

REMOTE SENSING FOR DETECTING AND MAPPING FLOWERING RUSH: A CASE
STUDY IN THE OTTAWA NATIONAL WILDLIFE REFUGE (ONWR), OHIO

Arisca Droog

A Thesis

Submitted to the Graduate College of Bowling Green
State University in partial fulfillment of
the requirements for the degree of

MASTER OF SCIENCE

December 2012

Committee:

Peter Gorsevski, Advisor

Helen Michaels

Enrique Gomezdelcampo

© 2012

Arisca Droog

All Rights Reserved

ABSTRACT

Peter Gorsevski, Advisor

Predicting and mapping invasive wetland plant species is an important process for future management decisions and strategies. Controlling and mapping such plant species requires robust methods that are applicable at different ecological scales to map and monitor their spread. In particular, this study tested the feasibility of classification tree analysis (CTA) by using a high resolution Applanix 439 Digital Sensor System (DSS) aerial imagery (< 20 cm) and linear spectral unmixing (LSU) analysis by using Landsat Thematic Mapper (TM) data to produce different distribution maps of invasive flowering rush (*Butomus umbellatus* L.) potential in the Ottawa National Wildlife Refuge (ONWR) wetlands, in Northwest Ohio. The classification accuracy from CTA maps derived from different splitting rules was evaluated by kappa statistics. The overall accuracy within the different runs varied between 35 to 56 % while the “Gini” splitting rule had the best performance. The endmembers from the best CTA performing map were utilized by the LSU method for estimating sub-pixel endmember fractions at a broader geographical scale. The results derived from the aerial imagery were slightly better than those from the Landsat imagery, as the goodness of fit between the flowering rush fraction map and the data measured in the field was lower. This study was intended to demonstrate the potential for flowering rush mapping over larger area using knowledge developed from smaller geographical scale using high resolution imagery. Results indicate that both methods show promising results for the prediction of flowering rush, but additional research

that encompass different field data collection techniques, datasets of imagery and modeling methods need to be explored.

I dedicate this to my mother MarieAnne, my father Bob, and to my brothers Bob Jr, Cilvani and Dilano. Thank you for all of your love, support, and encouragement.

ACKNOWLEDGMENTS

I first would like to thank my advisor Dr. Peter Gorsevski for his guidance, support, knowledge, and tolerance as he spend many hours helping me with this thesis. Also, a special thanks to my committee members Dr. Helen Michaels and Dr. Enrique Gomezdelcampo as they both provided guidance, advice, and continuous support of this thesis research. Thank you, Stephanie Kuck, Mike Plenzer, and Jacob Meier for spending many days in the marsh helping me collect data when temperatures were close to a hundred degrees.

I would like to thank Ron and Kathy Huffman, as well as the Ottawa National Wildlife Refuge who supported this project. Also, a big thank you to the BGSU Geology Department, both faculty and classmates, who introduced me to a whole new field of study and taught me how to enjoy rocks!

Lastly, I would like to thank my family and friends both overseas and here in the United States for their continuous love and support these last 6 ½ years, as I completed my undergraduate degree and now graduate degree far away from home. I would not have been able to complete this project without any of you. Again, thank you all!!

TABLE OF CONTENTS

	Page
INTRODUCTION	1
CHAPTER I. MODELING THEORY	6
Classification Tree Analysis	6
Linear Spectral Unmixing.....	9
CHAPTER II. MATERIALS AND METHODS.....	11
Study Area	11
Data Acquisition	11
Data Analysis and Development.....	14
CHAPTER III. RESULTS AND DISCUSSION.....	16
Aerial Imagery Results	16
Landsat TM Results	18
CHAPTER IV. CONCLUSION	22
REFERENCES	24
APPENDIX A. FIGURES	32
APPENDIX B. TABLES	40

LIST OF FIGURES

Figure		Page
1	Ottawa National Wildlife Refuge (ONWR) study area	32
2	Predetermined vegetation classes and the percent cover	33
3	Methodology used in current study.....	34
4	Classification tree Analysis (CTA)	35
5	Prediction map derived from aerial image using CTA	36
6	Percentage of total pixels for each predicted endmember	37
7	Endmember separability plot using Landsat TM bands	38
8	Linear Spectral Unmixing (LSU) fraction maps.....	39

LIST OF TABLES

Table		Page
1	Comparative accuracies for classification of the aerial imagery	40
2	Producer's Accuracy, Consumer's Accuracy, Error of Emission, and Error of Commission	41
3	Transformed divergence separability between three endmembers	42
4	Percentage of total pixels for each endmember	43
5	Validation of LSU fraction maps	44

INTRODUCTION

Flowering rush (*Butomus umbellatus* L.) is an invasive species introduced from temperate Europe and Asia that is becoming a major concern across North America. Climatic change, international trade, land conversion and grazing are all anthropogenic related disturbances that are likely to promote the spread of invasive plant species, thus creating a major challenge for land managers and ecologists (Underwood et al., 2003). Invasive species like the flowering rush have the ability to spread quickly and thus threaten the persistence of native flora and fauna, natural biodiversity and recreational and commercial activities (Joshi et al., 2004; Antonio et al., 2004; Brown and Eckert 2005). An example of this is the threat to native littoral species such as *Zizania aquatic*, also known as wild rice, an economically important plant in the Great Lakes. The invasion that can occur in wild rice habitat often requires the use of herbicides and can have adverse impact on the entire ecosystem (Brown and Eckert 2005; Johnson et al., 2008). Some estimates by the United States Department of Agriculture (USDA) suggest an economic loss of approximately \$137 billion dollar per year on the control of invasive species in the country; approximately \$35 billion annual cost for just invasive plants (Ustin et al., 2002; Windle et al., 2008). Costs associated with the control of wetland weeds alone are to be estimated around \$145 million dollars a year (Lavoie et al., 2003).

Flowering rush is an aquatic plant which comes from a monotypic family with emergent and fully submerged phenotypes, found in littoral lake habitats, rivers and ditch edges and other shallow wetland environments (Eckert et al. 2000; Klüber and Eckert 2005). The plant is easily identified by its whitish, pink flowering umbel and its spirally twisted leaf tips. It can be recognized by its three pink colored sepals and three large petals on a long pedicle arising from

the flowering stalk and can grow to about five to six feet in height (Johnson et al., 2008; Rice and Dupuis 2009). The leaves have distinctive triangular cross-sections where rigid vertical leaves are characteristic for the emerged phenotype and lax leaves that wave in the current are characteristic for the submerged phenotype. Flowering rush can be found in shallower areas such as along shorelines but can also be found in deeper water of lakes and rivers (Johnson et al., 2008). For example, Rice et al., (2010) reported that in clear oligotrophic environment such as the Flathead Lake in Montana flowering rush colonies exist in deep waters up to 20 foot. The sexual seed reproduction of the plant is combined with clonal reproduction via vegetative bulbils (Kliber and Eckert 2005; Rice and Dupuis 2009).

Literature suggest that flowering rush was first introduced into North America by export as horticultural plants from the Netherlands (Kliber and Eckert 2005; Rice and Dupuis 2009). Flowering rush was first recorded on the Saint Lawrence River in Eastern Canada in 1897 (Kliber and Eckert 2005) and quickly infested the northern part of the United States. Flowering rush then quickly spread into Eastern Lake Ontario and Lake Champlain over a time period of 30 years (Brown and Eckert 2005). Today flowering rush invades western states and provinces along the Canada/USA border such as Idaho, Oregon and Washington (Brown and Eckert 2005; Johnson et al., 2008; Rice and Dupuis 2009). Early detection and rapid assessments of the distribution and abundance of this species is crucial as it could indicate where these species may occur in the following spring (Santos et al., 2011).

However, within this broad geographical extent the invasion of flowering rush varies among different environments and factors, which drive the invasion across the country and present significant problems for managers and decision makers. Such problems require accurate mapping and monitoring for locating and controlling flowering rush and other invasive

communities over this large area. Remote sensing techniques for mapping invasive species have been used successfully at different spatial and temporal scales (Ustin et al., 2002; Turner et al., 2003; Joshi et al., 2004; Lawrence et al., 2005; Lass et al., 2005; Adam et al., 2010; He et al., 2011). Examples of successful mapping application by hyperspectral remote sensing sensors (i.e., Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), PROBE-1, Digital Compact Airborne Spectrographic Imager (CASI), and Hyperion include detection of the following invasive species: jubata grass (Underwood et al., 2003), Brazilian pepper (Lass and Prather 2004), and spotted knapweed and babysbreath (Lass et al., 2005). Multispectral remote sensing has also been a valuable tool for species detection due to its ability to record the intensity of reflected light at specific spectral wavelengths in the electromagnetic spectrum. Multispectral techniques simplify the complexity of the spectral reflectance curves for targeted plant species and communities (Turner et al., 2003; Joshi et al., 2004; Lass et al., 2005; Lawrence et al., 2006). Multispectral data from commercial satellites used for detection of invasive species include platforms such as Landsat TM/ETM+, ASTER, SPOT, IKONOS and Quickbird (Chong et al., 2001; Fuller 2005; Everitt et al., 2005; Bradley and Mustard 2006; D'iorio et al., 2007; Ghioca-Robrecht et al., 2008). Applications of high-spatial resolution imagery from IKONOS and Quickbird have been used to detect giant salvinia (Everitt et al., 2007) and malaleuca trees (*Malaleuca quinquenervia*) (Fuller 2005). Also, Landsat TM/ETM+ and Landsat MSS imagery has been used for mapping the spatial extents of the nonnative cheatgrass (Bradley and Mustard 2006), leafy spurge (Mladinich et al., 2006), and for the detection of spatial and temporal changes in semi-arid wetlands (Schmid et al., 2004). ASTER and SPOT imagery have been used for mapping invasive mangrove distributions (D'iorio et al., 2007), forest degradation (Souza et

al., 2003), and for classifying and mapping both native and nonnative rangeland plant communities (Clark et al., 2001).

However, the success of those satellite-based methods requires adequate population density of invasion for extracting accurate spectral detail (Turner et al., 2003; Lass et al., 2005; Evangelista et al., 2009) which is not the case with the flowering rush. On the other hand, aerial imagery can yield more detail than satellite imagery but they are often impractical for large area mapping and monitoring (Harvey and Hill 2001; Underwood et al., 2003; Everitt et al., 2004; Mladinich et al., 2006; Morgan et al., 2010; Grapentine and Kowalski 2010; He et al., 2010). Both satellite and aerial imagery often utilize Global Positioning System (GPS) for mapping the target area, ground truthing, and building correlations between the observed flowering rush and the imagery (Everitt et al., 2004; Mladinich et al., 2006; D'iorio et al., 2007). Although such methods provide high accuracies, they are still time consuming, labor intensive, and are only applicable for smaller management areas or collection of datasets for validation purposes (Lass et al., 2005; Lawrence et al., 2006; Mladinich et al., 2006; He et al., 2010; Tuanmu et al., 2010).

Many different mapping techniques have been explored for detecting invasive species using aerial imagery or Landsat TM such as neural network (Joshi et al., 2006; Wang et al., 2009), multiple endmember spectral mixture analysis (Wu 2004; Kärđi 2007), maximum likelihood (Ustin et al., 2002; Yang et al., 2011), Support Vector Machine (SVM) (Kelly et al., 2007; Yang et al., 2011), Spectral Angle Mapper (SAM) (Yang et al., 2011), and classification tree analysis (CTA) (Yuan et al., 2005; Andrew and Ustin 2008). CTA is a modern statistical technique that is ideal for the analysis of complex ecological data (De'ath and Fabricius 2000). It is a rule based classification of remotely sensed imagery and is useful in analyzing large data sets with complex structure without making distributional assumptions (Zambon et al., 2006;

Wright and Gallant 2007). The classification tree provides a variation of single response variables by repeatedly splitting the data into homogenous groups. Results however, are simple and easy to interpret. Also, more recent techniques such as linear spectral unmixing (LSU) (Peterson and Stow 2003; Miao 2006; Zhang et al., 2011) allows the quantification of sub-pixel abundance of species coverage from Landsat TM satellite imagery. The sub-pixel analyses are intended to provide the relative abundance of species from mixed pixels that result from a systematic combination of component spectra present especially when dealing with medium to coarse resolution satellite images. However, there are many other site specific challenges such as data availability, availability of time series, spatial or spectral data properties, cost constraints, and population density that complicate the mapping and the monitoring of invasive species spread using specific mapping procedures.

This thesis presents a suitable procedure to use high-spatial resolution from aerial imagery and moderate resolution from Landsat TM imagery for mapping flowering rush. In particular, the approach presented here included testing the applicability of CTA and LSU for the prediction of flowering rush. Specifically, a high resolution imagery of the ONWR was acquired to extract endmembers through CTA and obtain representative pixels from homogenous land covers. These results were then used for broad scale prediction using LSU. The proposed approach is demonstrated using a case study in the Ottawa National Wildlife Refuge (ONWR) wetlands, in Northwest Ohio. The CTA and LSU modeling approaches are presented in Section 2. Section 3 discusses the study area and the methodology applied in this study, while Section 4 discusses the results and the conclusions from this approach.

1. MODELING THEORY

1.1 Classification Tree Analysis (CTA)

The use of CTA or also known as classification and regression tree analysis (CART) in remote sensing and digital processing applications has been rapidly increasing (DeFries and Chan 2000; Lawrence et al., 2004; Zambon et al., 2006). CTA is a type of machine learning algorithm that has been successfully used for classification and regression problems of multispectral and hyperspectral imagery (Lawrence et al., 2004; Xu et al., 2005; Laliberte et al., 2007). The results of CTA are often in the form of an easily interpretable dichotomous tree that can be used as classification rules either by themselves or combined with expert knowledge (Lawrence et al., 2004). CTA can handle continuous and categorical information equally well and has resulted in higher accuracies than other methods such as maximum likelihood classifiers and linear discriminant function classifiers (Lawrence et al., 2004; Zambon et al., 2006; Laliberte et al., 2007).

A decision tree is composed of a root node which contains all the data, a set of internal nodes, also known as splits, and a set of terminal nodes or leaves (Xu et al., 2005). In use, the decision tree splits a dataset into increasingly homogenous subsets until terminal nodes are determined (Laliberte et al., 2007). CTA typically operates based on binary decisions that are made at each node. These binary splitting measures are applied to explanatory variables such as spectral responses (Lawrence et al., 2004; Xu et al., 2005). Furthermore, there are splitting rules associated with each split. There are several splitting rules that are widely available in current software implementations, such as Idrisi Taiga (Eastman 2006), of CTA for creating decision trees. Common splitting rules include: entropy, gain ratio, and Gini (Zambon et al., 2006; Eastman 2006). The method used is an important aspect of image classification as it controls the

pace of improvement in predictive accuracy (Gao 2009). First, entropy is a measure of impurity and therefore selects the attribute with the minimum entropy to divide the data set (Huang and Jensen 1997; Jin and Agrawal 2003). Entropy is becoming one of the most popular impurity functions when performing tree classification (Jin and Agrawal 2003). Entropy measures the degree of disorder or heterogeneity in an image and therefore provides a measurement of the impurity in the data set (Jin and Agrawal 2003; Laliberte et al., 2007). It is defined as:

$$Entropy(t) = - \sum_i p_i \log_2 p_i \quad \text{Eq. (1)}$$

where p_i is the relative frequency of class label i at node t (Jin and Agrawal 2003). The main aim of the entropy algorithm is to split and group the data by minimizing the within group diversity where the rule is associated with expected value of the minimized negative log-likelihood of a given split result.

One of the drawbacks in the entropy algorithm is oversplitting because its preference is to use attributes with large number of branches (i.e., unique values determine classes). To overcome this potential bias the gain ratio algorithm is applied which intends to normalize the process. The gain ratio measures the reduction in entropy in the data produced by a split; the subdivision of the data that maximizes the reduction in entropy of the descendant node is how each final node is selected (DeFries and Chan 2000). The gain ratio is calculated into two stages. First, the information gain is calculated for all attributes as:

$$Gain(t) = entropy(t) - \sum_{i=1}^n \frac{|t_i|}{|t|} \times entropy(t_i) \quad \text{Eq. (2)}$$

where the information gain of a single classification t is defined as the entropy after classification t and the sum of the entropy of each subset t_i , is weighted by its proportion $|t_i| / |t|$ which is used for normalization and accounting for large number of splits. The split information is defined as:

$$Splitinfo(t) = - \sum_{i=1}^n \frac{|t_i|}{|t|} \times \log_2 \left(\frac{|t_i|}{|t|} \right) \quad \text{Eq. (3)}$$

which represents the potential information generated by dividing t into n subsets.

Because information gain is not effective for some variables, dividing the information gain by the splitting information generates the gain ratio that tests for maximization of the information gain as:

$$Gain\ ratio(X) = gain(X) / split\ info(X) \quad \text{Eq. (4)}$$

Lastly, the Gini impurity is a measure of heterogeneity where the impurity measure at node t is defined as:

$$Gini(t) = \sum_i p_i(1 - p_i) \quad \text{Eq. (5)}$$

where p_i are the relative frequency of class i at node t that represents any parent or child node at which a given split of the data is performed. The sequence of the Gini splitting algorithm is to first isolate the largest homogeneous category within the dataset before subsequent nodes are then segregated and further divisions are not possible (Laliberte et al., 2007).

1.2 Linear Spectral Unmixing (LSU)

Linear spectral unmixing analysis is based on the concept that every image pixel is a mixture of different components that is based on the principle that reflectance recorded in each pixel is a weighted average of the reflectance from all endmembers in that pixel (Kärdi 2007; Lu, et al., 2003; Pacheco and McNairn 2010). A pixel is mixed by a number of different materials present in a scene called endmembers (Heinz and Chang 2001). Endmembers are a coherent set of spectra that represent physical components on the surface but also model the spectral variability inherent to the scene (Gong et al., 1994; Elmore et al., 2000). The intent of the spectral unmixing is to estimate the fractional abundance of surface targets such as different vegetation classes at sub-pixel level. The unmixing process assumes that composite values measured by the sensor represent a linear combination of the spectra of all components within a pixel (Lu et al., 2003). Thus, if we know the spectral reflectance in each pixel it is possible to describe these mixed pixels as linear mixtures of the endmembers and estimate the best fitting combination of endmember fractions for the observed reflectance spectra (Small 2003). Unmixing of individual pixels is possible by estimating the fraction of each endmember in the composite reflectance of a pixel (Small 2001):

$$R = \sum_{i=1}^N f_i R_i \quad \text{Eq. (6)}$$

where R is the effective reflectance of the mixed pixel, R_i is the reflectance of the i^{th} material (endmember), f_i is the spatial fraction covered by the i^{th} material and N is the number of materials in the pixel.

Spectral unmixing analysis has been widely used in remote sensing for various studies. Some of those studies include using linear spectral unmixing for monitoring long term vegetation

dynamics in Mediterranean rangelands (Hostert et al., 2003), evaluating remote sensing and spectral unmixing analysis for crop residue mapping (Pacheco and McNairn 2010), and for the unmixing of soils and vegetation (Asner and Lobell 2000).

2. MATERIALS AND METHODS

2.1 Study Area

Ottawa National Wildlife Refuge (ONWR) is located along the western Lake Erie shore in Benton Township, Ottawa County in Northwest Ohio (Figure 1). The refuge consists of approximately 40.5 km² (10,000 acres) of marshes, open water, wooded and coastal wetlands, shrub lands, grasslands, cropland, and an estuary. ONWR is known for its wide variety of plant and animal species and particularly for its critical migratory bird resting area (USFWS ONWR 2011). The refuge has a rich biodiversity and significant habitat for migrating birds and has been identified by the American Bird Conservancy as an Important Bird Area (USFWS ONWR 2012). A 4 m opening in a dike along Lake Erie connects the marshes to the lake and protects the marshes from offshore waves. However, water levels are driven by Lake Erie conditions as well as precipitation events (Bowers et al., 2005; Grapentine and Kowalski 2010). Annual precipitation in Ottawa County is approximately 805 mm with February typically being the driest month and August the wettest (OSU Extension Fact Sheet 2012). Seasonal weather extremes are influenced by its close proximity to Lake Erie (Hamilton and Limbird 1982). The general area consists of poorly drained glacial lake plain deposits that are nearly flat and approximately 175 m above sea level.

2.2 Data Acquisition

The field data collection used aerial imagery from July 2010 to aid the location of the plots followed by flowering rush *in situ* data collection that occurred during the last two weeks of July 2011. Another set of aerial imagery from August 2011 that coinciding with the field data collection was used for the spatial prediction analyses and accuracy assessments of the flowering rush. Both aerial imagery sets were cloud-free and acquired by an Applanix 439 Digital Sensor

System (DSS) (Mostafa, 2003). The DSS has ground sample distance of 0.16 m and a total of 3 spectral bands (60 mm CIR) including blue/green 500-600 nm; green/red 600-720 nm; and red/NIR 850-1200 nm. The imagery was orthorectified by U.S. Fish and Wildlife Service (UFWS) using RapidOrtho software (Hogan, 2012) to the World Geodetic System 1984 (WGS 84) datum and the Universal Transverse Mercator (UTM) zone 17 coordinate system. The aerial imagery was preprocessed using a dark object subtraction (DOS) atmospheric correction method to reduce the effects of atmospheric haze (Vincent et al., 2004). DOS is where the minimum pixel values of a dark object are subtracted with an assumption that no energy is reflected from that dark object (Chavez 1988; Qi et al., 2000).

A total of seven subset areas were extracted from the aerial imagery from July 2010 for the design of the sampling plots using a stratified-random approach. Initially, the subset areas were predetermined and used by ONWR who classified these into homogenous subset areas of the following categories: flowering rush (FR), flowering rush mixed with water (H₂O), and flowering rush mixed with other vegetation (MIX). Exploratory data analysis showed that those three categories were broadly defined with significant occurrence in variation that may have been driven by previous mapping standards. Therefore, the subset areas were subjected to unsupervised iterative self-organizing data analysis (ISODATA) classification to ensure better sample representation and control of variation within each subset areas. The ISODATA technique used a minimal spectral distance algorithm to assign a total of three arbitrary classes within each subset areas. A minimum of 10 sample points were assigned to each subset area (a minimum of 3 points per class) in a stratified random pattern using ArcGIS software (ESRI 2011). The geographic coordinates of each sampling point were extracted and a handheld Trimble Pathfinder GPS (GeoExplorer XH) receiver was used to navigate to the sample points

for subsequent data collection. The data collection used a 1×1 m quadrat placed on the ground or water where within each plot the percentage of flowering rush, water, and other vegetation cover was recorded. The categories were assigned based on the flowering rush abundance being greater than 50 % or assigned by the most dominant class. Thus, 21 points were recorded as flowering rush, 44 points were recorded as flowering rush mixed with other vegetation, and 18 points were recorded as flowering rush mixed with water. Figure 2 shows the original cover data for the 3 classes that was collected during the field data collection. First, Figure 2 shows that the median is above 90 % for the mixed vegetation cover (MIX) suggesting that the population of FR in this class is low. The boxplot for the H₂O class shows that the FR population is sparse because the median percent cover is still above 80 %, therefore suggesting that water dominates in most of the points that represent this class. Finally, the percent cover for FR shows that the median percent cover from the field data collection is 75 %, and thus suggests that a majority of the points represent a FR monoculture.

In addition, a Landsat TM scene from June 6, 2011 was used to determine the feasibility of broader scale mapping of flowering rush. While the spatial resolution of aerial imagery is very useful for detailed wetland mapping, the superiority of the spectral resolution of Landsat TM allows a collection of information for large geographic areas in a timely and costly effective manner. The scene used in this analysis covered Path/Row 20/31 in the spatial reference of Universal Transverse Mercator (UTM) 17N projection using a World Geodetic System (WGS) 1984 datum. Landsat TM has a ground sample distance of 30 m and includes spectral band 1 (0.45-0.52 μm), band 2 (0.52-0.60 μm), band 3 (0.63-0.69 μm), band 4 (0.76-0.90 μm), band 5 (1.55-1.75 μm), band 6 (10.40-12.50 μm), and band 7 (2.08-2.35 μm). Landsat TM spectral bands two (green), three (red), four (near infrared), along with a Normalized Vegetation Index

(NDVI), and two ratio vegetation indices of RVI (gni), and RVI (rni) were chosen as inputs for the LSU. The first ratio vegetation index (RVI) was derived by dividing band 2 (green) by band 4 (NIR), thus $RVI (gni) = \text{Green}/\text{NIR}$. The second ratio vegetation index was derived from band 3 (red) divided by band 4 (NIR), therefore $RVI (rni) = \text{Red}/\text{NIR}$. These ratios were chosen because of their ability to separate vegetation from soil and water by combining surface reflectance at two or more wavelengths. Thus, high values will indicate healthy living vegetation because of the high reflectance in the near-infrared and low reflectance in the red regions of the spectrum. The Landsat TM data tested here does have an advantage over the high spectral resolution data, which includes its cost-free availability, its utility for mapping larger areas at a time, and its repeatability for monitoring purposes. Challenges associated with this type of data are misclassification of pixels due to land cover mixtures, and mapping abundances in heavily fragmented areas (Kumar et al., 2007).

2.3 Data Analysis and Model Development

For the development of the proposed model (Figure 3), the 2011 aerial imagery's pixel size was resampled from 0.16 m to 1 m for it to match the 1×1 m quadrat for the field work collection. The CTA model was constructed using Idrisi Taiga software (Eastman 2009) and includes independent and dependent variables. The independent variables of 3 bands from the aerial imagery and a Normalized Difference Vegetation Index (NDVI) derived from the red and the near-infrared (NIR) bands. The significance of the NDVI is in its ability to highlight relevant vegetation changes because it absorbs strongly in the red wavelengths of sunlight and reflects in the near-infrared wavelengths (Chen et al., 2004; Healey et al., 2005).

The dependent variable used the field data collection and included a total of 83 observation points categorized as FR, H₂O, and MIX classes. These three different categories are

also known as endmembers. The CTA was performed using different splitting rules while trees were pruned by removing leaves with observations less than a predetermined within class proportion of 2 %. The selection of splitting rules is the most important factor for the overall accuracy of the classification because decision trees often “over fit” the data; therefore the classification requires tree reduction or pruning for development of robust and parsimonious trees for image classification purposes (Brown de Colstoun et al., 2003; Lawrence et al., 2004; Zambon et al., 2006). The kappa statistic was used to evaluate the accuracy of the classification tree. The kappa statistic considers all cells in a matrix and thus provides a correction for the proportion of chance agreement between the known sample points and the classification tree. Values of kappa can range from -1.0 to 1.0, where -1.0 indicates perfect disagreement below chance, 0.0 indicates agreement that is equal to chance, and 1.0 indicates perfect agreement above chance.

The transformed divergence distance was then used to evaluate the spectral separability of the endmembers and to select the optimum subset of bands from the Landsat TM image. The transformed divergence is a measure of statistical distance between endmember pairs of interest and can provide information on their separability. Values of the transformed divergence distance can range from 0 to 2000 where anything less than 1700 is considered poor separation and anything greater than 1700 is considered good separation (Malik and Husain 2006). The fraction maps were generated by performing LSU on the Landsat TM data. The ground data collection acquired in the field was used to validate the results of the LSU by correlating the image fraction maps to the data assessed on the ground.

3. RESULTS AND DISCUSSION

3.1 Aerial imagery results

The accuracy results of the multiple CTA runs, where every run represents a different splitting rule, are shown in Table 1 through the kappa coefficients of agreement. The table represents comparative accuracies ranging from 0.359 to 0.567 depending on their splitting rule and pruning level. The results from the table suggest that the best classification model was predicted by the Gini algorithm where the leaves were pruned with a proportion of less than 1.3 %. The accuracy associated with this classification has a coefficient of agreement kappa (κ) of 0.567 which was calculated for the overall error matrix of each classification. A kappa of 0.567 can be interpreted as a moderate agreement between the imagery and the sample points (Eastman 2006). Accuracy of the decision tree was assessed by using a cross-tabulation and kappa statistics. A cross tabulation is a process that combines and summarizes data from two sources in the form of a contingency table (Campbell 2007). The tabulation shows the number of respondents which give a particular combination of replies, where in this study, it includes the total number of pixels corresponding to each combination of endmembers being compared. In Table 1, kappa (κ) statistics seem to increase when decreasing the pruning percentage for both Entropy and Gini, however, Gini's kappa statistics overall seem to remain higher. The Gini split type fits the data best because it finds the largest homogeneous category within the dataset and separates it from the remainder of the data (Zambon et al., 2006). Furthermore, Table 2 shows the producer's accuracy (PA) and consumer's accuracy (CA) for all three endmembers. The table shows how many pixels are correctly or incorrectly classified on the final map, as well as the error associated with this. The producer's accuracy calculates how much of the ground vegetation type was correctly classified as that species and is a measure of omission error (EO).

The consumer's accuracy, however, is a reliability measure and therefore a measure of commission error (EC). The CA indicates how likely that pixel classified on the map actually represents that category on the ground (Congalton 1991; Mladinich et al., 2006; Campbell 2007; Shao and Wu 2008). The producer accuracies for the CTA image ranged from 71.4 % (FR) to 72.7 % (MIX) with omission errors of 27.6 % for FR and 31.8 % for MIX. This indicates a moderate probability of correctly assigning these subclasses on the ground to the corresponding category on the map. The consumer's accuracy for correctly classifying the endmembers was 84.2 % (MIX), 68.2 % (FR), and 56.5 % (H₂O). Commission errors associated with these are 15.8 % for MIX, 31.8 % for FR, and 43.5 % for H₂O which suggests that those pixels classified on the map actually represent that category on the ground is moderate to high. Limitations associated with this form of accuracy assessment are that only data collection points are being used for validation purposes. Figure 4 shows the final classification tree ($\kappa = 0.567$) that was used to classify FR, H₂O, and MIX using three bands green, red, and NIR), and a vegetation index (NDVI) from the aerial imagery. The final decision tree has 11 terminal nodes. The tree illustrates that a value from NDVI less than 0.3567 would be classified as MIX and thus would need additional conditions to detect the flowering rush. Furthermore, the tree shows that NDVI values ranging from 0.255 to 0.3567 will have the best chance of detecting FR, while for the NIR this includes values from 97.5 to 130.5 (Figure 4). The map generated by the classification rules from Figure 4 is shown in Figure 5 and shows that the majority of the flowering rush can be found near ditches along roads, as well as near the edges of the sampling subset areas. Physical variability, including fluctuating water levels and management tools, can also have an effect on the locations of flowering rush. Figure 6 shows the results that were extracted from the final classification map in Figure 5. Figure 6 (a) shows the predicted locations for the three

endmembers within the subset areas. The figure shows that 28 % of the total pixels are predicted to be FR, while 30 % represents MIX, and 42 % is H₂O. Figure 6 (b) shows the predicted endmember locations for the entire map. Out of the total pixels, it appears that 16 % is predicted to be FR, 41 % is H₂O, and 43 % is MIX, respectively. Results suggest that the small percentage of FR could be influenced by the number of sampling points collected during field work and therefore making it more difficult to distinguish the flowering rush from other endmembers. In addition to including more sampling points, more specific information from the sample data is also required to acquire more accurate results. Lastly, collecting the data at a different time of the season, for example, when the flowers are in bloom versus when they are not, or adding additional endmembers to better distinguish the flowering rush from other vegetation, can also add to better accuracy and needs to be considered for future studies.

3.2 Landsat TM results

Figure 7 illustrates the scaled reflectance curve from Landsat TM band 2 (green), band 3 (red), band 4 (NIR), NDVI, RVI (rni), and RVI (gni). The x-axis represents the selected Landsat TM bands and the vegetation indices to illustrate the separability between the three endmembers. The vegetation indices were chosen because they produced the best separability. The y-axis illustrates the scaled reflectance (DN) values ranging from 0 to 256 because all layers are standardized to an unsigned 8-bit (256 range) data type. The visible spectrum associated with water have low values because no light is being reflected back, however, this is dependent upon the water depth as clearer bodies of water can have higher reflectivity. For the NIR region water absorbs the light, therefore making it darker. Values greater than 0 can indicate that the sensor is detecting light reflected from an object depending on the bands. The illustrated graph of the spectral reflectance of these objects as a function of wavelength is termed the spectral reflectance curve.

This is important in determination of the wavelength regions to give insight into the spectral characteristics of these particular endmembers (Govender et al., 2007; Xie et al., 2008). The reflectance curve in Figure 7 shows that the separability between the endmembers is different across different bands and vegetation indices. For instance, the NDVI shows that the largest separability is between the MIX and the H₂O. NDVI values for MIX and FR are significantly higher than H₂O because the spectral reflectance is based on the chlorophyll absorption of the vegetation. Furthermore, RVI (gni) and RVI (rni) show a decrease in reflectance value because, although it separates vegetation from soil and water, it still lowers the value because the green and red portions of the spectrum have a low reflectance. Both RVI's illustrate substantial separabilities among the endmembers. Table 3 quantifies the spectral separability between the three endmembers. The highest separability was associated with H₂O and MIX (2000.0), while the lowest was associated with FR and MIX (1919.6), therefore indicating that the endmembers are clearly distinct from each other. This provides an indication of the ability to numerically classify the endmembers correctly, as well as to what degree of accuracy. Average class separability was 1971.8 and therefore considered good separability between endmembers.

Figure 8 shows the percentage coverage per endmember in each pixel obtained from the LSU. The legend represents a measure of the percentage of cover (endmember) in each pixel expressed on a scale range between 0.0 and 1.0, with 0.0 indicating absence of the endmember and increasing values showing higher abundance. Overall, the distribution and abundance values show that vast majorities of high coverage percentage (1.0) for flowering rush are located outside of the sampling subset areas. This includes areas near the edges of the sampling subset areas and roads, where no sample points were collected. Therefore, in comparison to the CTA classification from the aerial imagery, where FR was also found outside of the sampling subset

areas, the LSU results do indicate some similar patterns as the CTA map. Figures 8 (a) and Table 4 show that 90 % of the study area only has a 0 – 20 % likelihood of flowering rush being present, while only 4 % of the total pixels are in the 80 – 100 % likelihood of being present. Therefore, only a small area is associated with FR monoculture. Figures 8 (b) and Table 4 however, shows that 51 % of the total MIX pixels are associated with 80 – 100 % coverage while 43 % of the MIX pixels are associated with a probability 0 – 20 %. Similar to this is Figure 8 (c) and Table 4 where the majority of the H₂O pixels (57 %) are associated with a low likelihood of being present (0 – 20 %) while 35 % is associated with a high likelihood of being present (80 – 100 %). Figure 8 (d) represents the residual map and shows the error for each pixel showing higher errors in areas further away from where the sample points were taken. Also, when comparing the areas with higher errors to the aerial image, it is evident that they correspond to areas where high quantities of trees are present, and therefore unlikely for flowering rush to be present. Therefore it can be concluded that the highest uncertainty is with the MIX because of its close spectral reflectance to the FR, and because of higher residual values associated with areas where other vegetation is present. The LSU accuracy assessment for each fraction map was conducted by refitting the original field data set (Table 5). The table shows the number of sample points for each endmember corresponding to the coverage per pixel. Table 5 shows that 10 sample points of FR are associated with a low percentage of coverage per pixel (0 – 20 %), thus indicating a disagreement between the sample points and the pixels from the fraction map. In addition, 3 sample points of FR are associated with a high coverage per pixel (80 – 100 %). Sample points that are found in areas of poor agreement with the pixels are forested areas where in fact, no flowering rush was observed. Thus, the LSU validation results suggest that a total of 3 FR points were correctly classified while 10 were misclassified. For MIX however, 21 points are

correctly classified while 20 were misclassified. This indicates that the lack of data collection points is causing some uncertainty in the final results of LSU.

Based on the information from the figures and tables, it is evident that the overall model fit is poor and needs significant improvements, however, this methodology had demonstrated the prospective for mapping flowering rush at broader geographical scale. Improving the accuracies of LSU can include collecting ground data simultaneously to the imagery, and using longer monitoring periods. Accuracies could be greatly affected by the fact that the Landsat TM image was obtained from June 6, 2011, while the aerial imagery along with the data collection was performed in late July. No other Landsat TM images were acquired due to cloud coverage and thus were unable to provide useful scenes. Field data collection was completed during the first two weeks of July when the flowering rush was in high bloom. In addition, seasonal variability also needs to be considered for additional research as it can effect the distribution of the flowering rush for the upcoming spring. Lastly, assigning endmembers in consultation with the sample data can be critical as it did help discern the affiliation between the predicted endmember classes and sample points that were used for validation of the LSU results.

4. CONCLUSION

This study evaluated the suitability of using CTA to map flowering rush at a small geographical scale (<20 cm) while LSU was used to obtain abundance fraction maps of flowering rush from broader scale imagery (30 m). This thesis used CTA to generate endmembers and obtain knowledge to extrapolate flowering rush to a broader scale using LSU. The approach used in this study provided different methodological ideas for distinguishing flowering rush at different scales but additional work is needed to fully understand the capability of the proposed approach. Both CTA and LSU provide an effective approach, further highlighting the applicability of remote sensing and vegetation indices for detecting flowering rush.

The results indicate that aerial imagery coupled with classification tree analysis demonstrated its prospective for being an effective tool for mapping flowering rush across the study area. CTA results show a high potential for accurately mapping flowering rush and will be useful as a starter point when adopting more complex methodologies. Accuracy results show that the performance of the LSU technique was equitable for the predicting and mapping of flowering rush. Challenges associated with this type of data include misclassification of pixels due to land cover mixtures, and the uncertainties in the endmember spectra measurements. Also, adequate population density is needed to obtain higher accuracy results from satellite based methods. Field data collection, however, showed that flowering rush does not ensure high density infestations and therefore could complicate the mapping process of flowering rush. By integrating additional datasets of satellite imagery (i.e. hyperspectral imagery), and including more accurate spectral measurements of scene endmembers could increase the accuracy for predicting and mapping flowering rush. Another possibility to higher accuracies could be including spectral

measurements of the pink flowers through a spectrometer as it could provide spectral uniqueness. However the focus in this research was to test an applicable method to detect and map flowering rush by using high resolution aerial imagery and multispectral imagery.

Overall, this research indicates potential for successfully detecting and mapping flowering rush by using CTA for extracting endmembers and applying these to LSU to obtain fraction maps for predicting locations of flowering rush. Although these results are encouraging, more research is required to fully understand whether the CTA model and the LSU procedure are appropriate for extraction of endmembers from coarser resolution imagery. Uncertainties with this research included the field data collection, the different dates of the imagery acquisition, fusing these data together, and lastly the scale of the imagery. Also, this research was carried out during one season only and therefore the results are preliminary.

Based on this work and the discussion on how to improve the accuracy of this model, it is likely that this approach can result in yet higher classification accuracies. This would make satellite based imagery an even more attractive mapping tool for robust management and detection of invasive species across different spatial and spectral scales.

REFERENCES

- Adam, E.; Mutanga, O.; Rugege, D. 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management* 18: 281-296.
- Andrew, M.E.; Ustin, S.L. 2008. The role of environmental context in mapping invasive plants with hyperspectral image data. *Remote Sensing of Environment* 112: 4301-4317.
- Antonio, C. M. D.; Jackson, N.E.; Horvitz, C. C.; Hedberg, R. 2004. Invasive Plants in Wildland Ecosystems: Merging the study of invasion processes with management needs. *The Ecological Society of America* 2(10): 513-21.
- Asner, G.P.; Lobell, D.B. 2000. A Biogeophysical Approach for Automated SWIR Unmixing of Soils and Vegetation. *Remote Sensing of Environment* 74: 99-112.
- Bowers, R.; Sudomir, J.C.; Kershner, M.W.; de Szalay, F.A. 2005. The effect of predation and unionid burrowing on bivalve communities in a Laurentian Great Lake Coastal wetland. *Hydrobiologia* 545: 93-102.
- Bradley, B.A.; Mustard, J.F. 2006. Characterizing the Landscape Dynamics of an Invasive Plant and Risk of Invasion Using Remote Sensing. *Ecological Applications* 16(3): 1132-1147.
- Brown, J.S.; Eckert, C.G. 2005. Evolutionary increase in sexual and clonal reproductive capacity during biological invasion in an aquatic plant *Butomus umbellatus* (Butomaceae). *American Journal of Botany* 92(3): 495-502.
- Brown de Colstoun, E.; Story, M.H.; Thompson, C.; Commisso, K.; Smith, T.G.; Irons, J.R. 2003. National Park vegetation mapping using multitemporal Landsat 7 data and a decision tree classifier. *Remote Sensing of Environment* 85: 316-327.
- Campbell, J.B. Introduction to Remote Sensing. 4th ed. New York: The Guilford Press; 2007.
- Chavez, P., 1988, An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* 24: 459-479.
- Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter. *Remote Sensing of Environment* 91: 332-344.
- Chong, G.W.; Reich, R.M.; Kalkhan, M.A.; Stohlgren, T.J. 2001. New Approaches for Sampling and Modeling Native and Exotic Plant Species Richness. *Western North American Naturalist* 61(3): 328-335.
- Clark, P.E.; Seyfried, M.S.; Harris, B. 2001. Intermountain plant community classification using Landsat TM and SPOT HRV Data. *Journal of Rangeland & Management* 54: 142-160.

- Congalton, R.G. 1991. A Review of Assessing the Accuracy of Classifications of Remote Sensed Data. *Remote Sensing of Environment* 37: 35-46.
- D'iorio, M.; Jupiter, S.D.; Cochran, S.A.; Potts, D.C. 2007. Optimizing Remote Sensing and GIS Tools for Mapping and Managing the Distribution of an Invasive Mangrove (*Rhizophora mangle*) on South Molokai, Hawaii. *Marine Geodesy* 30: 125-144.
- De'ath, G.; Fabricius, K.E. 2000. Classification and Regression Trees: A powerful yet simple technique for ecological data analysis. *Ecology* 81(11): 3178-3192.
- DeFries, R.S.; Chan, J.C. 2000. Multiple Criteria for Evaluating Machine Learning Algorithms for Land Cover Classification from Satellite Data. *Remote Sensing of Environment* 74: 503-515.
- Eastman, R. J. 2006. IDRISI Andes. Clark Labs.
- Eastman, R.J. 2009. Idrisi Taiga Guide to GIS and Image Processing. Clark labs.
- Eckert, C.; Massonnet, B.; Thomas, J. J. 2000. Variation in sexual and clonal reproduction among introduced populations of flowering rush, *Butomus umbellatus* (Butomaceae). *Canadian Journal of Botany* 78: 437-446.
- Elmore, A.J.; Mustard, J.F.; Manning, S.J.; Lobell, D.B. 2000. Quantifying Vegetation Change in Semiarid Environments: Precision and Accuracy of Spectral Mixture Analysis and the Normalized Difference Vegetation Index. *Remote Sensing of Environment* 73: 87-102.
- ESRI. 2011. ArcGIS Desktop Help: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Evangelista, P.H.; Stohlgren, T.J.; Morisette, J.T.; Kumar, S. 2009. Mapping Invasive Tamarisk (*Tamarix*): A Comparison of Single-Scene and Time Series Analyses of Remotely Sensed Data. *Remote Sensing* 1: 519-533.
- Everitt, J. H.; Yang, C.; Fletcher, R. S.; Davis, M. R. 2004. Using aerial color-infrared photography and QuickBird satellite imagery for mapping wetland vegetation. *Geocarto International* 19(4): 15-22.
- Everitt, J.H.; Yang, C.; Deloach, C.J. 2005. Remote Sensing of Giant Reed with Quickbird Satellite Imagery. *Journal of Aquatic Plant Management* 43: 81-85.
- Everitt, J.H.; Fletcher, R.S.; Elder, H.S.; Yang, C. 2007. Mapping giant salvinia with satellite imagery and image analysis. *Environmental Monitoring and Assessment* 139: 35-40.

- Fuller, D.O. 2005. Remote detection of invasive *Melaleuca* trees (*Melaleuca quinquenervisa*) in South Florida with multispectral IKONOS imagery. *International Journal of Remote Sensing* 26(5): 1057-1063.
- Gao, J. 2009. *Digital Analysis of Remotely Sensed Imagery*. The McGraw-Hill Companies. 351-387.
- Ghioca-Robrecht, D.M.; Johnston, C.A.; Tulbure, M.G. 2008. Assessing the Use of Multiseason Quickbird Imagery for Mapping Invasive Species in a Lake Erie Coastal Marsh. *Wetlands* 28(4): 1028-1039.
- Gong, P.; Miller, J. R.; Spanner, M. 1994. Forest Canopy Closure from Classification and Spectral Unmixing of Scene Components—Multisensor Evaluation of an Open Canopy. *IEEE Transactions on Geoscience and Remote Sensing* 32(5): 1067-1080.
- Govender, M.; Chetty, K.; Bulcock, H. 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA* 33(2): 145-152.
- Grapentine, J. L.; Kowalski, K. P. 2010. Georeferencing Large-Scale Aerial Photographs of a Great Lakes Coastal Wetland: a Modified Photogrammetric Method. *Wetlands* 30: 369-374.
- Hamilton, E. S.; Limbird, A. 1982. Selective Occurrence of a Arborescent Species on Soils in a Drainage Toposequence, Ottawa County, Ohio. *The Ohio Journal of Science* 82(5): 282-292.
- Harvey, K.R.; Hill, G.J.E. 2001. Vegetation mapping of a tropical freshwater swamp in the Northern Territory, Australia: a comparison of aerial photography, Landsat TM and SPOT satellite imagery. *International Journal of Remote Sensing* 22(15): 2911-2925.
- He, K.S.; Rocchini, D.; Neteler, M.; Nagendra, H. 2011. Benefits of hyperspectral remote sensing for tracking plant invasions. *Diversity and Distributions* 17: 381-392.
- He, M.; Zhao, B.; Ouyang, Z.; Yan, Y.; Li, B. 2010. Linear spectral mixture analysis of Landsat TM data for monitoring invasive exotic plants in estuarine wetlands. *International Journal of Remote Sensing* 31(16): 4319-4333.
- Healey, S.; Cohen, W.B.; Zhiqiang, Y.; Krankina, O.N. 2005. Comparison of Tasseled Cap-Based Landsat data structures for use in forest disturbance detection. *Remote Sensing of Environment* 97: 301-310.
- Heinz, D.C.; Chang, C. 2001. Fully Constrained Least Squares Linear Spectral Mixture Analysis Method for Material Quantification in Hyperspectral Imagery. *Transactions on Geoscience and Remote Sensing* 39(3): 529-545.

- Hogan, B. 2012. High Altitude Tactical Mapping with Applanix DSS. Applanix, a Trimble Company.
http://calval.cr.usgs.gov/JACIE_files/JACIE08/2008_JACIE_DVD/Wednesday_AM/Hogan_DSS.pdf
- Hostert, P.; Röder, A.; Hill, J. 2003. Coupling spectral unmixing and trend analysis for monitoring of long-term vegetation dynamics in Mediterranean rangelands. *Remote Sensing of Environment* 87: 183-197.
- Huang, X.; Jensen, J.R. 1997. A Machine-Learning Approach to Automated Knowledge-Base Building for Remote Sensing Image Analysis with GIS Data. *Photogrammetric Engineering & Remote Sensing* 63(10): 1185-1194.
- Jin, R.; Agrawal, G. 2003. Efficient Decision Tree Construction on Streaming Data. 9th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD): 1-6.
- Johnson, M.; Rice, P.M.; Dupuis, V.; Ball, S. 2008 Addressing the Invasive Aquatic Flowering Rush (*Butomus Umbellatus*) in the headwaters of the Columbia River System-a multi-partner, interdisciplinary project. *Weeds Across Borders*: 76-84.
- Joshi, C.; deleeuw, J.; van Duren, I.C. 2004. Remote Sensing and GIS applications for mapping and spatial modeling of invasive species. *Proceedings of the XXth ISPRS Congress: Geo-Imagery bridging continents* 35: 669-677.
- Joshi, C.; Deleeuw, J.; Vanandel, J.; Skidmore, A.; Lekhak, H.; Vanduren, I.; Norbu, N. 2006. Indirect remote sensing of a cryptic forest understory invasive species. *Forest Ecology and Management* 225: 245-256.
- Kärđi, T. 2007. Remote sensing of urban areas: linear spectral unmixing of Landsat Thematic Mapper images acquired over Tartu (Estonia). *Proceedings of the Estonian Academy of Sciences, Biology and Ecology* 56(1): 19-32.
- Kelly, M.; Guo, Q.; Liu, D.; Shaari, D. 2007. Modeling the risk for a new invasive forest disease in the United States: An evaluation of five environmental niche models. *Computers, Environment and Urban Systems* 31: 689-710.
- Kliber, A.; Eckert, C.G. 2005. Interaction between founder effect and selection during biological invasion in an aquatic plant. *Evolution* 59(9): 1900-13.
- Kumar, U.; Kerle, N.; Ramachandra, T.V. 2007. Constrained Linear Spectral Unmixing Technique for Regional Land Cover Mapping Using MODIS Data. *International Joint Conferences on Computer, Information, and System Sciences, and Engineering (CIS²E 07)*: 1-8.
- Laliberte, A.S.; Fredrickson, E.L.; Rango, A. 2007. Combining Decision Trees with Hierarchical Object-oriented Image Analysis for Mapping Arid Rangelands. *Photogrammetric Engineering & Remote Sensing* 73(2): 197-207.

- Lass, L. W.; Prather, T.S. 2004. Detecting the Locations of Brazilian Pepper Trees in the Everglades with a Hyperspectral Sensor. *Weed Technology* 18: 437-442.
- Lass, L.W.; Prather, T.S.; Glenn, N.F.; Weber, K.T.; Mundt, J.T.; Pettingill, J. 2005. A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (*Centaurea maculosa*) and babysbreath (*Gypsophila paniculata*) with a hyperspectral sensor. *Weed Science* 53(2): 242-251.
- Lavoie, C.; Jean, M.; Delisle, F.; L  tourneau, G. 2003. Exotic plants species of the St. Lawrence River wetlands: a spatial and historical analysis. *Journal of Biogeography* 30: 537-549.
- Lawrence, R.; Bunn, A.; Powell, S.; Zambon, M. 2004. Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. *Remote Sensing of Environment* 90: 331-336.
- Lawrence, R.L.; Wood, S.D.; Sheley, R. L. 2006. Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment* 100: 356-362.
- Lu, D.; Moran, E.; Batistella, M. 2003. Linear mixture model applied to Amazonian vegetation classification. *Remote Sensing of Environment* 87: 456-469.
- Malik, R.N.; Husain, S.Z. 2006. Land Cover Mapping: A remote sensing approach. *Pakistan Journal of Botany* 38(3): 559-570.
- Miao, X.; Gong, P.; Swope, S.; Ruiliang, P.; Carruthers, R.; Anderson, G.L.; Heaton, J.S.; Tracy, C.R. 2006. Estimation of yellow starthistle abundance through CASI-2 hyperspectral imagery using linear spectral mixture models. *Remote Sensing of Environment* 101: 329-341.
- Mladinich, C.S.; Bustos, M.R.; Stitt, S.; Root, R.; Brown, K.; Anderson, G.L.; Hager, S. 2006. The Use of Landsat 7 Enhanced Thematic Mapper Plus for Mapping Leafy Spurge. *Journal of Rangeland Ecology & Management* 59(5): 500-506.
- Morgan, J.L.; Gergel, S.E.; Coops, N,C. 2010. Aerial Photography: A Rapidly Evolving Tool for Ecological Management. *BioScience* 60: 47-59.
- Mostafa, M.M.R.; 2003. Design and Performance of the DSS. Fritsch, D. (Ed.) proceedings, 49th photogrammetric week, Stuttgart, Germany, September 1-5: 77-87.
- Ohio State University Extension Fact Sheet (OSU Extension Fact Sheet). 2012. Food, Agricultural and Biological Engineering at http://ohioline.osu.edu/aex-fact/0480_87.html Accessed March 7, 2012.

- Pacheco, A.; McNairn, H. 2010. Evaluating multispectral remote sensing and spectral unmixing analysis for crop residue mapping. *Remote Sensing of Environment* 114(10): 2219-2228.
- Peterson, S.H.; Stow, D.A. 2003. Using multiple image endmember spectral mixture analysis to study chaparral regrowth in southern California. *International Journal of Remote Sensing* 24(22): 4481-4504.
- Qi, J.; Marsett, R.C.; Moran, M.S.; Goodrich, D.C.; Heilman, P.; Kerr, Y.H.; Dedieu, G.; Chehbouni, A.; Zhang, X.X. 2000. Spatial and temporal dynamics of vegetation in the San Pedro River basin area. *Agricultural and Forest Meteorology* 105: 55-68.
- Rice, P.M.; Dupuis, V. 2009. Flowering rush: An invasive aquatic macrophyte infesting the headwaters of the Columbia river system. 1-11. Available at Center for Invasive Plant Management (CIPM). <http://www.weedcenter.org/projects-spatial.html>.
- Rice, P.M.; Reddish, M.; Dupuis, V.; Mitchell, A. 2010. Flowering Rush Mapping and Spatial Prediction Model. Available at Center for Invasive Plant Management (CIPM) <http://www.weedcenter.org/research/docs/Flowering%20Rush%20Spatial%20Final%20Report.pdf>
- Santos, M.J.; Anderson, L.W.; Ustin, S.L. 2011. Effects of invasive species on plant communities: an example using submersed aquatic plants at the regional scale. *Biological Invasions* 13: 443-457.
- Schmid, T.; Koch, M.; Gumuzzio, J.; Mather, P.M. 2004. A spectral library for a semi-arid wetland and its application to studies of wetland degradation using hyperspectral and multispectral data. *International Journal of Remote Sensing* 25(13): 2485-2496.
- Shao, G.; Wu, J. 2008. On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecology* 23: 505-511.
- Small, C. 2001. Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing* 22(7): 1305-1334.
- Small, C. 2003. High spatial resolution spectral mixture analysis of urban reflectance. *Remote Sensing of Environment* 88: 170-186.
- Souza Jr, C.; Firestone, L.; Silva, L.M.; Roberts, D. 2003. Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. *Remote Sensing of Environment* 87: 494-506.
- Tuanmu, M.; Viña, A.; Bearer, S.; Xu, W.; Ouyang, Z.; Zhang, H.; Liu, J. 2010. Mapping understory vegetation using phenological characteristics derived from remotely sensed data. *Remote Sensing of Environment* 114: 1833-1844.

- Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. 2003. Remote Sensing for biodiversity science and conservation. *TRENDS in Ecology and Evolution* 18(6): 307-16.
- Underwood, E.; Ustin, S.; Dipietro, D. 2003. Mapping Nonnative Plants Using Hyperspectral Imagery. *Remote Sensing of Environment* 86(2): 150-61.
- Ustin, S.L.; Dipietro, D.; Olmstead, K.; Underwood, K.; Scheer, G.J. 2002. Hyperspectral Remote Sensing for Invasive Species Detection and Mapping. *Geoscience and Remote Sensing Symposium 3*: 1658-1660.
- U.S. Fish and Wildlife Service Ottawa National Wildlife Refuge (USFWS ONWR). 2012. <http://www.fws.gov/refuges/profiles/index.cfm?id=31540> Accessed March 18, 2012.
- Vincent, R.; Qin, X.; McKay, R.; Miner, J.; Czajkowski, K.; Savino, J.; Bridgeman, T. 2004. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment* 89: 381-392.
- Wang, T.; Skidmore, A.K.; Toxopeus, A.G.; Liu, X. 2009. Understory Bamboo Discrimination Using a Winter Image. *Photogrammetric Engineering & Remote Sensing* 75(1): 37-47.
- Windle, P.N.; Kranz, R.H.; La, M. 2008. Invasive Species in Ohio Pathways, Policies, and Costs. *Union of Concerned Scientists*: 1-109.
- Wright, C.; Gallant, A. 2007. Improved wetland remote sensing in Yellowstone National Park using classification trees to combine TM imagery and ancillary environmental data. *Remote Sensing of Environment* 107: 582-605.
- Wu, C. 2004. Normalized spectral mixture analysis for monitoring urban composition using ETM+ imagery. *Remote Sensing of Environment* 93: 480-492.
- Xie, Y.; Sha, Z.; Yu, M. 2008. Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology* 1(1): 9-23.
- Xu, M.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.J. 2005. Decision tree regression for soft classification of remote sensing data. *Remote Sensing of Environment* 97: 322-336.
- Yang, C.; Everitt, J.H.; Murden, D. 2011. Evaluating high resolution SPOT 5 satellite imagery for crop identification. *Computers and Electronics in Agriculture* 75: 347-354.
- Yuan, F.; Sawaya, K.E.; Loeffelholz, B.C.; Bauer, M.E. 2005. Land cover classification and change analysis of the Twin Cities (Minnesota) Metropolitan Area by multitemporal Landsat remote sensing. *Remote Sensing of Environment* 98: 317-328.

- Zambon, M.; Lawrence, R.; Bunn, A.; Powell, S. 2006. Effect of Alternative Splitting Rules on Image Processing Using Classification Tree Analysis. *Photogrammetric Engineering & Remote Sensing* 72(1): 25-30.
- Zhang, Y.; Lu, D.; Yang, B.; Sun, C.; Sun, M. 2011. Coastal wetland vegetation classification with a Landsat Thematic Mapper image. *International Journal of Remote Sensing* 32(2): 545-561.

APPENDIX A: FIGURES

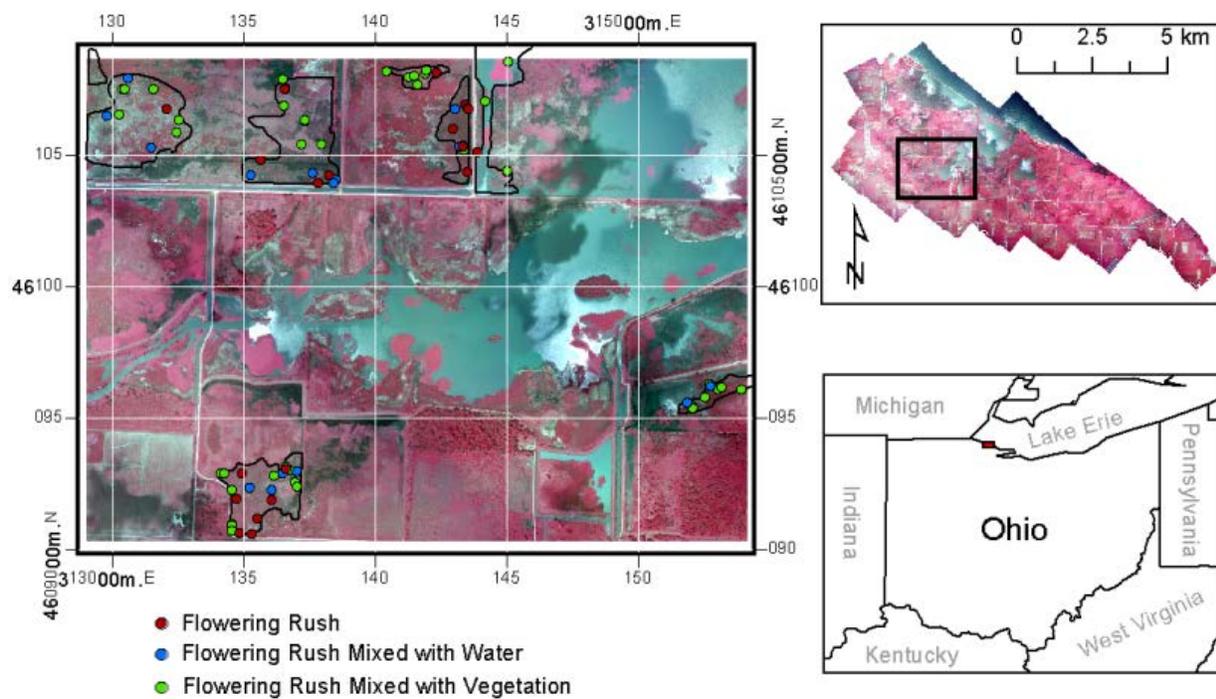


Figure 1: Ottawa National Wildlife Refuge (ONWR) study area with subarea boundaries, and sampling locations.

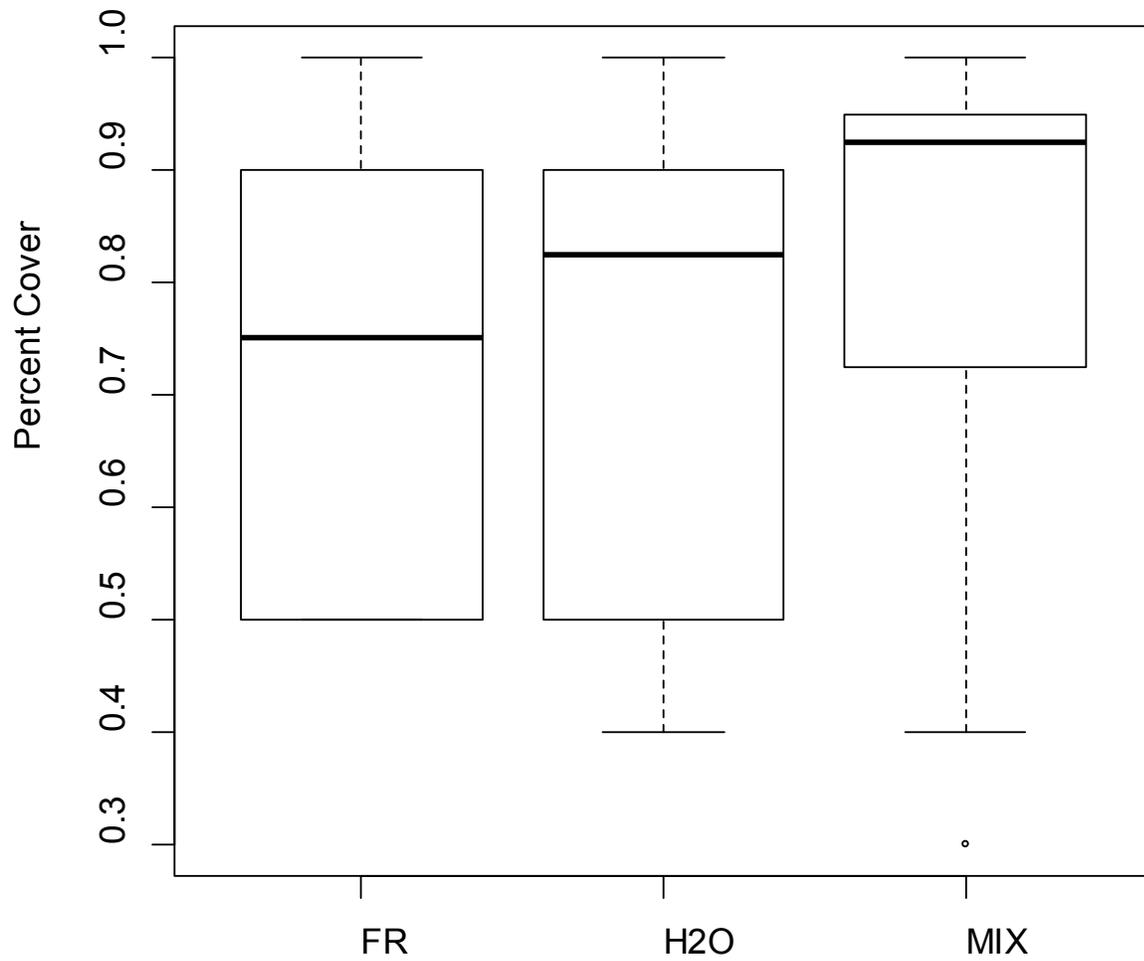


Figure 2: Predetermined vegetation classes and the percent cover associated with the observations made during field data collection

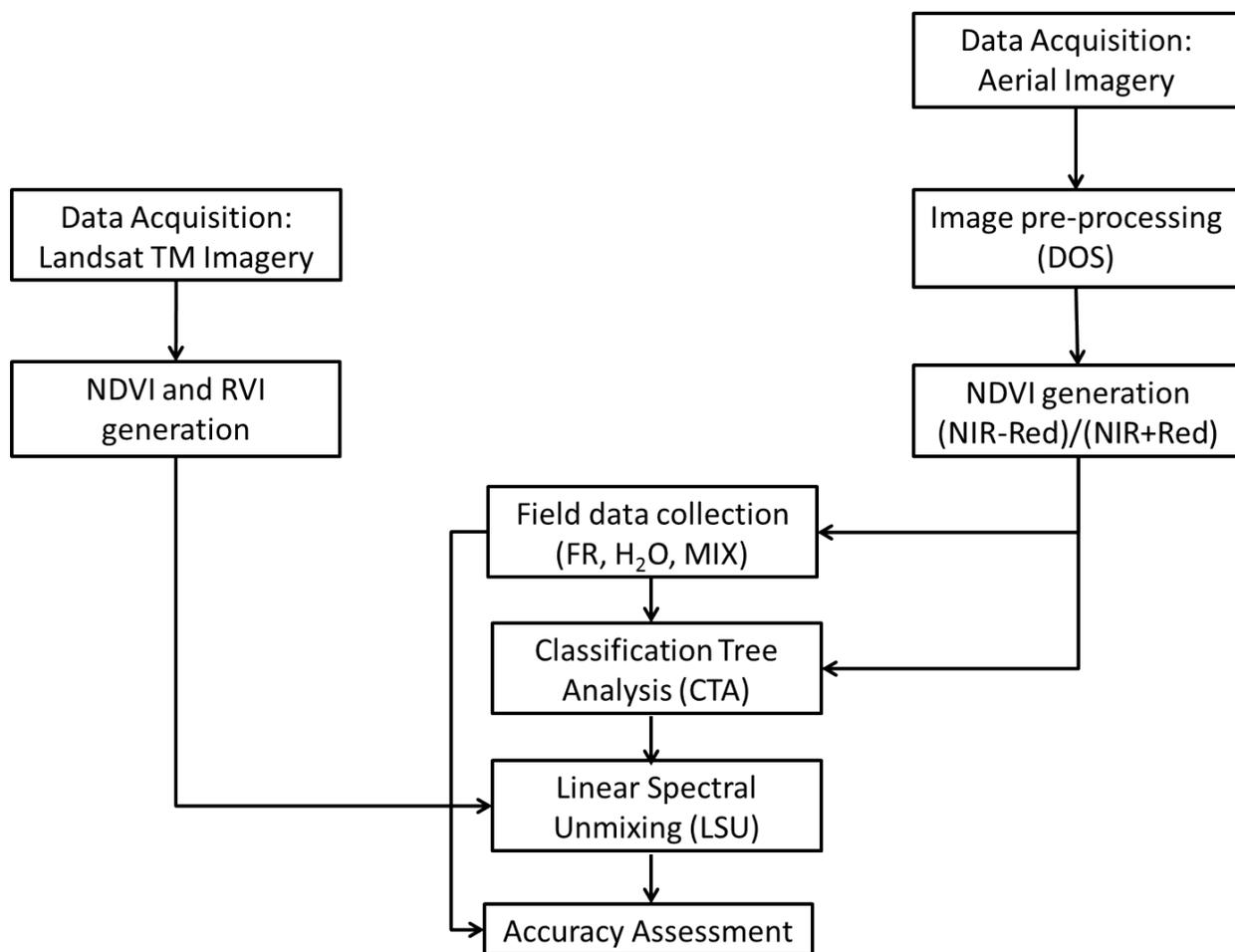


Figure 3: Methodology used in current study

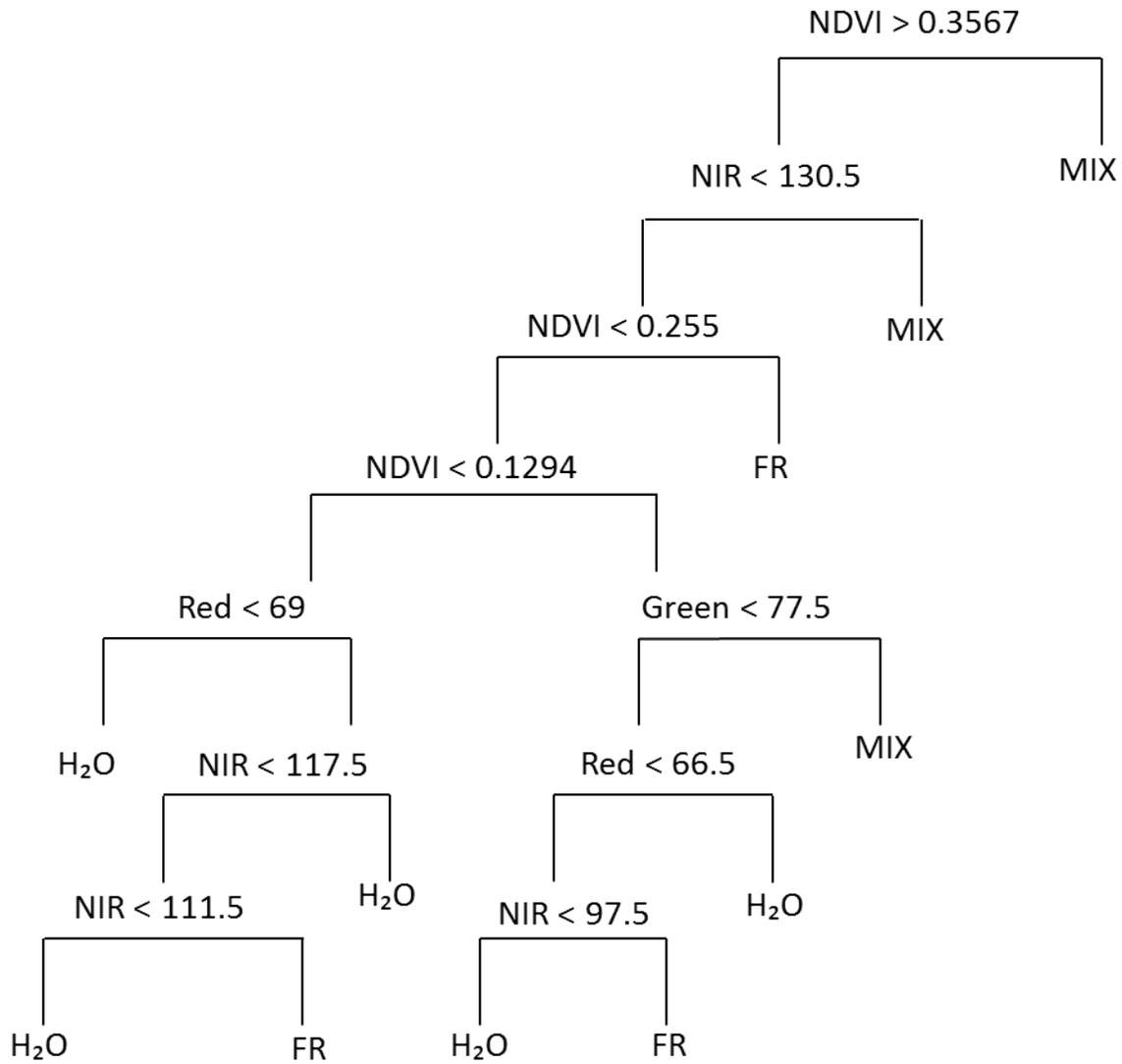


Figure 4: Classification tree with a kappa of 0.567. Predictor variables include bands 1 (green), 2 (red), 3 (near-infrared), and NDVI.

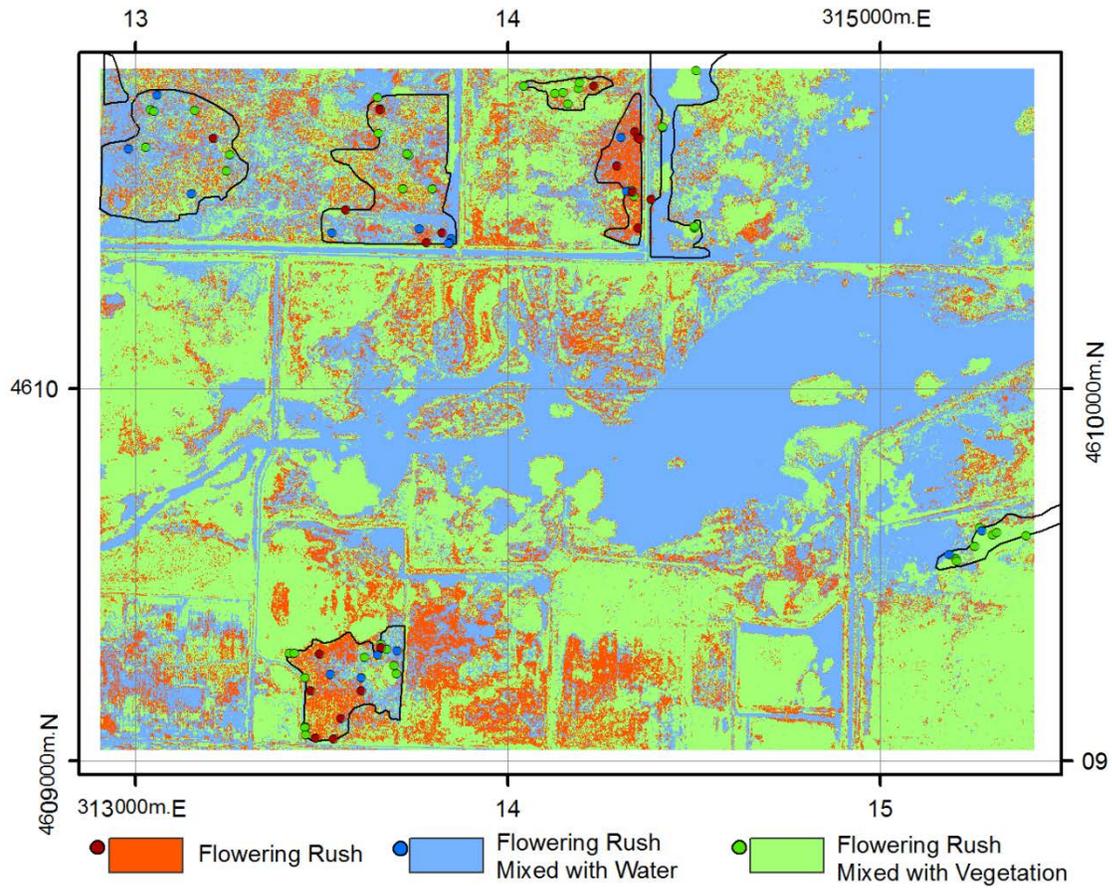


Figure 5: Prediction map derived from the aerial imagery using CTA. Map includes Flowering rush (FR), Flowering Rush mixed with Water (H_2O), and Flowering Rush Mixed with Vegetation (MIX).

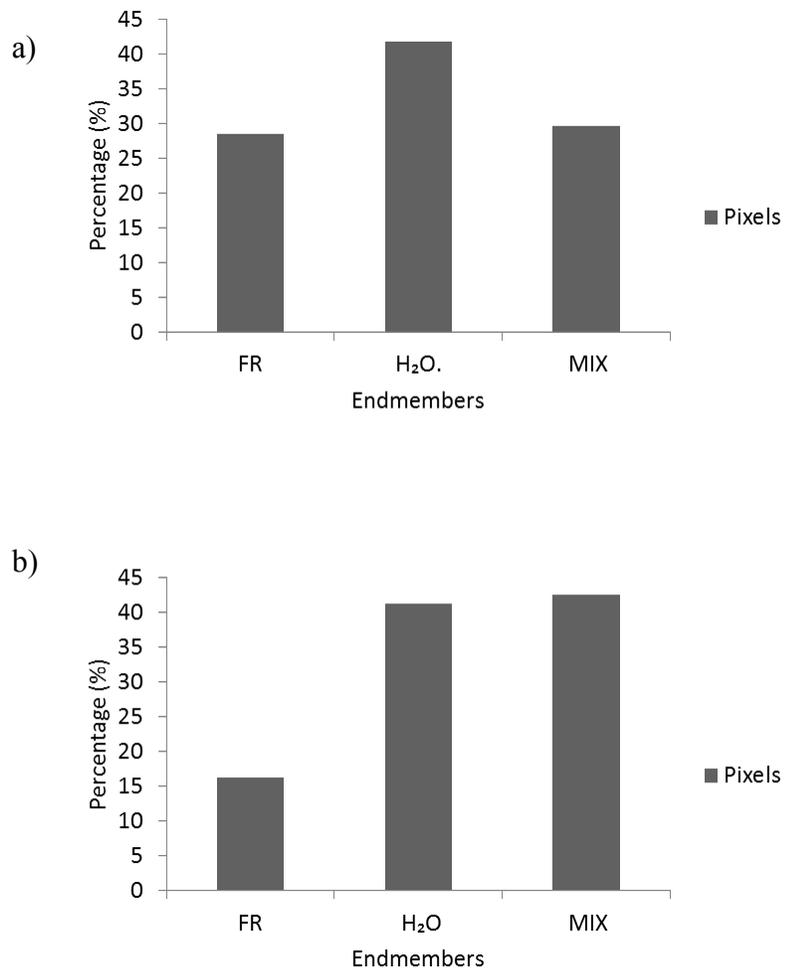


Figure 6: Percentage of the total pixels for each predicted endmember. a) Percentage of the total pixels for the subsets, and b) percentage of the total pixels for the entire map (Figure 5).

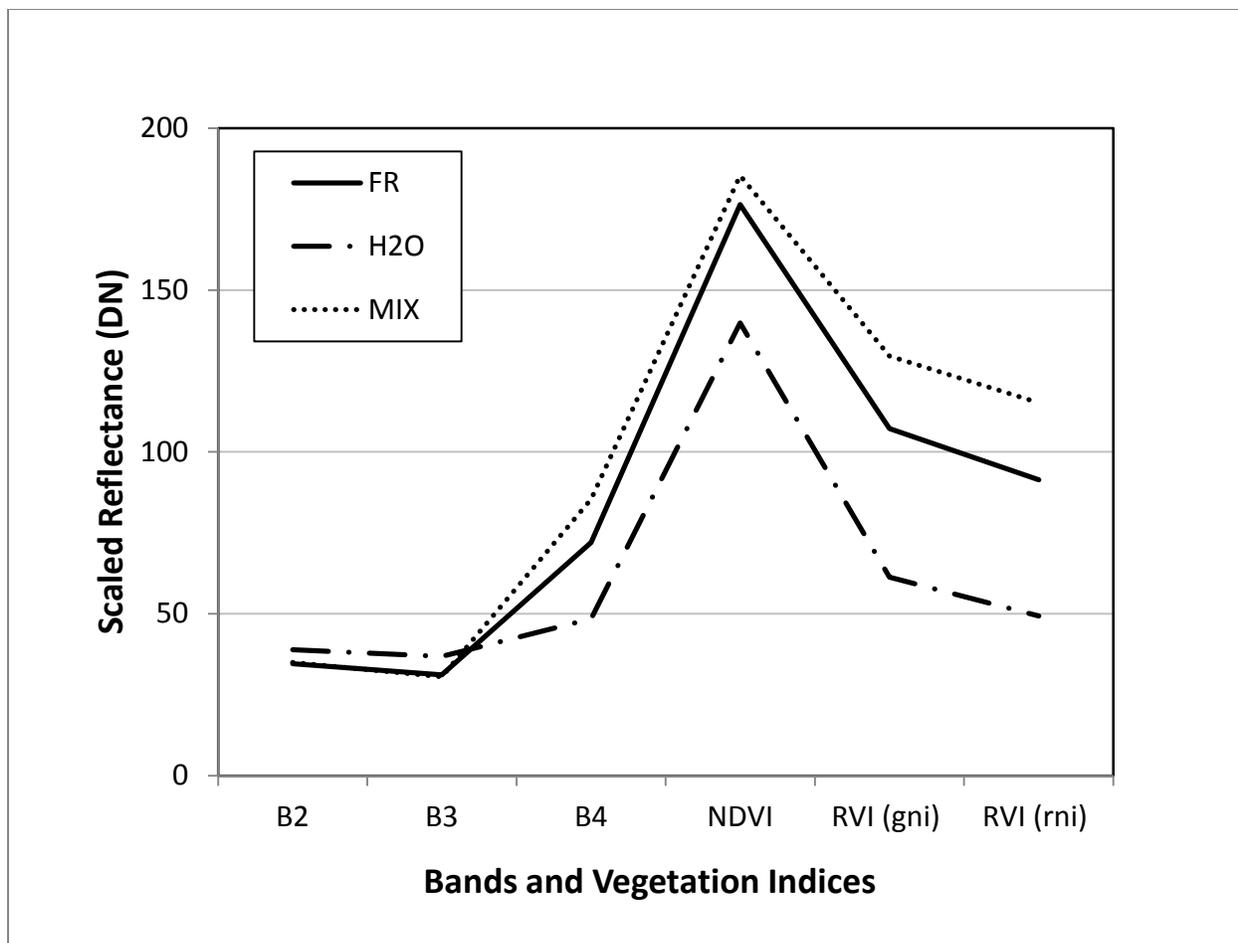


Figure 7: Endmembers separability plot using Landsat TM bands 2(green), 3(red), 4(nir), NDVI, ratio (2/4), and ratio (3/4).

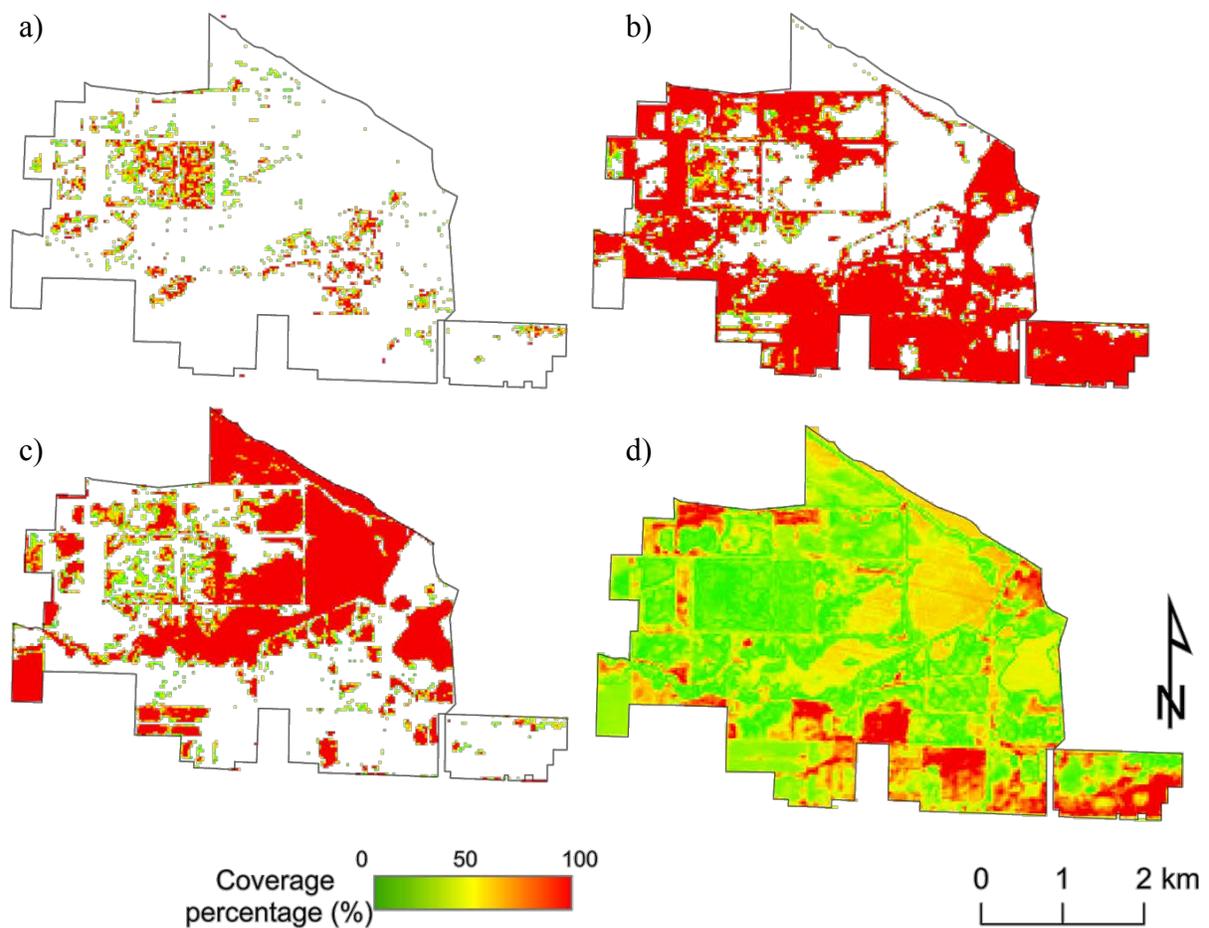


Figure 8: Linear spectral unmixing fraction maps for the three selected endmembers derived from Landsat TM imagery. a) flowering rush, b) flowering rush mixed with other vegetation, c) flowering rush mixed with water, and the d) residual image with the total error for each pixel.

APPENDIX B: TABLES

Split Type	Pruning (%)	Kappa (%)
Gini	1.8	0.459
Entropy	1.8	0.393
Ratio	1.8	0.359
Gini	1.5	0.465
Gini	2.0	0.465
Gini	1.3	0.567
Ratio	1.5	0.359
Ratio	1.0	0.435
Entropy	1.0	0.472
Entropy	2.0	0.382
Entropy	0.8	0.497

Table 1: Comparative accuracies for classification of the 2011 aerial imagery of ONWR. Inputs: G,R,NIR, and NDVI

	FR	H₂O	MIX	Total	PA (%)	EO (%)	EC (%)
FR	15	3	3	21	71.4	28.6	31.8
H₂O	2	13	3	18	72.2	27.8	43.5
MIX	5	7	32	44	72.7	27.3	15.8
Total	22	23	38	83			
CA (%)	68.2	56.5	84.2				

Table 2: The number of correctly classified pixels, the producer's accuracy (PA), consumer's accuracy (CA), error of emission (EO), and error of commission (EC) for each endmember from the final CTA map.

Separability between:	Transformed Divergence
Flowering rush (FR) and flowering rush mixed with water (H ₂ O)	1995.79
Flowering rush (FR) and flowering rush mixed with other vegetation (MIX)	1919.61
Flowering rush mixed with water (H ₂ O) and flowering rush mixed with other vegetation (MIX)	2000.00
Average over all pairwise combinations	1971.80

Table 3: Transformed divergence separability between the three endmembers of FR, H₂O, and MIX

Coverage per pixel (%)	%		
	FR	H₂O	MIX
0-20	90	57	43
20-40	2	3	2
40-60	2	3	2
60-80	2	3	2
80-100	4	35	51

Table 4: Percentage of the total pixels for each endmember corresponding to the coverage (%) per pixel.

Coverage per Pixel (%)	FR	MIX	H ₂ O
	0-20	10	20
20-40	2	0	1
40-60	1	2	2
60-80	5	1	4
80-100	3	21	4
Total	21	44	18

Table 5: Validation of fraction maps