spectral signatures can be considered ‘pure’.

5.4 Level 1 Classification

Separating the broad land-cover types in the first level of classification is necessary in order to ensure that only pixels containing forested areas are included in the subsequent classification and unmixing procedures. The use of SVM in separating the land-cover types yielded excellent results with an overall accuracy of 97.5 %.

The object of this level of classification was separating forest cover from the other land- cover classes. The accuracies obtained for this class were 99.4 % for producer’s accuracy and 98.9 % for user’s accuracy. The reason for the high accuracies across all land-cover types can be attributed to the broad groupings used and the fact that each grouping aimed to exploit large differences in the EM spectrum (i.e., forest compared to

water). However, there was confusion between the grass and wetland classes, where 18.7 % of the wetland pixels were misclassified as grass and 6.6 % of the grass pixels were classified as wetland. This is due to the similar vegetation types present in grasslands and wetlands. Both land-cover types are typically covered by short bushes, such as willow, rather than trees. Although the understory differs between the two, these difference may not be sufficient to spectrally separate the two land-cover types at this time of year and, thus, confusion arises when trying to classify these types. While the results of this classification level are qualitatively good, the number of validation pixels could be increased, especially for wetland to provide more confidence in the classification results. Ideally, the sampling size should be about 10 times the number of bands in the image (Schott, 2007). However, due to the proportions of the land-cover types in the scene, this was not possible. This is more of a note on the accuracy of the classifier and has little bearing on this study, because all land-cover types other than forest were masked for subsequent processing.

5.5 Level 2 Classification

Because trees can be either coniferous or deciduous, this was the next logical level of classification. Due to the relatively large pixel size and mixture of tree canopies (i.e., deciduous and coniferous trees), a third mixed coniferous/deciduous class was also used for classification purposes. At this stage of the classification hierarchy, additional LiDAR data were used together with the hyperspectral data in order to investigate if there is an increase in accuracy with their inclusion. The SVM classification was again used at this level. The results indicate that the inclusion of the LiDAR dataset in the classification yielded better results than using the CHRIS data alone with overall accuracies of 74.4 %

and 70.7 %, respectively. These results are significantly different from each other at a 95-% confidence level when using McNemar's test with the continuity correction.

The class with the lowest accuracies, both producer and user, was the mixed coniferous/deciduous class. For the CHRIS dataset alone, these accuracies are 34.9 % and 62.0 %, respectively with higher accuracies of 45.6 % and 66.2 % for the CHRIS/LiDAR merged dataset, respectively. These low accuracies are expected for a mixed class, because the pixels being classified contain a portion of both coniferous and deciduous tree types. Accordingly, the dominant tree type is likely better represented by the classifier.

While every effort was made to represent the mixed class with representative training samples, some of the dominant tree types may have been selected when creating the training class. Although the overall accuracy of the CHRIS/LiDAR merged dataset was higher than the CHRIS data alone, this may be attributed to the different canopy cover and canopy height present in the coniferous and deciduous classes. Conifers tend to have a cone shaped canopy, while deciduous trees produce a broader canopy (Hosie, 1979). While the LiDAR data can be used to model the shape of a tree, given enough data points, this was not done for this thesis. This distinction can help a classifier determine a tree type in the presences of spectral\similarity between classes.

While these results are promising, especially for the coniferous and deciduous classes, the mixed class had a great deal of confusion. Using a higher spatial resolution may have helped to better classify the tree types by using the canopy shape to further separate the classes. Shape could have been determined by using a valley approach, whereby tree crowns are treated similar to high-elevation topographic features and the

low-elevation features surrounding the former would be the lower leaves (Leckie *et al*., 2003; Gougeon. 1995). However, this approach would not work well in a deciduous forest, where tree canopies are more spread out than cone shaped trees. Shape could have been included in the classification process as another variable used to differentiate tree type and potentially species. Imagery with a spatial resolution of less than 1 m would be ideal to distinguish shape, because tree crowns could be identified. Furthermore, the mixed class could have benefited from spectral unmixing to aid in the classification, because this class, by its nature, contains proportions of both coniferous and deciduous trees.

5.6 Level 3 Classification

The final level of classification involved separating the tree types into tree species using the CHRIS dataset alone and the CHRIS/LiDAR merged data. Classification was performed using the SVM classifier and MESMA unmixing on both datasets. In the case of the latter classification method, the model which best represented the pixel was used to classify the pixel. In comparing the SVM classification results, the CHRIS/LiDAR combination had a superior overall accuracy with 83.4 % compared to the CHRIS dataset alone with 74.4 %. For MESMA, the CHRIS/LiDAR combination again had a higher overall accuracy at 63.8 % compared to the CHRIS dataset alone at 43.2 %. A discussion of the different classification accuracies and reasoning behind them is discussed in the following sections.

5.6.1 SVM

The most accurate tree-type classified in both datasets was the Aspen class with a producer's accuracies of 94.4 % for CHRIS data alone and 93.6 % for CHRIS/LiDAR merged data. These high accuracies in the Aspen class can be attributed to the large proportion of training samples used in the classification. As a result, this class may be over represented and bias the overall accuracy assessment towards that of the Aspen class itself. The least accurate classification of tree species was Jack Pine with a producer's accuracy of 35.3 % for CHRIS data alone and Poplar with a producer's accuracy of 46.2 % for CHRIS/LiDAR merged data. Significant increases in accuracies occurred for coniferous tree species by including LiDAR data in the classification process (Table 5.1). The largest increase was observed for Black Spruce with a difference of 47.1 %. These differences may be attributed to the height differences between Black Spruce and White Spruce (Hosie, 1979). Within the CHRIS classification, Black Spruce was most often confused with White Spruce. However, with the inclusion of the LiDAR structural parameters, these differences could be taken into considerations when classifying the image.

Comparing the results of the classifications of SVM with CHRIS as opposed to CHRIS/LiDAR showed that the results are significantly different when using McNemar's test with the continuity correction

Table 5.1: Comparison of SVM producer's accuracies for conifer species.

	Producers Accuracies	
Tree Type	CHRIS	CHRIS/LiDAR
Tamarack	46.2 %	64.3 %
Jack Pine	30.8 %	76.9 %
Black Spruce	35.3 %	82.4 %
White Spruce	72.7 %	75.8 %

The moderate pixel size of 34 m resulted in mixed pixels containing spectra from both understory and other tree species. This mixing coupled with a low number of representative reference pixels may be painting an un-realistic picture of the true merit of the classification methods and accuracies. Although relatively high accuracies were obtained using the SVM classification, this classification method at this spatial resolution may be better suited for broad stand-level studies rather than small well-pad size studies.

5.6.2 MESMA

Similar to the SVM classification, Aspen was again the species most accurately classified with a producer's accuracies of 56.0 % for CHRIS alone and 78.9 % for CHRIS/LiDAR when using MESMA. Again, this can be attributed to the large proportion of Aspen training samples used for the classification. The least accurate class for the CHRIS and CHRIS/LiDAR data was Jack Pine with a producer's accuracy of 15.4 %. The same accuracies for both classifications is peculiar but not unexpected. Due to the small validation sample size of 13 pixels there is less of a chance for pixels to differ from the training sample. Within the CHRIS classification, Tamarack and Jack Pine classes were the most often confused with 38.5 % of the Jack Pine pixels being classified as Tamarack. Because these trees typically grow in different environmental regimes, Jack Pine in elevated, drier areas and Tamarack in lower, wetter areas (Hosie, 1979).This

result suggests that greater pixel separability and more pure Jack Pine endmembers are required to successfully classify the Jack Pine species. Another 38.5 % of the Jack Pine training pixels were determined to be unclassified. Within the CHRIS/LiDAR classification, White Spruce was most often confused with Tamarack with 46.1% of the Tamarack pixels being classified as White Spruce. This result is understandable as these tree species favour the same environmental regimes.

The results of the McNemar's tests on the CHRIS and CHRIS/LiDAR MESMA classified images indicate that the classifications are significantly different from one another.

Due to the moderate pixel size, the chances of a pixel being spectrally pure are minute. This necessitates the need for a spectral unmixing procedure in order to determine what the contributing spectra of a pixel are. MESMA is able to determine the contributions of different spectra to the overall pixel spectrum. From these contributions, the spectrum that contributes the most is assigned to that pixel. This method works well on high-spatial resolution pixels where the number of contributing spectra is lower. However, at moderate spatial resolutions, this method creates confusion in classifying an image as there can be a large number of spectrally distinct features within the spatial distribution of a pixel. Logically grouping tree species based on their preference of growing areas can be more representative of what is happening on the ground. For example, Tamarack and Black Spruce are often found growing together in low lying wetland areas (Hosie, 1979). Creating a model with endmembers made up of these two tree species is a better representation of these low lying areas. The same can be said for species such as White Spruce and Jack Pine. Creating a spectral library from trees within

the study area using high quality spectral signatures from which to base the model on, would also aid in increasing the classification accuracies.

A four end-member model, as opposed to a three-endmember model, was also run on the mixed and coniferous tree species in order to determine if the inclusion of another endmember into the MESMA modelling would increase the accuracies substantially. This was not found to be the case with an approximate 1-% decrease in overall accuracies from the three-endmember models to the four-endmember models.

5.7 Hierarchical Approach

The hierarchical approach used in this study is well suited to tree-species classification as it spatially constrains the data in order to decrease the complexity, in this case from broad land-cover to tree species. This method can be further improved upon leveraging tree growing preferences. Using the DEM created from the LiDAR data, a further spatial constraint can be applied based on drainage patterns. Using this elevation constraint, tree types and furthermore tree species could have been better differentiated by observing their preferred growing regimes. Low-lying drainage areas are more likely to attract coniferous species, such as Tamarack and Black spruce. Accordingly, confining tree species to these areas would have allowed for better classification.

In any system that builds upon previous methods, there is a chance for error propagation from the first level to the last level. In the case of this study, the actual classification of tree species is predicated on correctly identifying land-cover types. For example, forests that are misclassified as another land-cover type are removed from any further classification. Conversely, land-covers that are misclassified as being forest contribute spectra that can reduce classification accuracies as they are significantly

different than training spectra. This can occur in transitional zones, such as when a forest ends and grassland begins. The inclusion of a mixed class attempted to account for any error propagation from level 2. Rather than declaring a tree to be either coniferous or deciduous, the mixed class allowed for some ambiguity to occur within the spectral profile of pixels within this class. This ambiguity can increase the classification error for tree species identified as being of the mixed class in level 2 when in reality their classification should be either coniferous or deciduous.

The area for which this study was conducted has a wide variation in land-cover types, which helps with the hierarchical constraints. A heavily treed area or an area with heavily mixed coniferous and deciduous trees may not benefit from this approach as there would an abundance of mixed pixels which would hamper further constrain mapping efforts. However, the results of this study highlight the merit of a hierarchical approach.

Chapter 6

Conclusions

Understanding what tree species are present and the distribution of those tree species in oil and gas well sites and surrounding areas is an important part of a cumulative reclamation plan. Utilizing remote sensing has the following two main benefits for reclamation management: a) The technology is able to cover large areas of land in a timely fashion and b) it allows for repeat measurements almost whenever a user desires.

This thesis investigated the utility of combining two remote sensing datasets, hyperspectral satellite and airborne LiDAR imaging in order to classify tree species in and around abandoned well-sites in Cold Lake, Alberta. To facilitate the classification procedure, the following two methods of classifying tree species were investigated, SVM and MESMA. A hierarchical classification approach was used to spatially constrain the remote sensing data at different classification levels in order to reduce spectral confusion. The accuracy of the different classification approaches and the different remote sensing datasets was evaluated and compared. The methods used in this thesis demonstrate an approach for reclamation personnel when attempting to determine tree species in and around abandoned well-sites.

This thesis answers the hypothesis of whether combining structural and spectral data over a forested reclamation site will yield classification results that are more accurate than spectral data alone. Indeed, the results of this study show that through the inclusion of structural data, an increase of 9 % for SVM and 22 % for MESMA in classification accuracy can be observed. These results were tested against spectral data alone and are

significantly different at a 95 % confidence level. Furthermore, while spectral unmixing methods can provide the estimated proportions of endmembers within each pixel, their use with medium spatial resolution images is not as good as compared to discrete classification methods.

The results from the third and final level of classification using SVM and MESMA showed that a combined dataset of hyperspectral and LiDAR data achieved the best result distinguishing tree species. Results show that classifying a combined CHRIS/LiDAR dataset with SVM produced the most accurate tree species classification with 83.4 % compared to the CHRIS data alone (74.4 %). SVM was also superior to MESMA in the final level of classification. MESMA classification achieved overall accuracies of 64.3 % using CHRIS/LiDAR data and 44.2 % using CHRIS data alone. All classified datasets showed confusion between Black Spruce and White Spruce and Jack Pine and Tamarack. However, the addition of LiDAR data to the analysis improved the ability of the classifiers to identify these species. The most accurate tree species identified in all classifications was Aspen. This is to be expected as Aspen had a larger training and validation size and also due to the fact that the majority of tree species that were measured on the ground were Aspen.

The combined dataset approach used in this thesis demonstrates that through the inclusion of structural data in the form of Canopy Height Models and Canopy Cover Models, tree-species classification accuracies can be increased. The approaches identified in this study can be improved upon through the use of higher spatial resolution imagery and through acquiring data at different time periods in order to build a temporal dataset. The increased spatial resolution will help in reducing mixed pixels that were a major

source of confusion in this study, and the use of a temporal dataset will further help in separating coniferous and deciduous tree species through leaf-off and leaf-on conditions. However, with an increase in spatial resolution the amount of area that can be covered by a single image diminishes. This can reduce the effectiveness of tree-species identification for reclamation monitoring over large areas.

This study has demonstrated that using hyperspectral and LiDAR datasets to determine tree species in and around abandoned well-sites is feasible and indeed can be used to assess reclamation objectives. However, problems can arise when attempting to differentiate certain tree-species from one another. For example, Aspen and Balsam Poplar are extremely difficult to separate both spectrally and in the field. This is demonstrated in both CHRIS and CHRIS/LiDAR datasets when using SVM or MESMA. A more encompassing approach to reclamation management using tree-species classification may be to determine if there are significant differences in the types of trees (coniferous or deciduous) in a well-site compared to the trees around a well-site. This approach could act as a flag, whereby a more in depth field analysis should be conducted.

References

Adams, J.B., Smith, M.O., and Gillespie, A.R. (1993). Imaging spectroscopy: Interpretation based on spectral mixture analysis. *Remote geochemical analysis: Elemental and mineralogical composition*, *7*: 145-166.

Alberta Environment and Sustainable Resource Development (AESRD). (1995). Reclamation Criteria for Wellsites and Associated Facilities for Forested Lands - 1995 Update. Edmonton, Alberta, p. 81.

Alberta Environment and Sustainable Resource Development (AESRD) (2011) Reclamation Criterea - Fact Sheet. Retrieved on October 10, 2015. From <http://esrd.alberta.ca/lands-forests/land-industrial/programs-and-services/reclamation-and-remediation/upstream-oil-and-gas-reclamation-and-remediation-program/documents/2010-ReclamationCriteria-Factsheet.pdf>.

Alberta Environment and Sustainable Resource Development (AESRD). (2012). Update Report on Alberta Environment and Sustainable Resource Development's Upstream Oil and Gas Reclamation Certificate Program, March 31, 2012. Edmonton, Alberta. (p. 9).

Alberta Environment and Sustainable Resource Development (AESRD) (2013a) Oil and Gas Well Reclamation. Retrieved on October 10, 2015. From <http://esrd.alberta.ca/lands-forests/land-industrial/programs-and-services/reclamation-and-remediation/upstream-oil-and-gas-reclamation-and-remediation-program/documents/2010-ReclamationCriteria-Factsheet.pdf>.

Alberta Environment and Sustainable Resource Development (AESRD). (2012b). Update Report on Alberta Environment and Sustainable Resource Development's Upstream Oil and Gas Reclamation Certificate Program, March 31, 2012. Edmonton, Alberta, p. 9.

Alberta Environment and Sustainable Resource Development (AESRD). (2013a). 2010 Reclamation Criteria for Wellsites and Associated Facilities for Forested Lands (Updated July 2013). Edmonton, Alberta, p. 81.

Alberta Environment and Sustainable Resource Development (AESRD). (2013b). 2010 Reclamation Criteria for Wellsites and Associated Facilities for Cultivated Lands (Updated July 2013). Edmonton, Alberta, p. 92.

Alberta Environment and Sustainable Resource Development (AESRD). (2013c). 2010 Reclamation Criteria for Wellsites and Associated Facilities for Native Grasslands (Updated July 2013). Edmonton, Alberta, p. 92.

Alberta Environment and Sustainable Resource Development (AESRD). (2014) Oil and Gas Wells Reclamation. Retrieved on September 12, 2014 from

<http://esrd.alberta.ca/focus/state-of-the-environment/land/response-indicators/oil-and-gas-wells-reclamation.aspx>.

Anderson, J.E., Plourde, L.C., Martin, M.E., Braswell, B.H., Smith, M.-L., Dubayah, R.O., Hofton, M.A., and Blair, J.B. (2008). Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. *Remote Sensing of Environment, 112*: 1856-1870.

Bachelet, D., Neilson, R.P., Lenihan, J.M., & Drapek, R.J. (2001). Climate change effects on vegetation distribution and carbon budget in the United States. *Ecosystems, 4*: 164-185.

Bardos, R.P., Mariotti, C., Marot, F., and Sullivan, T. (2001). Framework for Decision Support used in Contaminated Land Management in Europe and North America, Land Contaminated and Reclamation. *Evaluation of demonstrated and emerging technologies for the treatment of contaminated land and groundwater (phase III) 2000 special session decision support tools* (pp. 149-163).

Bartzen, B.A., Dufour, K.W., Clark, R.G., and Caswell, F.D. (2010). Trends in agricultural impact and recovery of wetlands in prairie Canada. *Ecological Applications, 20*: 525-538.

Berk, A., Anderson, G.P., Bernstein, L.S., Acharya, P.K., Dothe, H., Matthew, M.W., Adler-Golden, S.M., Chetwynd Jr, J.H., Richtsmeier, S.C., and Pukall, B. (1999). MODTRAN4 radiative transfer modeling for atmospheric correction. Proceedings of SPIE's International Symposium on Optical Science, Engineering, and Instrumentation, (Denver, Colorado, U.S.A) pp. 348-353.

Boardman, J.W., Kruse, F.A., and Green, R.O. (1995). Mapping target signatures via partial unmixing of AVIRIS data. Summaries 5th JPL Airborne Earth Science Workshop, vol 1. 995: 23–26.

Booth, D., and Cox, S. (2009). Dual-camera, high-resolution aerial assessment of pipeline revegetation. *Environmental Monitoring and Assessment, 158*: 23-33.

Bradshaw, A.D. (1984). Ecological principles and land reclamation practice. *Landscape Planning, 11*: 35-48.

Brandtberg, T., Warner, T.A., Landenberger, R.E., and McGraw, J.B. (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment, 85*: 290-303.

(2010) CHRIS Proba - Theoretical Details. Retrieved on January 15 from <http://www.brockmann-consult.de/beam/chris-box/theory-details.html#Atmospheric_Correction>.

Brockmann-Consult. (2013) CHRIS Toolbox for BEAM. Retrieved on November 2, 2013 from <http://www.brockmann-consult.de/beam/chris-box/index.html>.

Burger, J. (2008). Environmental management: Integrating ecological evaluation, remediation, restoration, natural resource damage assessment and long-term stewardship on contaminated lands. *Science of The Total Environment, 400*: 6-19.

Canty, M.J. (2006). *Image analysis, classification and change detection in remote sensing: with algorithms for ENVI/IDL*. CRC Press, 513 pages.

Carter, G.A. (1993). Responses of Leaf Spectral Reflectance to Plant Stress. *American Journal of Botany, 80*: 239-243.

Canadian Forestry Service (CFS). (2008) Canada's National Forest Inventory ground sampling guidelines [electronic resource]: specifications for ongoing measurement. Version 5.0. Retrieved on February 9, 2014 from <https://nfi.nfis.org/ground_plot.php?lang=en>.

Champagne, C.M., Abuelgasim, A., Staenz, K., Monet, S., and White, H.P. (2004). Ecological Restoration from Space: The use of Remote Sensing for Monitoring Land Reclamation in Sudbury. Proceedings of the16th International Conference Society for Ecological Restoration. Victoria, British Columbia, Canada. pp.1-7.

Chi, M., Feng, R., and Bruzzone, L. (2008). Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem. Advances in Space Research, *41*: 1793-1799.

Clark, M.L., Roberts, D.A., and Clark, D.B. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment, 96*: 375-398.

Coburn, C., and Roberts, A. (2004). A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing, 25*: 4287-4308.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin, 70*: 213.

Congalton, R.G. (1991a). Remote Sensing and Geographic Information System Data Integration: Error Sources and. *Photogrammetric Engineering and Remote Sensing, 57*: 677-687.

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

Congalton, R.G., and Green, K. (2008). *Assessing the accuracy of remotely sensed data: principles and practices*. (2 ed.). CRC Press, 192 pages.

Cook, B.D., Bolstad, P.V., Næsset, E., Anderson, R.S., Garrigues, S., Morisette, J.T., Nickeson, J., and Davis, K.J. (2009). Using LiDAR and quickbird data to model plant production and quantify uncertainties associated with wetland detection and land cover generalizations. *Remote Sensing of Environment, 113*: 2366-2379.

Craig, N., Turner, R., and Day Jr, J. (1979). Land loss in coastal Louisiana (U.S.A). *Environmental Management, 3*: 133-144.

Curtis, D.R., and Campopiano, M. (2013). Medieval land reclamation and the creation of new societies: comparing Holland and the Po Valley, c.800–c.1500. *Journal of Historical Geography, in press.*

Dalponte, M., Bruzzone, L., and Gianelle, D. (2008). Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas. *Ieee Transactions on Geoscience and Remote Sensing, 46*: 1416-1427.

Dalponte, M., Bruzzone, L., and Gianelle, D. (2012). Tree species classification in the Southern Alps based on the fusion of very high geometrical resolution multispectral/hyperspectral images and LiDAR data. *Remote Sensing of Environment, 123*: 258-270.

Dalponte, M., Orka, H.O., Gobakken, T., Gianelle, D., and Næsset, E. (2013). Tree species classification in boreal forests with hyperspectral data. *Ieee Transactions on Geoscience and Remote Sensing, 51*: 2632-2645.

United Kingdom Department for Environment, Food and Rural Affairs (DEFRA). (2004). CLR 11: Model Procedures for the Management of Land Contamination. Bristol, United Kingdom. 192 pages.

Delegido, J., Fernandez, G., Gandia, S., and Moreno, J. (2008). Retrieval of chlorophyll content and LAI of crops using hyperspectral techniques: Application to PROBA/CHRIS data. *International Journal of Remote Sensing, 29*: 7107-7127.

Deng, J., Wang, K., Deng, Y., and Qi, G. (2008). PCA-based land-use change detection and analysis using multitemporal and multisensor satellite data. *International Journal of Remote Sensing, 29*: 4823-4838.

Dennison, P.E., Halligan, K.Q., and Roberts, D.A. (2004). A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment, 93*: 359-367.

Dennison, P.E., and Roberts, D.A. (2003a). The effects of vegetation phenology on endmember selection and species mapping in southern California chaparral. *Remote Sensing of Environment, 87*: 295-309.

Dennison, P.E., and Roberts, D.A. (2003b). Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sensing of Environment, 87*: 123-135.

Dubayah, R.O., and Drake, J.B. (2000). Lidar Remote Sensing for Forestry. *Journal of Forestry, 98*: 44-46.

Duda, R.O., and Hart, P.E. (1973). Pattern recognition and scene analysis. In: Wiley, New York, 512 pages.

Environment Canada. (2013) Canadian Climate Normals. Retrieved on February 4, 2014 from <http://climate.weather.gc.ca/climate_normals>.

United States Environmental Protection Agency (EPA). (1992). Reclamation and Redevelopment of Contaminated Land: Volume II. European Case Studies. Washington, DC, U.S.A. 302 pages.

Eurpean Space Agency (ESA). (2014) PROBA-1 CHRIS Data Product. Retrieved on January 2, 2014 from <http://earth.esa.int/web/guest/-/proba-chris-level-1a-1488>.

Evans, J.S., Hudak, A.T., Faux, R., and Smith, A. (2009). Discrete return lidar in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote sensing, 1*: 776-794.

Federal Law Gazette I, Waste Disposal Act (1977), Federal Law Gazette, 36 pages.

Foody, G.M. (1992). On the compensation for chance agreement in image classification accuracy assessment. *Photogrammetric Engineering and Remote Sensing, 58*: 1459-1460.

Foody, G.M., and Embashi, M.R.M. (1995). Mapping despoiled land cover from Landsat thematic mapper imagery. *Computers, Environment and Urban Systems, 19*: 249-260.

Foody, G.M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*: 185-201.

Foody, G.M. (2004). Thematic Map Comparison. *Photogrammetric Engineering and Remote Sensing, 70*: 627-633.

Foody, G.M. (2006). The evaluation and comparison of thematic maps derived from remote sensing. *7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences.* Lisbon, Portugal, pp. 18-31.

Foody, G.M., and Mathur, A. (2004). A relative evaluation of multiclass image classification by support vector machines. IEEE *Transactions on Geoscience and Remote Sensing, 42*: 1335-1343.

Foody, G.M., and Mathur, A. (2006). The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment, 103*: 179-189.

Franke, J., Roberts, D.A., Halligan, K., and Menz, G. (2009). Hierarchical multiple endmember spectral mixture analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment, 113*: 1712-1723.

Franklin, S.E. (2001). *Remote sensing for sustainable forest management*. CRC Press, 424 pages.

Galidaki, G., Gitas, I., and Katagis, T. (2012). Forest species differentiation and mapping using EO-1 Hyperion imagery. Proceedings of First Sentinel-2 Preparatory Symposium. Frascati, Italy, pp. 707.

Gao, B.-C., Montes, M.J., Davis, C.O., and Goetz, A.F. (2009). Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean. *Remote Sensing of Environment, 113*: S17-S24.

Genderen, J. L. van, and Pohl, C. (1994). Image fusion: Issues, techniques and applications., Proceedings EARSeL Workshop on Intelligent Image Fusion, Strasbourg, France, 11 September 1994, edited by J. L. van Genderen and V. Cappellini (Enschede: ITC), pp. 18- 26.

Ghiyamat, A., and Shafri, H.Z.M. (2010). A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. *International Journal of Remote Sensing, 31*: 1837-1856.

Glenn, N.F., Spaete, L.P., Sankey, T.T., Derryberry, D.R., Hardegree, S.P., and Mitchell, J.J. (2011). Errors in LiDAR-derived shrub height and crown area on sloped terrain. *Journal of Arid Environments, 75*: 377-382.

Gómez-Chova, L., Alonso, L., Guanter, L., Camps-Valls, G., Calpe, J., and Moreno, J. (2008). Correction of systematic spatial noise in push-broom hyperspectral sensors: application to CHRIS/PROBA images. *Applied Optics, 47*: F46-F60.

Gong, P., and Howarth, P.J. (1990). The use of structural information fo imporovinf land-cover classification accuracies at the rural-urba fringe. Photogrammetric Engineering and Remote Sensing, 56.1: 67-73

Goodenough, D.G., Dyk, A., Niemann, K.O., Pearlman, J.S., Hao, C., Tian, H., Murdoch, M., and West, C. (2003a). Processing Hyperion and ALI for forest classification. *IEEE Transactions on Geoscience and Remote Sensing, 41*: 1321-1331.

Goodenough, D.G., Hao, C., Dyk, A., Tian, H., McDonald, S., Murdoch, M., Niemann, K.O., Pearlman, J., and West, C. (2003b). EVEOSD forest information products from AVIRIS and Hyperion. Proceedings of the International *Geoscience and Remote Sensing Symposium, (IGARSS 2003). Toulouse, France, Vol. 281* pp. 284-287.

Goodwin, N.R., Coops, N.C., Wulder, M.A., Gillanders, S., Schroeder, T.A., and Nelson, T. (2008). Estimation of insect infestation dynamics using a temporal sequence of Landsat data. *Remote Sensing of Environment, 112*: 3680-3689.

Government of Alberta, Surface Reclamation Act (1963, chapter. 64), Alberta Queen's Printer, p. 10.

Government of Alberta, Land Surface Conservation and Reclamation Act (1973, chapter. 34), Alberta Queen's Printer, p. 49.

Government of Alberta, Environmental Protection and Enhancement Act (1992, chapter. E-13), Alberta Queen's Printer, p. 152.

Government of Alberta, Conservation and Reclamation Regulation (1993, AR 115/1993), Alberta Queen's Printer, p. 17.

Government of Alberta, Environmental Protection and Enhancement Act, RSA 2000 (2013, chapter. E-12), Alberta Queen's Printer, p. 158.

Guanter, L., Alonso, L., and Moreno, J. (2005). A method for the surface reflectance retrieval from PROBA/CHRIS data over land: Application to ESA SPARC campaigns. *IEEE Transactions on Geoscience and Remote Sensing, 43*: 2908-2917.

Guanter, L., Richter, R., and Moreno, J. (2006). Spectral calibration of hyperspectral imagery using atmospheric absorption features. *Applied Optics, 45*: 2360-2370.

Hall, D. L. (1992). Mathematical techniques in multisensor data fusion (Norwood: Artech House Inc.).

Halpern, C.B., and Spies, T.A. (1995). Plant species diversity in natural and managed forests of the Pacific Northwest. *Ecological Applications, 5*: 913-934.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., Egorov, A., Chini, L., Justice, C.O., and Townshend, J.R.G. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science, 342*: 850-853.

Heinzel, J., and Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation, 18*: 101-110.

Hill, R., Wilson, A., George, M., and Hinsley, S. (2010). Mapping tree species in temperate deciduous woodland using time-series multi-spectral data. *Applied Vegetation Science, 13*: 86-99.

Hill, R.A., and Thomson, A.G. (2005). Mapping woodland species composition and structure using airborne spectral and LiDAR data. *International Journal of Remote Sensing, 26*: 3763-3779.

Hodgson, M. E., Jensen, J. R., Mackey, J. E., Jr, and Coulter, M. C. (1988), Monitoring wood stork foraging habitat using remote sensing and geographic information systems. *Photogrammetric Engineering & Remote Sensing*. 54: 1601–1607.

Hoeksema, R.J. (2007). Three stages in the history of land reclamation in the Netherlands. *Irrigation and Drainage, 56*: S113-S126.

Holmgren, J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scandinavian Journal of Forest Research, 19*: 543-553.

Holmgren, J., Persson, Å., and Söderman, U. (2008). Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. *International Journal of Remote Sensing, 29*: 1537-1552.

Hopkinson, C., and Chasmer, L. (2009). Testing LiDAR models of fractional cover across multiple forest ecozones. *Remote Sensing of Environment, 113*: 275-288.

Hosie, R. (1979). *Native Trees of Canada*. (8th ed.). Fitzhenry and Whiteside Ltd. in co-operation with the Canadian Forestry Service (Environment Canada) and the Canadian Government Publishing Centre, 380 pages.

Huang, C., Davis, L., and Townshend, J. (2002). An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing, 23*: 725-749.

Huete, A., and Didan, K. (2004). MODIS seasonal and inter-annual responses of semiarid ecosystems to drought in the Southwest U.S.A. Proceedings of the International *Geoscience and Remote Sensing Symposium, (IGARSS 2004). Anchorage, Alaska, U.S.A., Vol. 1533* pp. 1538-1541.

Hurd, J.D., Wilson, E.H., Lammey, S.G., and Civco, D.L. (2001). Characterization of forest fragmentation and urban sprawl using time sequential Landsat imagery. Proceedings of the *ASPRS Annual Convention, St. Louis, Missouri, U.S.A., pp.2001*.

Husch, B., Beers, T.W., and Kershaw Jr, J.A. (2002). *Forest mensuration*. (3rd ed.). J.Wiley, 456 pages.

Hyde, P., Dubayah, R., Walker, W., Blair, J.B., Hofton, M., and Hunsaker, C. (2006). Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment, 102*: 63-73.

INRA-CSE, Avignon. (2002) Comparing VALERI sampling schemes to better represent high spatial resolution satellite pixel from ground measurements: How to characterize an ESU. Retrieved on February 12, 2014 from <http://w3.avignon.inra.fr/valeri/methodology/samplingschemes.pdf>.

INRA (2013) CAN-EYE, Version 6.11 from <http://www6.paca.inra.fr/can-eye/>.

Jennings, S., Brown, N., and Sheil, D. (1999). Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. *Forestry, 72*: 59-74.

Jensen, J.R. (2005). *Introductory digital image processing: a remote sensing perspective*. (3 ed.). Upper Saddle River, NJ: Pearson Prentice Hall, 544 pages.

Jones, T.G., Coops, N.C., and Sharma, T. (2010). Assessing the utility of airborne hyperspectral and LiDAR data for species distribution mapping in the coastal Pacific Northwest, Canada. *Remote Sensing of Environment, 114*: 2841-2852.

Jung, J. (2011). Utilization of full waveform LIDAR and hyperspectral data for forest structure characterization. *Ph.D dissertation., Purdue University, Indiana, U.S.A. 118 pages.*

Kraus, K., and Pfeifer, N. (1998). Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 53*: 193-203.

Kraus, K., and Pfeifer, N. (2001). Advanced DTM generation from LIDAR data. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences, 34*: 23-30.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., and Goetz, A.F.H. (1993). The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment, 44*: 145-163.

Lathrop, R. G., Aber, J. D., Bognar, J. A., Ollinger, S. V., Casset, S., and Ellis, J. M. (1994), GIS development to support regional simulation modeling of northeastern (USA) forest ecosystems. In *Environmental Information Management and Analysis* (W. Michener, J. W. Brunt, and S. Stafford, Eds.), Taylor and Francis, London, pp. 431–451

Lee, K. S., Cohen, W. B., Kennedy, R. E., Maiersperger, T. K., and Gower, S. T. (2004). Hyperspectral versus multispectral data for estimating leaf area index in four different biomes. Remote Sensing of Environment, 91(3), 508-520.

Lefsky, M.A., Cohen, W.B., and Spies, T.A. (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Canadian Journal of Forest Research, 31*: 78-87.

Lévesque, J., and Staenz, K. (2004). A method for monitoring mine tailings revegetation using hyperspectral remote sensing. *Proceedings of the International Geoscience and Remote Sensing Symposium, (IGARSS 2004). Anchorage, Alaska, U.S.A., Vol. 1 pp. 575-587.*

Lévesque, J., and Staenz, K. (2008). Monitoring mine tailings revegetation using multitemporal hyperspectral image data. *Canadian Journal of Remote Sensing, 34*: 172-186.

Lévesque, J., Staenz, K., Shang, J., Neville, B., Yearwood, P., and Singhroy, V. (1999). Temporal monitoring of mine tailings revegetation using hyperspectral data. *Proceedings of the Thirteenth International Conference on Applied Geologic Remote Sensing*. Sudbury, Ontario, Canada. pp. 21-28.

Losee, S.T.B. (1942) Air Photographs and Forest Sites. *The Forestry Chronicle, 18(3)*: 129-144.

Li, L., Ustin, S.L., and Lay, M. (2005). Application of AVIRIS data in detection of oil-induced vegetation stress and cover change at Jornada, New Mexico. *Remote Sensing of Environment, 94*: 1-16.

Liang, S., Fang, H., and Chen, M. (2001). Atmospheric correction of Landsat ETM+ land surface imagery. I. Methods. *IEEE Transactions on Geoscience and Remote Sensing, 39*: 2490-2498.

Lillesand, T.M., Kiefer, R.W., and Chipman, J.W. (2004). *Remote sensing and image interpretation (5th ed.)*. John Wiley and Sons Ltd. 784 pages.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography, 27*: 88-106.

Liu, G., Ding, X., Chen, Y., Li, Z., and Li, Z. (2001). Ground settlement of Chek Lap Kok Airport, Hong Kong, detected by satellite synthetic aperture radar interferometry. *Chinese Science Bulletin, 46*: 1778-1782.

Lu, D., and Weng, Q. (2004). Spectral mixture analysis of the urban landscape in Indianapolis with Landsat ETM+ imagery. *Photogrammetric Engineering and Remote Sensing, 70*: 1053-1062.

Mallet, C., and Bretar, F. (2009). Full-waveform topographic lidar: State-of-the-art. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*: 1-16.

Martin, M.E., Newman, S.D., Aber, J.D., and Congalton, R.G. (1998). Determining Forest Species Composition Using High Spectral Resolution Remote Sensing Data. *Remote Sensing of Environment, 65*: 249-254.

Masek, J., Lindsay, F., and Goward, S. (2000). Dynamics of urban growth in the Washington DC metropolitan area, 1973-1996, from Landsat observations. *International Journal of Remote Sensing, 21*: 3473-3486.

Maune, D.F. (2007). *Digital elevation model technologies and applications: The DEM users manual*. (2nd ed.). American Society for Photogrammetry and Remote Sensing. 655 pages.

McGaughey, R.J. (2007) FUSION/LDV: Software for LIDAR Data Analysis and Visualization., Version 3.4.2 from <http://www.fs.fed.us/eng/rsac/fusion/>.

Ministry of Infrastructure and the Environment (1996). Soil Protection Act (Wet bodembescherming - Wbb) (p. 36).

Mountrakis, G., Im, J., and Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*: 247-259.

Mumby, P.J., Green, E.P., Edwards, A.J., and Clark, C.D. (1999). The cost-effectiveness of remote sensing for tropical coastal resources assessment and management. *Journal of Environmental Management, 55*: 157-166.

Myneni, R.B., Maggion, S., Iaquinta, J., Privette, J.L., Gobron, N., Pinty, B., Kimes, D.S., Verstraete, M.M., and Williams, D.L. (1995). Optical remote sensing of vegetation: Modeling, caveats, and algorithms. *Remote Sensing of Environment, 51*: 169-188.

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing, 52*: 49-56.

Næsset, E., and Økland, T. (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment, 79*: 105-115.

Nascimento, J.M., and Bioucas Dias, J.M. (2005). Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 43*: 898-910.

Neville, R.A., Sun, L., and Staenz, K. (2004). Detection of keystone in imaging spectrometer data. *Defense and Security* (pp. 208-217).

Neville, R.A., Sun, L., and Staenz, K. (2008). Spectral calibration of imaging spectrometers by atmospheric absorption feature matching. *Canadian Journal of Remote Sensing, 34*: S29-S42.

Nelson, R. F., Latty, R. S., and Mott, G. (1985), Classifying northern forests using Thematic Mapper Simulator data. *Photogrammetric Engineering & Remote Sensing*. 50: 607–617.

North Atlantic Treaty Organization - Committee on the Challenges of Modern Society (NATO/CMSS). (1998). Evaluation of Demonstrated and Emerging Technologies for the Treatment and Clean Up of Contaminated Land and Groundwater - (Phase II). Washington, DC., U.S.A. 282 pages.Olshansky, R.B. (2006). Planning after hurricane Katrina. *Journal of the American Planning Association, 72*: 147-153.

Oil Sands Research and Information Network. (2011). Equivalent Land Capability Workshop Summary Notes. Edmonton, Alberta, Canada. (p. 83).

Ozesmi, S.L., and Bauer, M.E. (2002). Satellite remote sensing of wetlands. *Wetlands Ecology and Management, 10*: 381-402.

Parliament of the United Kingdom (1936). Public Health Act 1936,The National Archives. (p. 247).

Parliament of the United Kingdom (1947), Town and County Planning Act 1947, Blackwell Publishing Ltd. (pp. 72-85).

Peddle, D.R., Hall, F.G., and LeDrew, E.F. (1999). Spectral Mixture Analysis and Geometric-Optical Reflectance Modeling of Boreal Forest Biophysical Structure. *Remote Sensing of Environment, 67*: 288-297.

Peddle, D.R., Johnson, R.L., Cihlar, J., and Latifovic, R. (2004). Large area forest classification and biophysical parameter estimation using the 5-Scale canopy reflectance model in Multiple-Forward-Mode. *Remote Sensing of Environment, 89*: 252-263.

Peddle, D.R., Johnson, R.L., Cihlar, J., Leblanc, S.G., Chen, J.M., and Hall, F.G. (2007). Physically based inversion modeling for unsupervised cluster labeling, independent forest classification, and LAI estimation using MFM-5-Scale. *Canadian Journal of Remote Sensing, 33*: 214-225.

Pezeshki, S., Hester, M., Lin, Q., and Nyman, J. (2000). The effects of oil spill and clean-up on dominant US Gulf coast marsh macrophytes: a review. *Environmental Pollution, 108*: 129-139.

Plourde, L.C., Ollinger, S.V., Smith, M.-L., and Martin, M.E. (2007). Estimating Species Abundance in a Northern Temperate Forest Using Spectral Mixture Analysis. *Photogrammetric Engineering and Remote Sensing, 73*: 11.

Popescu, S.C., and Wynne, R.H. (2004). Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering and Remote Sensing, 70*: 589-604.

Popescu, S.C., Wynne, R.H., and Nelson, R.F. (2002). Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture, 37*: 71-95.

Powter, C., Chymko, N., Dinwoodie, G., Howat, D., Janz, A., Puhlmann, R., Richens, T., Watson, D., Sinton, H., and Ball, K. (2012). Regulatory history of Alberta's industrial land conservation and reclamation program. *Canadian Journal of Soil Science, 92*: 39-51.

Price, J.C. (1994). How unique are spectral signatures? *Remote Sensing of Environment, 49*: 181-186.

Pringle, M.J., Schmidt, M., and Muir, J.S. (2009). Geostatistical interpolation of SLC-off Landsat ETM+ images. *ISPRS Journal of Photogrammetry and Remote Sensing, 64*: 654-664.

Puttonen, E., Suomalainen, J., Hakala, T., Räikkönen, E., Kaartinen, H., Kaasalainen, S., and Litkey, P. (2010). Tree species classification from fused active hyperspectral reflectance and LIDAR measurements. *Forest Ecology and Management, 260*: 1843-1852.

Rabe, A., Jakimow, B., Held, M., van der Linden, S., and Hostert, P. (2014) EnMAP-Box, Version 2.0 from <www.enmap.org>.

Richter, R. (1997). Correction of atmospheric and topographic effects for high spatial resolution satellite imagery. *International Journal of Remote Sensing, 18*: 1099-1111.

Roberts, D.A., Dennison, P.E., Gardner, M.E., Hetzel, Y., Ustin, S.L., and Lee, C.T. (2003). Evaluation of the potential of Hyperion for fire danger assessment by comparison to the Airborne Visible/Infrared Imaging Spectrometer. *IEEE Transactions on Geoscience and Remote Sensing, 41*: 1297-1310.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., and Green, R.O. (1998). Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment, 65*: 267-279.

Rozenstein, O., and Karnieli, A. (2011). Comparison of methods for land-use classification incorporating remote sensing and GIS inputs. *Applied Geography, 31*: 533-544.

Running, S.W., Loveland, T.R., and Pierce, L.L. (1994). A vegetation classification logic based on remote sensing for use in global biogeochemical models. *AMBIO, 23*: 77-77.

Schmid, T., Rico, C., Rodríguez-Rastrero, M., José Sierra, M., Javier Díaz-Puente, F., Pelayo, M., and Millán, R. (2013). Monitoring of the mercury mining site Almadén implementing remote sensing technologies. *Environmental Research, 125*: 92-102.

Schott, J.R. (2007). *Remote Sensing: The Image Chain Approach*. (2 ed.). Oxford University Press. 688 pages.

Schriever, J.R., and Congalton, R.G. (1995). Evaluating seasonal variability as an aid to cover-type mapping from Landsat Thematic Mapper data in the Northeast. *Photogrammetric Engineering and Remote Sensing, 61*: 321-327.

Sellers, P.J., Meeson, B.W., Hall, F.G., Asrar, G., Murphy, R.E., Schiffer, R.A., Bretherton, F.P., Dickinson, R.E., Ellingson, R.G., Field, C.B., Huemmrich, K.F., Justice, C.O., Melack, J.M., Roulet, N.T., Schimel, D.S., and Try, P.D. (1995). Remote sensing of the land surface for studies of global change: Models — algorithms — experiments. *Remote Sensing of Environment, 51*: 3-26.

Selley, H.E. (1949). Air Photography and its Application to Forestry. Department of Mines and Resources. Mines, Forest and Scientific Services Branch. Forest Air Survey Publication No. 6. 11 pages.

Serrano, L., Ustin, S.L., Roberts, D.A., Gamon, J.A., and Peñuelas, J. (2000). Deriving Water Content of Chaparral Vegetation from AVIRIS Data. *Remote Sensing of Environment, 74*: 570-581.

Shen, S. S., Badhwar, G. D., and Carnes, J. G. (1985), Separability of boreal forest species in the Lake Jennette area, Minnesota. *Photogrammetric Engineering & Remote Sensing*. 51: 1775–1783.

Solberg, S., Næsset, E., Lange, H., and Bollandsås, O.M. (2004). Remote sensing of forest health. *Natcan International Conference on Laser-Scanners for Forest and Landscape Assessment*. Freiburg, Germany. Vol. XXXVI-8/W2, pp.161-166.

Sonnentag, O., Chen, J., Roberts, D., Talbot, J., Halligan, K., and Govind, A. (2007). Mapping tree and shrub leaf area indices in an ombrotrophic peatland through multiple endmember spectral unmixing. *Remote Sensing of Environment, 109*: 342-360.

Sousa, C.D. (2001). Contaminated sites: The Canadian situation in an international context. *Journal of Environmental Management, 62*: 131-154.

Sullivan, A.A., McGaughey, R.J., Andersen, H.-E., and Schiess, P. (2009). Object-Oriented Classification of Forest Structure from Light Detection and Ranging Data for Stand Mapping. *Western Journal of Applied Forestry, 24*: 198-204.

Sun, L., Neville, R., Staenz, K., and White, H.P. (2008). Automatic destriping of Hyperion imagery based on spectral moment matching. *Canadian Journal of Remote Sensing, 34*: S68-S81.

Swedish Environmental Protection Agency (1977), Environmental Protection Act (p. 36).

Teillet, P. (1986). Image correction for radiometric effects in remote sensing. *International Journal of Remote Sensing, 7*: 1637-1651.

Thorp, K.R., French, A.N., and Rango, A. (2013). Effect of image spatial and spectral characteristics on mapping semi-arid rangeland vegetation using multiple endmember spectral mixture analysis (MESMA). *Remote Sensing of Environment, 132*: 120-130.

Tømmervik, H., Johansen, B., and Pedersen, J. (1995). Monitoring the effects of air pollution on terrestrial ecosystems in Varanger (Norway) and Nikel-Pechenga (Russia) using remote sensing. *Science of the Total Environment, 160*: 753-767.

Townsend, P.A., Helmers, D.P., Kingdon, C.C., McNeil, B.E., de Beurs, K.M., and Eshleman, K.N. (2009). Changes in the extent of surface mining and reclamation in the Central Appalachians detected using a 1976–2006 Landsat time series. *Remote Sensing of Environment, 113*: 62-72.

Institut national de la recherche agronomique. (2014) VAlidation of Land European Remote sensing Instruments (VALERI). Retrieved on February 3, 2014 from <http://w3.avignon.inra.fr/valeri/>.

van der Sluijs, J. (2012). Characterizing tree specie in the Northwest Territories using Sepctral Mixute Analysis and Mult-Temporal Satellite Imagery. (Master's Thesis). Retrieved from https://www.uleth.ca/dspace/bitstream/handle/10133/3497/VANDERSLUIJS_JURJEN_MSC_2014.pdf?sequence=1

Van Leeuwen, W.J. (2008). Monitoring the effects of forest restoration treatments on post-fire vegetation recovery with MODIS multitemporal data. *Sensors, 8*: 2017-2042.

Vaughn, N.R., Moskal, L.M., and Turnblom, E.C. (2012). Tree species detection accuracies using discrete point lidar and airborne waveform lidar. *Remote sensing, 4*: 377-403.

Wang, Q., and Li, P. (2011). Identification of robust hyperspectral indices on forest leaf water content using PROSPECT simulated dataset and field reflectance measurements. *Hydrological Processes, 8:* 1230-1241.

Weiss, M., Baret, F. (2010). CAN-EYE V 6. 1 User Manual. Institut national de la recherche agronomique (INRA). 48 pages.

Winter, M.E. (1999). N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data. In, *Storage and Retrieval for Image and Video Databases.*

Wulder, M. (1998). Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. *Progress in Physical Geography, 22*: 449-476.

Wulder, M.A., Bater, C.W., Coops, N.C., Hilker, T., and White, J.C. (2008a). The role of LiDAR in sustainable forest management. *The Forestry Chronicle, 84*: 807-826.

Wulder, M.A., Dymond, C.C., White, J.C., Leckie, D.G., and Carroll, A.L. (2006). Surveying mountain pine beetle damage of forests: A review of remote sensing opportunities. *Forest Ecology and Management, 221*: 27-41.

Wulder, M.A., White, J.C., Cranny, M., Hall, R.J., Luther, J.E., Beaudoin, A., Goodenough, D.G., and Dechka, J.A. (2008b). Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Canadian Journal of Remote Sensing, 34*: 549-562.

Xu, D., Guo, X., Li, Z., Yang, X., and Yin, H. (2014). Measuring the dead component of mixed grassland with Landsat imagery. *Remote Sensing of Environment, 142*: 33-43.

Yokoya, N., Miyamura, N., and Iwasaki, A. (2010). Preprocessing of hyperspectral imagery with consideration of smile and keystone properties. *SPIE Asia-Pacific Remote Sensing*. International Society for Optics and Photonics. (pp. 78570B-78570B-78579).

Zarco-Tejada, P.J., Miller, J.R., Noland, T.L., Mohammed, G.H., and Sampson, P.H. (2001). Scaling-up and model inversion methods with narrowband optical indices for chlorophyll content estimation in closed forest canopies with hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing, 39*: 1491-1507.

Zhu, X., Liu, D., and Chen, J. (2012). A new geostatistical approach for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment, 124*: 49-60.

Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., and Gerard, P.D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment, 87*: 171-182.

Appendix 1

Site Plot	Number of Trees	Field Measured Fractional Cover (%)	Majority Field Measured Tree Species	LiDAR Measured Fractional Cover (%)	LiDAR Measured Height (m)	Species From CHRIS SVM	Species From CHIRS LiDAR SVM	Species From CHRIS MESMA	Species From CHRIS LiDAR MESMA	License Number	Surface Location	Well Construction Date	Reclamation Certification Date	Time Between Construction and Reclamation	Age of plot after reclamation to August 2012	Reclamation Certification Number
F10025689	92	43.55	Pb	42.75	5.50	N/A	N/A	N/A	N/A	25689	4-04-066-02-01	1964-03-03	1967-02-13	2 years 11 months	45 years 5 months	55-377
F10027735	32	0.74	Pb	8.67	1.67	Aw	Aw	Unknown	Pj	27735	4-04-065-03-14	1965-02-20	1967-11-15	2 years 8 months	44 years 8 months	55-816
F10087962	0	0.00	No Trees Present	0.00	0.40	Aw	Aw	Aw	Aw	87962	4-04-064-32-10	1981-02-07	1985-12-31	4 years 10 months	26 years 7 months	
F10088251	9	20.04	Aw	2.75	1.25	Pb	Pb	Aw	Pb	88251	4-04-064-28-15	1981-02-27	1985-12-31	4 years 10 months	26 years 7 months	
F10088252	16	27.14	Pb	43.50	7.75	Aw	Aw	Aw	Aw	88252	4-04-064-28-02	1981-03-01	1985-12-31	4 years 9 months	26 years 7 months	
F10135076	34	0.28	Pb	5.60	2.60	Aw	Aw	Pj	Sw	135076	4-04-064-24-14	1988-07-26	1993-02-11	4 years 6 months	19 years 5 months	NER 93 0053
F10159571	57	9.70	Lt	13.20	3.20	N/A	N/A	N/A	N/A	159571	4-04-066-14-04	1994-01-29	2007-01-26	12 years 11 months	5 years 6 months	NE1-07-30660
F10176511	21	0.00	Sb	0.00	0.00	N/A	N/A	N/A	N/A	176511	4-04-066-10-03	1996-02-13	1999-05-17	3 years 3 months	13 years 2 months	NEB 99 2E017
F10181679	5	0.00	Lt	0.00	0.00	N/A	N/A	N/A	N/A	181679	4-04-066-03-15	1996-02-09	2012-02-13	16 years	5 months	
F10182632	6	8.07	Pj	4.33	1.67	N/A	N/A	N/A	N/A	182632	4-04-066-10-11	1996-01-14	1999-05-17	3 years 4 months	13 years 2 months	NEB 99 2E032
F10191968	0	0.00	No Trees Present	0.00	0.00	Aw	Aw	Aw	Pj	191968	4-04-065-05-11	1997-03-05	1999-05-17	2 years 2 months	13 years 2 months	NEB 99 2E040
F10195902	0	0.00	No Trees Present	0.00	0.00	N/A	N/A	N/A	N/A	195902	4-04-066-27-13	1997-03-16	1999-05-17	2 years 2 months	13 years 2 months	NEB 99 2E022
F10196012	0	0.00	No Trees Present	0.75	5.50	N/A	N/A	N/A	N/A	196012	4-04-066-26-13	1997-03-21	1999-05-17	2 years 1 month	13 years 2 months	NEB 99 2E033
F20051406	7	41.94	Pb	16.75	3.00	Aw	Aw	Pj	Pb	51406	4-04-064-13-01	1974-10-26	1985-12-31	11 years 2 months	26 years 7 months	
F20069598	0	0.00	No Trees Present	0.00	0.00	Aw	Aw	Aw	Unknown	69598	4-04-064-20-13	1978-07-25	2009-01-23	30 years 5 months	3 years 6 months	NE1-09-80307
F20087929	2	0.00	Pb	0.75	0.75	Aw	Aw	Lt	Pj	87929	4-04-064-21-10	1981-01-30	2007-01-19	25 years 11 months	5 years 6 months	NE1-07-02382
F20087960	0	47.49	No Trees Present	46.67	5.67	Pb	Sb	Aw	Sb	87960	4-04-064-20-02	1981-01-23	1985-12-31	4 years 11 months	26 years 7 months	
F20088253	0	0.00	No Trees Present	0.00	0.00	Unknown	Unknown	Unknown	Unknown	88253	4-04-064-32-06	1981-02-15	1985-12-31	4 years 10 months	26 years 7 months	
F20104481	0	0.00	No Trees Present	0.00	0.00	Aw	Aw	Pb	Aw	104481	4-04-064-28-03	1984-02-05	1986-10-23	2 years 8 months	25 years 9 months	NER 86 0062
F20105627	25	30.55	Pb	0.40	0.80	Pb	Pb	Unknown	Pb	105627	4-04-064-28-07	1984-02-18	2008-10-14	24 years 7 months	3 years 9 months	NE1-08-40100
F20112558	0	0.00	No Trees Present	3.75	1.75	Unknown	Unknown	Unknown	Unknown	112558	4-04-064-31-10	1985-01-25	1986-10-23	1 year 8 months	25 years 9 months	NER 86 0063
F20150494	0	0.00	No Trees Present	0.00	1.00	N/A	N/A	N/A	N/A	150494	4-04-066-26-11	1992-01-07	2008-10-21	16 years 9 months	3 years 9 months	NE1-08-10885
F20367845	0	0.00	No Trees Present	3.00	5.67	Aw	Aw	Unknown	Aw	367845	4-04-064-31-08	2007-01-09	2011-03-22	4 years 2 months	1 year 4 months	
F20368906	0	0.00	No Trees Present	0.00	0.00	Aw	Aw	Unknown	Aw	368906	4-04-064-31-01	2007-02-03	2011-03-22	4 years 1 month	1 year 4 months	
RF10025689	22	37.89	Sw	67.75	15.50	N/A	N/A	N/A	N/A							
RF10027735	21	61.63	Aw	70.60	18.40	Pb	Pb	Aw	Aw							
RF10087962	39	66.30	Aw	80.33	17.67	Aw	Aw	Aw	Aw							
RF10088251	27	53.00	Aw	67.50	16.50	Aw	Aw	Aw	Aw							
RF10088252	13	68.87	Aw	73.00	16.75	Pb	Pb	Aw	Lt							
RF10135076	23	66.87	Aw	83.50	18.50	Pj	Lt	Aw	Pb							
RF10159571	26	43.08	Sb	62.00	12.50	N/A	N/A	N/A	N/A							
RF10176511	38	34.25	Sb	72.00	13.67	N/A	N/A	N/A	N/A							
RF10181679	36	4.65	Lt	11.50	3.00	N/A	N/A	N/A	N/A							
RF10182632	12	28.62	Aw	26.75	12.75	N/A	N/A	N/A	N/A							
RF10191968	51	68.47	Lt	79.40	16.40	Sw	Sb	Pj	Sb							
RF10195902	20	72.42	Aw	64.67	23.00	N/A	N/A	N/A	N/A							
RF10196012	50	63.75	Aw	81.40	22.20	N/A	N/A	N/A	N/A							
RF20051406	24	71.12	Pb/Aw	75.25	18.25	Aw	Aw	Pb	Sw							
RF20069598	21	86.94	Aw	88.33	20.00	Aw	Aw	Unknown	Aw							
RF20087929	18	57.48	Pb	65.60	16.20	Sw	Sw	Sw	Sw							
RF20087960	5	61.41	Pb	72.25	25.00	Aw	Pb	Aw	Pj							
RF20088253	12	66.55	Aw	81.25	18.25	Aw	Pb	Aw	Pb							
RF20104481	20	82.41	Aw	82.00	19.67	Aw	Aw	Pb	Aw							
RF20105627	25	53.55	Pb	77.00	25.33	Aw	Aw	Unknown	Sw							
RF20112558	22	67.10	Aw	72.14	18.57	Aw	Aw	Pb	Pb							
RF20150494	37	25.31	Sw	53.33	10.67	N/A	N/A	N/A	N/A							
RF20367845	18	77.58	Aw	76.50	18.00	Aw	Aw	Aw	Aw							
RF20368906	12	56.05	Pb	82.50	17.75	Aw	Aw	Pb	Aw							

Reference Sites do not have any oil and gas activity associated with them.