Vegetation change analysis from 2010-2018 using aerial photography and RTK-GNSS to assist Lake Mattamuskeet Restoration Efforts in North Carolina, USA

by

Rachel Smaby

July 2021

Director of Thesis: Hannah M. Cooper

Major Department: Geography, Planning, and Environment

Mapping vegetation species and documenting their changes is critical to achieving the Lake Mattamuskeet National Wildlife Refuge restoration and habitat management efforts of controlling the spread of invasive species. This study maps the dominant vegetation species and inventories their changes in the Lake Mattamuskeet waterfowl impoundment 4 using time series aerial images collected 2010 – 2018. RTK-GNSS surveys were conducted in the field to record the locations of dominant species used as reference data. The reference data were matched to their respective vegetation patch delineated by object-based image analysis. The Random Forest machine learning classification algorithm was used to accurately predict the unknown locations of the dominant species. The algorithm had an overall accuracy ≥76 and Kappa statistic ≥66. Phragmites australis (phragmites) expanded in 2016 but was constrained in 2018 by Echinochloa walteri (Walter's millet). The Refuge's goal of achieving 50% good waterfowl species was not met during the time series investigated, as the largest cover of good waterfowl habitat was only 25% in 2018. The results of this research provide insights about the effectiveness of current vegetation management techniques implemented at the Refuge for waterfowl impoundment 4 and have implications for global wetland mapping and change analysis.

Vegetation change analysis from 2010-2018 using aerial photography and RTK-GNSS to assist Lake Mattamuskeet Restoration Efforts in North Carolina, USA

A Thesis

Presented to the Faculty of the Department of Geography, Planning, and Environment

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Geography

by

Rachel Smaby

July 2021

Vegetation change analysis from 2010-2018 using aerial photography and RTK-GNSS to assist Lake Mattamuskeet Restoration Efforts in North Carolina, USA

by

Rachel Smaby

APPROVED BY:

DIRECTOR OF

THESIS:____

(Hannah Cooper, PhD)

COMMITTEE MEMBER:____

(Thad Wasklewicz, PhD)

COMMITTEE MEMBER:____

(Yong Wang, PhD)

CHAIR OF THE DEPARTMENT

OF GEOGRAPHY, PLANNING, AND ENVIRONMENT:________________________

(Thad Wasklewicz, PhD)

DEAN OF THE

GRADUATE SCHOOL:____

(Paul J. Gemperline, PhD)

DEDICATIONS

I would like to dedicate my thesis to my family for their support and motivation in helping me reach this goal of earning an advanced degree.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Hannah Cooper, and my committee members (Dr. Thad Wasklewicz, Dr. Yong Wang) here in the Department of Geography, Planning, and Environment at East Carolina University. I would also like to thank the Mattamuskeet National Wildlife Refuge staff for their support and assistance with this research.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

1.1 Significance of coastal wetland monitoring

Coastal wetland areas are known to be biologically diverse ecosystems that provide a number of ecosystem functions such as flood protection, wildlife food and habitat, and improved water quality (Rojas et al. 2017). Given the importance of these wetland ecosystems, it's of great concern that wetland acreage loss is occurring in many areas globally. In an analysis examining the Atlantic, Pacific, Gulf of Mexico, and Great Lakes watersheds between 2004 and 2009, it was found that wetlands were lost on an average rate of over 80,000 acres per year (Dahl and Stedman 2013). Many factors including human activity, climate change, and invasive species have contributed to coastal wetland loss (Xie et al. 2010). These changes make it necessary to implement accurate and detailed monitoring techniques in order to assess the rate of change. In areas where further management efforts are possible, the mapping of wetland change can help to determine the effectiveness of management strategies against factors such as invasive species.

1.2 Significance of vegetation management at Lake Mattamuskeet

Coastal National Wildlife Refuges (NWRs) provide critical habitat areas of wetlands, estuaries, and beaches for migratory birds and many threatened or endangered species. In coastal North Carolina, Mattamuskeet NWR is home to the state's largest natural lake, Lake Mattamuskeet. It is a shallow coastal lake (up to 1 m depth) that provides critical habitat for various flora and fauna valued by coastal residents and visitors (North Carolina Coastal Federation 2018). However, the lake has a history of ecological problems that complicate managing its habitat. In the early 1900s, a series of canals were dug, hoping to drain the lake into the Pamlico Sound so that the fertile lakebed could be used for farming. Unfortunately, these efforts were deemed unsuccessful. The lake water became brackish as saltwater from the Pamlico Sound came in through the canals. This changed in 1930s when the first tide gates were installed in the channels, essentially transitioning the lake back from the brackish water to mostly freshwater ecosystems. The original tide gates have since been replaced multiple times, with the newest and most effective gates being installed in 2011 and 2014 for all four canal gates (North Carolina Coastal Federation 2018).

In addition to the lake itself, Mattamuskeet NWR staff control the waterfowl impoundment water levels using the established network of canals, water control structures, and pumps. Waterfowl impoundments are wetlands that are hydrologically managed to provide critical habitat and food for migratory birds. A total of 14 bird impoundments were created between 1967 and 1980 to provide critical habitat between the months of November and March. Most of the impoundments were originally classified as moist soil units, dictating what types of vegetation would be planted and managed in these areas for the waterfowl such as *Echinochloa walteri* (hereinafter Walter's millet) and *Ludwigia alternifolia* (seedbox). Walter's millet is known to grow in wet areas, often in shallow water, and can grow in both undisturbed and disturbed environments (Heller, and Michael 2021). Seedbox is typically found in wet, low lying areas such as swamps or roadside ditches (Duke 1955). The spread of invasive species, however, such as the resilient wetland grass *Phragmites australis* (phragmites) and *Centella asiatica* (centella) are a threat to the beneficial wetland plants. Ideal growing conditions for centella include damp habitats with partial to full sun (Devkota and Jhuan 2010). These concerns prompted the NC Wildlife Resources Commission and US Fish and Wildlife to contract the Coastal Federation to develop the Lake Mattamuskeet Watershed Restoration Plan (North Carolina Coastal Federation 2018).

A major objective of the Lake Mattamuskeet restoration plan is to expand efforts to control the spread of invasive species such as phragmites on the Refuge. The efforts include a combination of mechanical and chemical management tactics. Herbicide is currently the most used management method for controlling the spread of phragmites (Martin *et al.* 2013; Hazelton *et al.* 2014). Since 2008, Mattamuskeet NWR has been able to apply herbicide to phragmites annually (U.S. Department of the Interior, 2008). Accurate vegetation maps are useful for the Refuge as they make it possible to assess the effectiveness of treatment strategies over time in relation to containing invasive vegetation species.

1.2 Previous efforts to vegetation mapping at Lake Mattamuskeet

The U.S. National Wetland Inventory maps and Status Trends (https://www.fws.gov/wetlands) is the authoritative data source on the spatial distribution of a species change over time. These maps are an excellent tool for viewing wetland types displayed as groups of similar classifications (e.g., all freshwater emergent wetlands are grouped as one) as well as the area of each wetland type and how it has changed over time. However, resource managers are more interested in knowing where phragmites are spreading and which specific native wetland species are most impacted by the invasion. Thus, the classifications of wetland types displayed as individual species are needed. Since this data is not available for Mattamuskeet NWR, this requires the use of modern remote sensing techniques to create species maps.

A common method used to analyze data in vegetation surveys is object-based image analysis (OBIA). The OBIA algorithm partitions images into objects and assesses the objects' spatial, spectral, and temporal aspects (Hay, 2006). This can be useful for various fields since objects can depict features such as shape and texture that cannot be seen otherwise in a single pixel. Additionally, this can help eliminate observer bias in research and make studies more repeatable. This technology is a particularly powerful tool for assessing ground cover from photographs. Specifically, in one study, researchers determined the percent land cover of five vegetation types by segmenting the images based on color and shape (Luscier et al. 2006). Wetland plant species are also known to have low spectral contrast between species, making it difficult to use this information for OBIA (Luscier et al. 2006). However, Pande-Chhetri et al. (2017) found that the classification of wetland vegetation was more accurate using the OBIA approach compared to pixel-based classification methods.

Machine learning is another powerful analysis tool, which has been found useful when used alongside OBIA for classifying vegetation (Zhang et al. 2017). Machine learning works by learning a predictive model given a set of data and taking this information to predict the values for unseen properties or cases (Stojanova et al. 2010). In one study, researchers were able to use both OBIA and the machine learning method to map plant species with 85% accuracy (Zhang and Xie 2013). A separate study found using OBIA and machine learning to be suitable for mapping terrestrial coastal areas, and suggests that this technique may be a beneficial alternative to other methods such as manually interpreting aerial images or relying only on field work surveys (Juel et al. 2015). This technique has also been successful at classifying vegetation in a dry savanna ecosystem using the Random Forest classifier algorithm (Mishra and Crews 2014). The combination of OBIA and machine learning has also been used successfully in the historical analysis of vegetation growth. Specifically, research on the vegetation distribution in parts of the Florida Everglades effectively combines object-based image classification, object-based change analysis, and machine learning to map a series of historical changes in plant biodiversity (Zhang et al. 2017). The success of this study suggests that these digital techniques can provide an accurate and efficient alternative to manual aerial photography interpretation, but more research is necessary to determine its effectiveness for other land types, such as those in Mattamuskeet NWR.

1.3 Objectives

The primary objective of this study is to examine how the dominant vegetation species within the Mattamuskeet NWR's most popular bird impoundment has changed over time from 2010 to 2018. Changes are expected to be attributed to known management practices within the waterfowl impoundment such as staff efforts to control the spread of invasive species to managing water levels. Vegetation species maps are needed to explore these potential changes. However, limited species maps exist for Mattamuskeet NWR. Additional objectives for the study are based on the Lake Mattamuskeet Watershed Restoration Plan and the Habitat Management Plan for the Refuge. The purpose of the Habitat Management Plan to provide guidelines for management decisions and a long-term vision for the Refuge (U.S. Fish and Wildlife 2018). These goals are to determine whether the phragmites population has decreased during the study period, and to examine whether the area of the study site has consisted of 50% good waterfowl vegetation during the 10-year period (U.S. Fish and Wildlife Service 2018). The Refuge is working to achieve this goal through a series of vegetation management efforts including but not limited to techniques such as mowing or burning dead phragmites, applying herbicide, and disking treated areas (U.S Fish and Wildlife Service 2018). The objectives of this study were examined by collecting, *in situ* measures of dominant vegetation species using Real-time Kinematic Global Navigation Satellite Systems (RTK-GNSS) in the field and by using automated mapping methods implemented in the laboratory, including OBIA and machine learning classification algorithms to generate vegetation species maps. These changes are represented numerically in graphs, as well as visually with distribution and change maps. The change analysis will serve as a resource for restoration managers and planners in the future as the Lake Mattamuskeet Watershed Restoration Plan progresses.

2. Study site and data

2.1 Study site

Mattamuskeet NWR is located in Hyde County of coastal North Carolina (Figure 1A). Vegetation management strategies of the Refuge focus on the reduction of phragmites and *Centella asiatica* (hereinafter centella) to expand the growth of habitat that is good for waterfowl such as Walter's millet and seedbox. A total of 14 bird impoundments surround the lake (Figure 1B). The specific management plan for impoundment 4, the focus of this research, includes mowing dead phragmites in May, checking for phragmites and spraying herbicide in June, burning if possible in July, spraying herbicide in August, disking the treated areas in September, and reflooding the impoundment for the winter months in October (U.S. Fish and Wildlife Service 2018).

Impoundment 4 near the main visitor center and along the Refuge wildlife drive is the major attraction for bird watching. Per the Refuge's biologist's recommendation, this study focuses on vegetation change in impoundment 4 (Figure 1B).

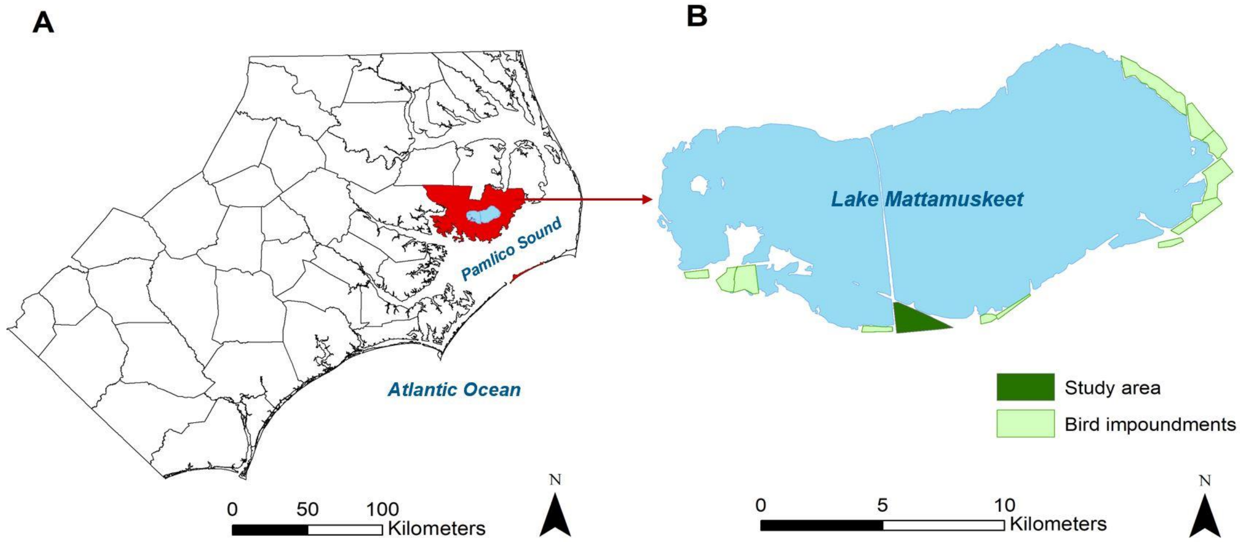

Figure 1. The study area located in Hyde County (A) in the Mattamuskeet National Wildlife Refuge which contains 14 bird impoundments and focuses on the study area impoundment (B).

2.2 Data

Data used in this study include aerial imagery to generate image objects, Light Detection and Ranging (LiDAR) Digital Elevation Model (DEM) to define the elevation of each object, and reference data to train and validate the classification of each image object into a vegetation species. The imagery used for this research was collected every two years from 2010-2018 through the National Agriculture Imagery Program (NAIP) during the agricultural growing season in the U.S. (April through October). The images consist of red/green/blue (RGB) bands as well as near-infrared (NIR). The 2018 data was collected in June 2018, with a 0.6 m resolution cell size. The imagery for the years 2016 (collected in May), 2014 (collected in August), 2012 (collected in June), and 2010 (collected in July) have a 1-meter resolution. For this study, the 2018 imagery resolution was resampled to 1 m to match the cell size and alignment of the 2010, 2012, 2014, and 2016 imagery. All NAIP imagery collected since 2009 meet a horizontal accuracy specification of ±6 m that matches to true ground specifications.

The LiDAR DEM was derived from ground returns collected in April 2014 by the North Carolina Flood Mapping Program and was acquired from NOAA Data Access Viewer (https://coast.noaa.gov/dataviewer/). The reported DEM's horizontal accuracy is 100 cm and vertical accuracy is 6 cm for bare earth. The LiDAR DEM from 2014 was used for all years because no other LiDAR data is available for the study area between 2010 and 2018. An underlying assumption in this research is that the study site did not undergo any major elevation changes during that time.

Reference data, or ground truth measurements taken in the field, are not available at the species level for Mattamuskeet NWR. Therefore, reference data for this study was collected during the growing season in August 2020, which was after some of the vegetation management practices had already occurred for Impoundment 4. This was due to delays and travel restrictions because of the coronavirus pandemic. Although herbicide was applied to some areas of the impoundment in August during the time frame when the reference field data was collected, dense stands were found in the field (see Figure 2).

We used a Trimble Spectra Precision SP80 Real-time Kinematic Global Navigation Satellite Systems (RTK-GNSS), which has high-precision static post-processed accuracy of 3 mm in the horizontal and 3.5 mm in the vertical (Root Mean Square Error, RMSE) (Trimble, 2017). During the RTK-GNSS surveys, three SP80 receivers were used. One was set up as a base station to establish a benchmark, and the other two were rovers positioned level with the ground surface for 60 epochs to record each location's species. The OPUS solution was used to later correct the data (Table 1). A total of 196 points were collected, centella = 33, phragmites = 61, seedbox = 13, Walter's millet = 19, loblolly pine = 20, and open water = 50. The points for both open water and loblolly pine were manually selected in ArcMap 10.8 using the 2018 imagery.

Figure 2. Five dominant vegetation species surveyed in September 2020 within bird impoundment 4 at Mattamuskeet NWR.

Table 1. OPUS solution for the benchmark established in this study to conduct vegetation surveys.

NAD83 2011 (EPOCH: 2010)	RMSE	State Plane Coordinates: SPC (3200 NC)
N 35º 27" 5.193'	0.001 (m)	Northing: 192332.198
W 76º 11" 50.949'	0.006 (m)	Easting: 863959.901
Ellipsoid Height: -37.882 (m)	0.011 (m)	
Orthometric height: 0.390 (m) [NAVD88 using	0.058 (m)	
Overall	0.012 (m)	

3. Methods

This study follows the workflow presented in Figure 3 to create vegetation species maps of bird impoundment 4 in Mattamuskeet NWR from 2010-2018. In the workflow, 4 bands (R, G, B, NIR) were used for the years 2010-2018 because the addition of the NIR band is useful for plant identification, as it allows for the identification of plants based on spectral characteristics for which provides more accurate information (Antonio and Zain 2001). A LiDAR DEM was also used as it may be useful for identification of plants based on elevations. LiDAR was included based on previous research which has found that microtopography can influence distribution of wetland plant communities (Courtwright and Findlay 2011). The use of ground LiDAR in combination with aerial imagery is thought to be helpful for machine learning algorithms, as this ensures that both spectral information as well as ground surface features are incorporated in predictions (Zurqani et al. 2020). Image normalization was then performed using the 2018 imagery as a base image to ensure that the images had similar radiometric scales to help with the accuracy of vegetation classification. Object-based Image Analysis (OBIA) was completed to distinguish vegetation patches for each year. Through this process, the mean brightness values of all bands for the five imagery years were extracted to the corresponding object location. Machine learning was also used for these research methods, a technology which has not been used previously to map vegetation species at Mattamuskeet NWR. Machine learning was used to predict vegetation distributions for all five years of this research to assess how the vegetation distribution has changed over the study period.

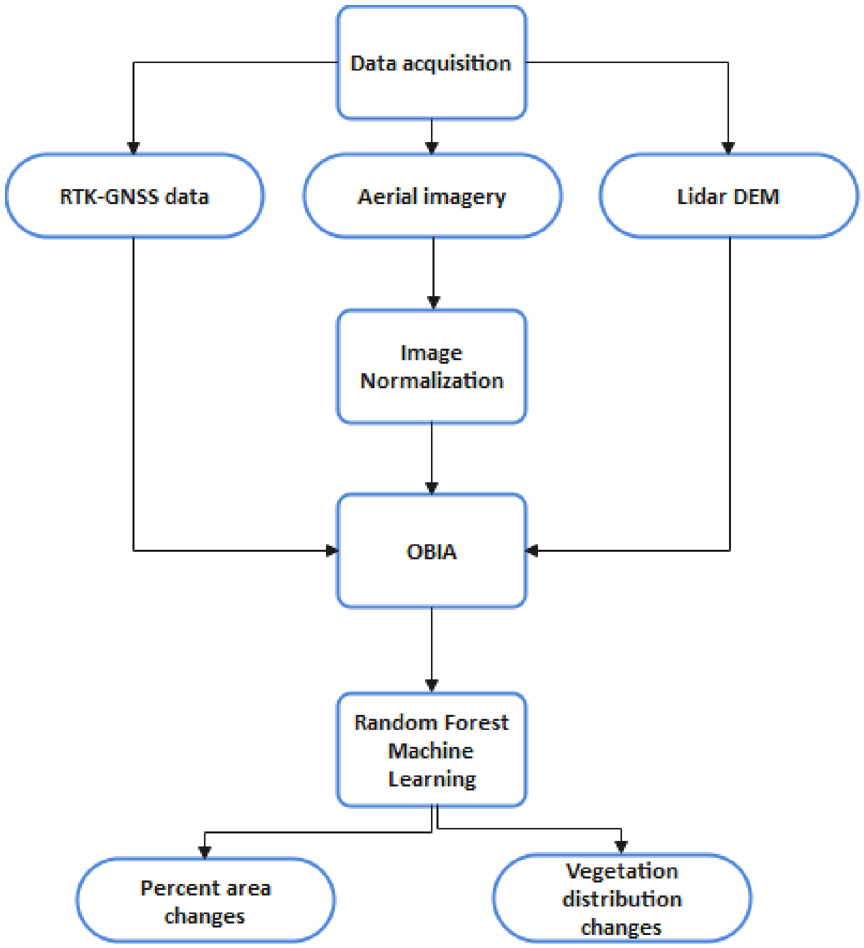

Figure 3. Framework used in this study to generate species maps and calculate area of change.

3.1 Image normalization

To assess the current and historical vegetation distribution, the aerial images need to be normalized. Image normalization works by using a reference image to reduce the variation of pixel intensity values band by band between multi-date images. The idea is that the images appear to have come from the same survey and thus take on the same spectral characteristics (Jensen 2015). To perform image normalization, the 2018 imagery was assigned as the base image. The cell resolution of the 2018 image was converted to 1m to match the reference imagery cell size. The 2018 raster was also snapped to the 2016 raster to align the pixels in each image. The reference images were then transformed so that the radiometric scales of each image matched the 2018 image radiometric scale. This was done by identifying areas where features or plant distributions remained relatively unchanged between years with the use of ENVI and ArcMap. These unchanging features are known as pseudo-invariant features (PIF's) (Zhang et al. 2017). In total, 13 PIF's were used for the years 2016 and 2014. 7 of these same PIF's were used for the 2012 image, and 8 of these same PIF's were used for the 2010 image. Less PIF's were used for the 2010 and 2012 images is because some of the unchanged man-made features in the study area had not yet been constructed at these points in time. The individual bands (R, G, B, NIR) of each of the 4 reference images were then exported as separate images. With the individual bands of each of the 4 reference images, regression equations were formed using the Brightness Values (BV's) of the selected PIF's and comparing the values to the corresponding brightness values of the 2018 image. R^2 values were also calculated. All bands of each image have a minimum of a 0.7 R^2 value (Table 3). Linear graphs were created to show the relationship between brightness values of the PIF's between the base image bands and reference image bands (Figure 2). The radiometrically normalized images are based on the original reference images and transformed with the use of the regression equations in ArcMap. The image normalization was completed by compiling the 4 normalized bands together for each of the reference images to create one normalized image for each of the reference years (Figure 3).

Table 2. Equations which were used in ArcMap to normalize all images, using 2018 for the base image. R^2 values are also shown for each year.

Year	Equation	R^2
2010	y = 1.2594 * band1 – 47.702	0.7880
	y = 0.9273 * band2 + 5.8409	0.8556
	y = 0.8076 * band3 + 32.762	0.8450
	y = 1.3389 * band4 – 119.19	0.7664
2012	y = 1.0922 * band1 – 33.046	0.9099
	y = 0.8141 * band2 + 25.865	0.8660
	y = 0.8175 * band3 + 24.231	0.8919
	y = 0.9872 * band4 + 32.002	0.7022
2014	y = 1.3045 * band1 – 73.82	0.8686
	y = 1.1637 * band2 – 46.548	0.8636
	y = 1.2213 * band3 – 59.096	0.8569
	y = 1.5673 * band4 – 71.789	0.7385
2016	y = 1.3933 * band1 – 44.393	0.8847
	y = 1.1188 * band2 – 26.047	0.8823
	y = 1.1170 * band3 – 20.754	0.8484
	y = 1.5678 * band4 – 89.461	0.8463

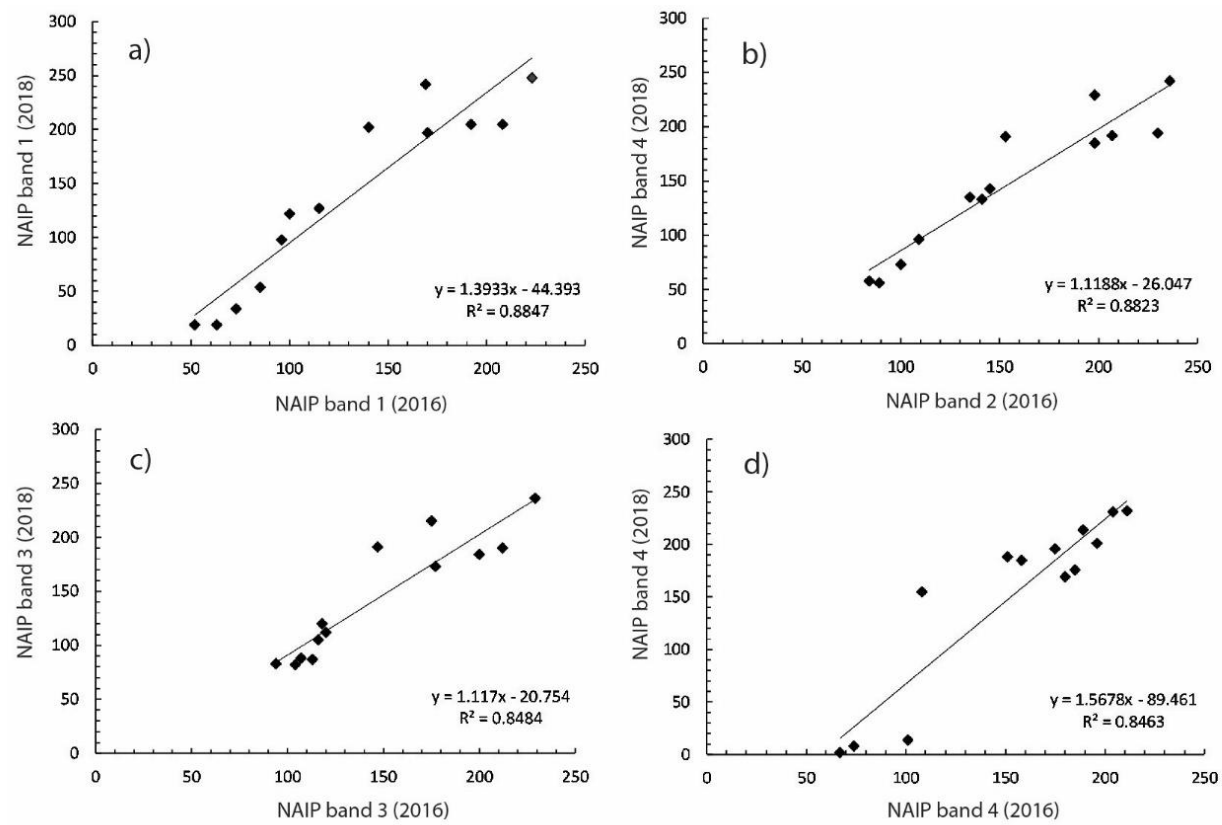

Figure 4. Graphs comparing the pseudo-invariant feature pixel brightness values of each of the 4 bands in the 2016 and 2018 images.

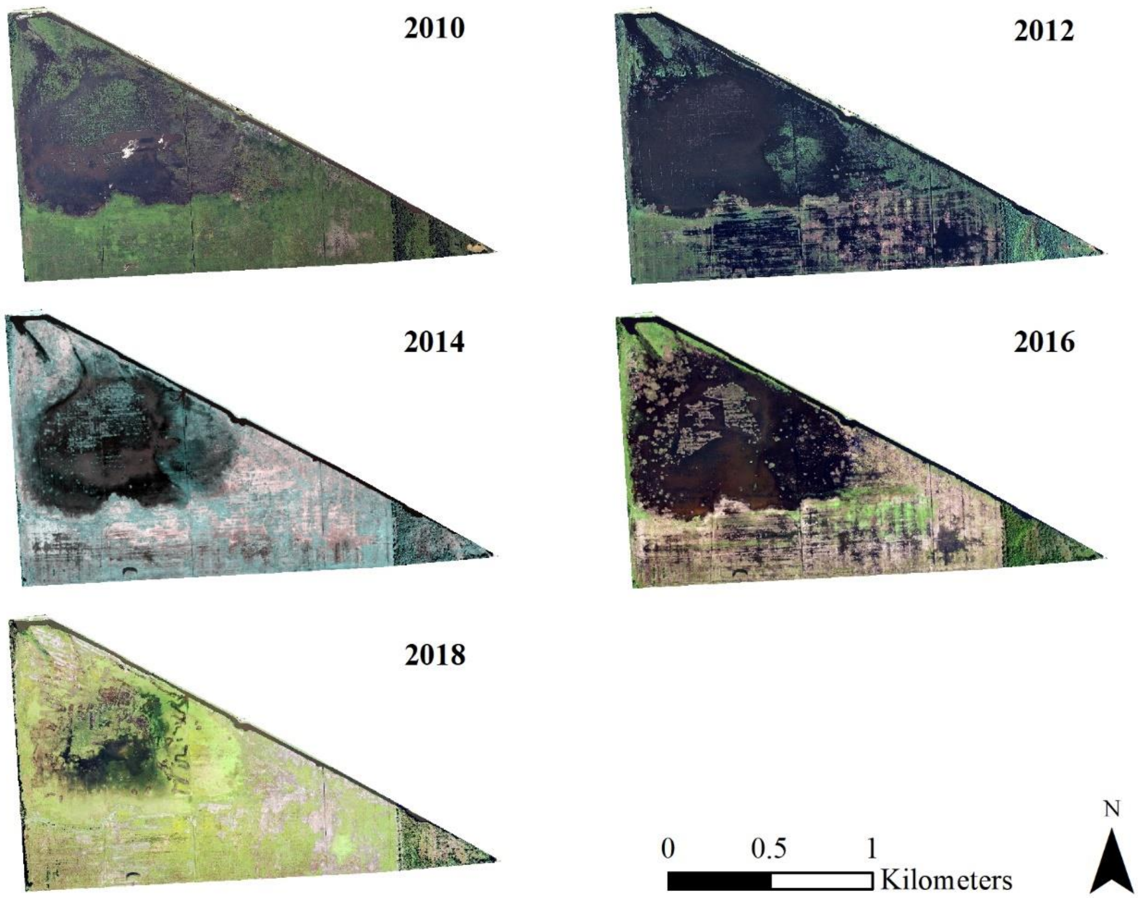

Figure 5. Normalized reference images based on the 2018 base image radiometric scale.

3.2 Object-Based Image Analysis (OBIA)

A multiresolution segmentation algorithm in eCognition 10.1 was used to create objects with the normalized study site images. The object scale of 15 was chosen to best represent the vegetation patches after testing sample scales of 15, 25, 50, 75, and 100. Visually, this scale of 15 appeared to be best fit based on identifiable vegetation patches, such as loblolly pines, which could be seen in the aerial imagery. A color/shape criteria of 0.2/0.8 was chosen in order to more heavily weigh the shape of the vegetation objects. A smoothness/compactness criteria of 0.5/0.5

was chosen to equally weigh compact and non-compact segments. This process was conducted using the aerial imagery for each of the five years in order to find the unique vegetation patches and mean brightness values of the image bands.

3.3 Random Forest Machine Learning classification

The Random Forest classifier works by creating multiple decision tree outputs and choosing the final output based on the most popular class (Pal 2005). Random Forest was chosen based on previous research which has found the algorithm to increase the classification accuracy for land cover classification studies (Friedl et al. 1999; Muchoney et al. 2000). A total of 196 samples were used to train the Random Forest machine learning classifications. The training data were used for all years because it is not possible to collect *in-situ* data for previous years, and it can be inaccurate to attempt to manually identify multiple plant species using aerial imagery (Zhang et al. 2017). Attributes for each of the five training datasets (one for each year) included the vegetation species, location, and mean brightness values of all four imagery bands for each object. The mean brightness values were calculated during the creation of image objects in eCognition 10.1 by Trimble. As a result, all five training dataset years were associated with different objects and mean brightness values, which is what differentiated the training datasets. In addition, LiDAR DEM statistics were included for all 5 training datasets, and it was assumed that the study site did not undergo any major elevation changes. Zonal statistics as Table tool within the Spatial Analyst toolbox in ArcMap 10.8.1. was used to calculate LiDAR DEM attributes of Count, Area, Minimum, Maximum, Range, Standard Deviation, and Sum. These attributes are different for each year’s set of objects and were incorporated into each of the training datasets.

Weka 3.8.5, an open source software developed by the University of Waikato, was used to conduct the machine learning component of this research (Frank et al. 2016). The classification algorithm chosen for this study was Random Forest. The classification results for the training data were found using a 10-fold cross validation. In 10-fold cross validation, the training dataset is randomly split into 10 subsets, and the cross-validation accuracy is the average of all 10 training accuracies (Berrar 2019).

Testing datasets for all 5 years were created using the same attributes as the training datasets in order to predict the vegetation type of each unknown object. Predictions for the test sets were then made in Weka using the training dataset for each year with the highest overall accuracy. The prediction results were visualized ArcMap for further analysis.

3.4 Classification accuracy assessment

The overall accuracy and the Kappa statistic were calculated to determine the best training dataset for each year. The overall accuracy represents the probability that the subject category will be classified correctly (Alberg et al. 2004). A high overall accuracy percentage is preferable, meaning that a higher amount of the data is correctly classified. The Kappa statistic determines the amount of agreement between the two data sets of interest, and the results are between 0 and 1. A high Kappa value indicates a strong agreement between the data categories (Sahoo and Kumar 2012).

4. Results

4.1 Maps of vegetation species distribution 2010 to 2018

The Random Forest classification algorithm was used to make predictions for each object. These objects were then symbolized with their respective classification using ArcMap Desktop 10.8 to show how the spatial distribution of each species changes from 2010 to 2018 (Figure 6). In the maps, centella (coral) and phragmites (teal) show areas of poor waterfowl food, and seedbox (light blue) and Walter's millet (pink) show areas of excellent waterfowl food. The most striking changes identified from the maps are those taking place near the open water (green). In year 2016, phragmites (teal) spreads into the open water areas (blue), which is then overtaken by Walter's millet in 2018 (pink). The distribution of seedbox (light blue) and Walter's millet (pink) typically appear to remain near the open water area (green) throughout the study years.

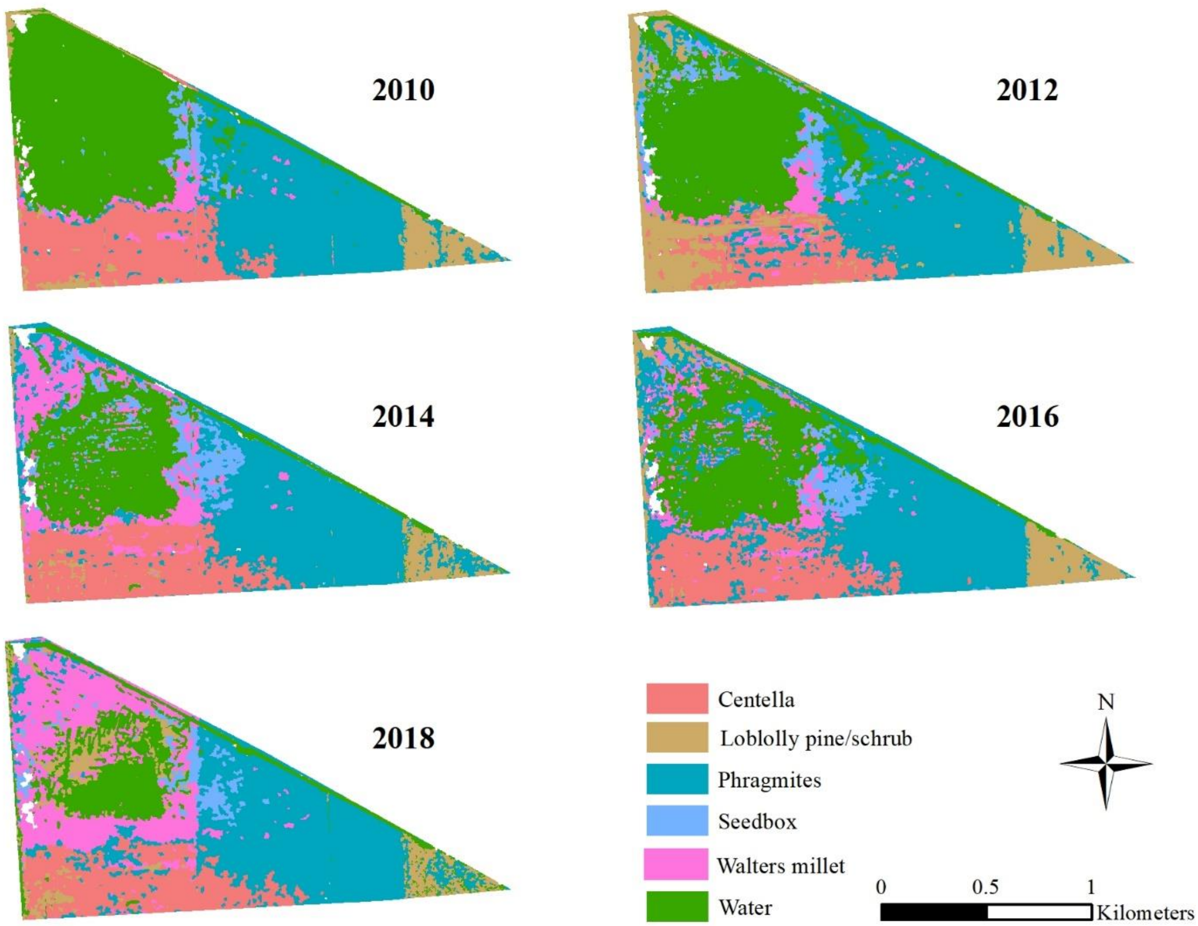

Figure 6. Weka output predictions mapped to show how the spatial distribution of each species changes from year 2010 to year 2018 within impoundment 4 of Mattamuskeet NWR.

In addition to visually observing vegetation change through the maps (Figure 6), numerical observations of the data displayed as graphs provide helpful insights (Figure 7). The total area of each classification category was also found for all years (Table 3). Phragmites occupied at least 30% of the total area for each year investigated, making it the dominant species in the impoundment. Phragmites is seen to occupy the highest percentage of the study site area (38%) in the year 2018 (Table 3). Centella is the second dominant species and steadily occupies at least 10% of the total study area. There is an inverse relationship between open water and Walter's millet, as a decrease in open water means an increase in Walter's millet (Figure 7).

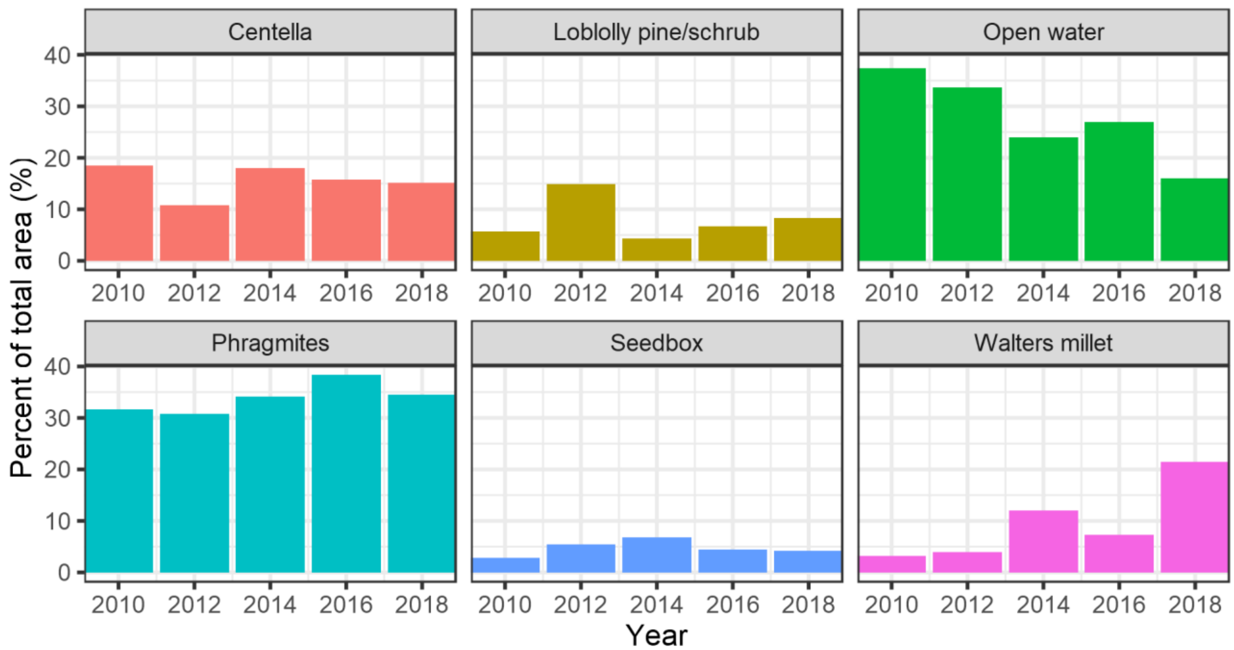

Figure 7. Each bar chart represents the percent of total area occupied by each vegetation species from 2010 - 2018. In the graphs, the y axis is the percent of total land area a species occupies each year (denoted by the x-axis).

Table 3. Area composition of each vegetation species by year.

Year	Vegetation species	Area (m²)	Percent of Study Area (%)
2010	Centella	318,996.3	18.4
	Phragmites	546,101.2	31.6
	Seedbox	47,980.77	2.8
	Walter's millet	53,803.99	3.1
	Loblolly pine/schrub	99,177.5	5.7
	Water	646,729	37.4
2012	Centella	185,270.6	10.7
	Phragmites	530,861.7	30.7
	Seedbox	93,014	5.4
	Walter's millet	67,792.71	3.9
	Loblolly pine/schrub	256,925.8	14.9
	Water	581,326.2	33.6
2014	Centella	309,578.9	17.9
	Phragmites	589,018.7	34.1
	Seedbox	117,330.3	6.8
	Walter's millet	208,338.6	12.0
	Loblolly pine/schrub	74,633.58	4.3
	Water	412,835.3	23.9
2016	Centella	271,121.9	15.7
	Phragmites	661,540.1	38.3
	Seedbox	75,562.44	4.4
	Walter's millet	126,308	7.3
	Loblolly pine/schrub	115,163.5	6.7
	Water	465,146.1	26.9
2018	Centella	261,610	15.1
	Phragmites	595,729	34.5
	Seedbox	70,343	4.1
	Walter's millet	370,373	21.4
	Loblolly pine/schrub	142,323	8.2
	Water	277,001	16.0

4.2 Classifier performance

The accuracies of the Random Forest algorithm for each year (2010, 2012, 2014, 2016, and 2018) are shown in Table 4. These overall accuracy and Kappa statistic were observed for each year in order to assess the accuracy of the Random Forest classifier for each year of the study period. The training datasets with the highest overall accuracy and Kappa statistic were chosen for each year. All study years had an overall accuracy ≤ 76% and a Kappa statistic ≤ 0.66 (Table 4). The lowest overall accuracy (76%) and Kappa statistic (0.66) used are associated with the year 2010), while data associated with the year 2016 had the highest overall accuracy (83.56%) and Kappa statistic (0.77) (Table 4). The training datasets included information such as mean brightness values for each imagery band, known species information collected in the field, and LiDAR statistics.

Table 4. Overall accuracy and Kappa statistic results for each imagery year.

Imagery Year	Overall accuracy (%)	Kappa statistic
2010	76.0	0.66
2012	80.80	0.73
2014	77.13	0.68
2016	83.56	0.77
2018	78.04	0.69

5. Discussion

This investigation examines the changes in vegetation distribution at Mattamuskeet NWR using a combination of OBIA and machine learning techniques. The results of this research provide insights about the effectiveness of current vegetation management techniques implemented at the Refuge, as well as the effectiveness of different machine learning classification algorithms in a managed wetland environment. Limitations of this research include the timing of when the aerial images were captured, as well as the availability of LiDAR data for the study years. The findings of this study show that the current management techniques are effective at containing the spread of invasive species but are not necessarily causing a decrease in the unwanted species. This suggests that a change in management techniques may be advised. Additionally, the Random Forest machine learning classification algorithm was found to be the best fit for this particular environment.

5.1 Species distribution and management efforts

The distribution of each vegetation species was expected to be influenced by known vegetation and management practices within the waterfowl impoundment. Vegetation management practices during 2010 to 2018 were effective at containing the distribution of phragmites, but they were not effective at reducing the total area the phragmites occupied in the impoundment (Figures 6, 7). A goal of attaining 50% cover of good waterfowl foods per bird impoundment was suggested in the most recent Habitat Management Plan for the Refuge (U.S. Fish and Wildlife Service 2018). Based on the results of this study, the combined percentage of Walter’s millet and seedbox did increase from 5.9% in 2010 to 25.5% in 2018 for the research impoundment. This may be due in part to the difference in water levels for 2010 and 2018, and the fact that the imagery for these years were collected in different months (June and July respectively). While the Refuge is typically finished with manually drawing down the water in the bird impoundments by early June, water is sometimes pumped back in to the impoundments during dry summers to help facilitate growth of the moist soil unit plants (U.S. Fish and Wildlife Service 2018). In either case, the year with the highest percentage of good waterfowl food (2018) of 25.5% is only halfway to reaching the goal of attaining 50% of the impoundment as good waterfowl vegetation. If future restoration efforts aim to target the removal of phragmites (North Carolina Coastal Federation 2018), then vegetation management strategies may need revised to accomplish that goal.

5.2 Limitations in availability and timing in data

Previous efforts to map the distribution of phragmites at Mattamukeet NWR's waterfowl impoundment 4 involved digitizing orthoimagery to estimate that 264 acres consisted of phragmites in 2012 (U.S. Fish and Wildlife Service 2018). In comparison, this study used the Random Forest machine learning algorithm with an overall accuracy of 80.8% where it was estimated that phragmites made up 131.2 acres (30% of the total area) of the same study site in 2012. Although the machine learning classification provided acceptable accuracy in comparison with previous mapping efforts, the results of our study are influenced by error due temporal discrepancies between the data.

It is best to have all imagery, reference data, and elevation collected the same month, year, and before any vegetation management strategies begin in the growing season. In this study, aerial imagery was not collected in the same month each year. Instead, it was collected within different months of the growing season (May to August). One approach the Refuge may want to consider is the use of small unmanned aerial systems to collect imagery consistent through the years with corresponding *insitu* measures. It is best to have *insitu* measures collected at the time of the aerial survey to ensure accurate species to spectral data matching, but this information is rarely available to compliment the imagery. It is also better to collect the *insitu* measures prior to any vegetation management. In this study, we collected *insitu* measures in August 2020 after vegetation management procedures were being implemented by the Refuge. The abundancy decreased through the years due to vegetation being different. The elevation data used in this study was also collected in the beginning of the growing season in 2014. Future research would benefit if the state of North Carolina increases its LiDAR data collection on the coastal plain to more than every ten years (the last LiDAR survey before year 2014 was 2001) to match the collection of the NAIP imagery.

5.3 Beneficial practices for future monitoring efforts

This study identifies a potential future workflow that could be implemented by restoration managers to better monitor changes in species distribution. First, annual data could be collected post spring bloom and pre vegetation management efforts. These data include sources from sUAS flights and *in situ* measures of a species' xyz location and height measured above ground level. sUAS surveys are an effective method to monitoring phragmites invasions and distinguishing colonies of phragmites from other species of vegetation (Kaneko *et al.* 2014). A camera sensor with RGB and NIR bands could be used on the sUAS to collect high resolution (e.g., 8 cm) imagery as well as a LiDAR sensor to collect dense xyz point clouds that estimate the ground (Cooper et al., 2021). OBIA or pixel-based methods could be used on the sUAS imagery based on the availability of software. *Insitu* measures would be matched to mean brightness values of all four sUAS imagery bands, and mean LiDAR ground point elevation, and mean vegetation height for each object or pixel. This matched dataset could then be used to train the machine learning classification and create species maps.

Many future studies may be conducted based on this research. Focusing specifically on Mattamuskeet NWR, these same techniques could be tested with different bird impoundments. Vegetation management techniques vary for the different bird impoundments on the Refuge. The timing and duration of manually flooding the impoundments also vary to provide a diversity of vegetation habitats (U.S. Fish and Wildlife Refuge 2018). Therefore, a future study might look at different bird impoundments to compare the effectiveness of each management plan with regards to the distribution and changes of desired and undesired plant species. In terms of the methods, future research could examine why certain species are more accurately identified than others in the Weka machine learning software, and how the accuracy can be improved for habitats with multiple dominant species. Perhaps this is due in part to the unique spectral signatures of each plant or the time of year that the aerial imagery is collected. The effectiveness of these methods could also be tested more broadly in habitats other than wetland environments.

6. Conclusion

Restoration managers interested in knowing where an invasive species is spreading, and which native species are most impacted by the invasion, need maps that show the change in individual species through time. Limited species maps exist for Mattamuskeet NWR, so this study used an Object-Based Image Analysis and machine learning approach to create accurate vegetation species distribution maps over a 10-year time period. The Random Forest machine learning classification algorithm was used for the machine learning process, with an overall accuracy of >76 for each training dataset. Focusing specifically on Mattamuskeet NWR, these same techniques could be tested with different bird impoundments. Management techniques vary for the different bird impoundments on the Refuge. Therefore, a future study might look at different bird impoundments to compare the effectiveness of each management plan with regards to the distribution and changes of desired and undesired plant species.

The results of this study demonstrate that the current management strategies for Mattamuskeet NWR bird impoundment 4 are effective at containing the area the invasive plant species occupy, but they are not effective at reducing the area. Additionally, the goal of reaching a total area of 50% good waterfowl vegetation was not reached during the study period. The year 2018 was closest to this goal, with a total of 25% of the area consisting of good waterfowl vegetation. This has implications for future restoration efforts that aim to remove invasive species, thus reducing its area of occupation. The species area maps created will allow for the assessment of how the species distribution has changed over time. These results also indicate a decline in water levels that lead to an increase in good waterfowl vegetation species (i.e., Walter's millet). It would be a valuable direction for future research to consider sUAS surveys in conjunction with *insitu* measures to monitor the change in species distribution through time to investigate the relationship of water levels, elevation, and vegetation species distribution at Mattamuskeet NWR.

References

Addink, E.A., Van Coillie, F., De Jong, S. 2012. Introduction to the GEOBIA 2010 special issue: from pixels to geographic objects in remote sensing image analysis. *International Journal of Applied Earth Observation and Geoinformation.* 15: 1-6

Akar, Ö., & Güngör, O. (2012). Classification of multispectral images using Random Forest algorithm. *Journal of Geodesy and Geoinformation*, *1*(2), 105-112

Alberg, A. J., Park, J. W., Hager, B. W., Brock, M. V., & Diener-West, M. 2004. The use of "overall accuracy" to evaluate the validity of screening or diagnostic tests. *Journal of general internal medicine*, *19*(5p1): 460-4

Antonio, T., & Zain, H. (2001). Plants Identification by Using Near-Infrared to Visible Spectrum Ratio Techniques. *IFAC Proceedings Volumes*, *34*(11), 306-310

Ash Meadows National Wildlife Refuge. 2006. Integrated pest management plan for Ash Meadows National Wildlife Refuge. https://ecos.fws.gov/ServCat/DownloadFile/56697?Reference=55846

AskUSDA, 17 July 2019, ask.usda.gov/s/article/What-is-the-horizontal-accuracy-of-the-National-Agriculture-Imagery-Program-imagery-and-how-is-it-verified

Beall, D. 1984. Brigantine Division - Marsh vegetation rehabilitation of Phragmites. USFWS, 8 https://soilcropandmore.info/crops/Grasses/Common_reed/phraaus.pdf

Berrar, D. 2019. Cross-validation. *Encyclopedia of bioinformatics and computational biology.* 1: 542-545

Bongiorno, S., Trautman, J., Steinke, T., Kawa-Raymond, S., Warner, D. 1984. A study of restoration in Pine Creek salt marsh, Fairfield, Connecticut. In Proceedings of the 11th Annual Conference on Wetlands Restoration and Creation, Hillsborough Community College Institute of Florida Studies (pp. 10-19) http://www.naturalareas.org/docs/66NAJ1404_285-294.pdf

Boone, J., Furbish, E., Turner, K. 1987. Control of Phragmites communis: Results of burning, cutting, and covering with plastic in a North Carolina Salt Marsh. CPSU Technical Report No 41 https://www.birdandhike.com/jlboone/papers/plants/_Phrag.htm.
Carlson, M., Kowalski, K., Wilcox, D. 2009. Promoting species establishment in a *Phragmites*-
dominated Great Lakes coastal wetland. *Natural Areas Journal* 29(3): 263-280

Cooper, H., Wasklewicz, T., Zhu, Z., Lewis, W., LeCompte, K., Heffentrager, M., Smaby, R., Brady, J., Howard, R. (2021). Evaluating the ability of multi-sensor techniques to capture topographic complexity. *Sensors*

Cross, D., Fleming, K. 1989. 13.4.12. Control of Phragmites or Common Reed. *Waterfowl Management Handbook*, 32 https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1031&context=icwdmwfm.

Courtwright, J. and Findlay, S. 2011. Effects of microtopography on hydrology, physicochemistry, and vegetation in a tidal swamp of the Hudson River. Wetlands, 31(2): 239-249

Dahl, T. and Stedman, S. 2013. Status and trends of wetlands in the coastal watersheds of the Conterminous United States 2004 to 2009. U.S. Department of the Interior, Fish and Wildlife Service and National Oceanic and Atmospheric Administration, National Marine Fisheries Service. (46 p.)

Devkota A; Jhan PK, 2010. Effects of different light levels on the growth traits and yield of Centella asiatica. Middle East Journal of Scientific Research, 5(4):226-230

Duke, J.A., 1955. Distribution and speciation of the genus Ludwigia in North Carolina. *Journal of the Elisha Mitchell Scientific Society*, 71(2): 255-269

Folmar, L., Sanders, H., Julin, A. 1979. Toxicity of the herbicide glyphosate and several of its formulations to fish and aquatic invertebrates. *Archives of Environmental Contamination and Toxicity* 8: 269-278

Frank, E., Hall, M., Witten, I. 2016. The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016

Friedl, MA, Brodley, CE and Strahler, AH. 1999. Maximizing land cover classification accuracies produced by decision tree at continental to global scales.. *IEEE Transactions on Geoscience and Remote Sensing*, 37: 969–977

Hay, G., Castilla, G. 2006. Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT). *Proc. 1st Int. Conf. OBIA* (pp. 4-5) https://www.isprs.org/PROCEEDINGS/XXXVI/4-C42/Papers/01_Opening%20Session/OBIA2006_Hay_Castilla.pdf

Hazelton, E., Mozdzer, T., Burdick, D., Kettenring, K., Whigham, D. 2014. *Phragmites australis* management in the United States: 40 years of methods and outcomes. *AoB PLANTS* 6

Heller, (Pursh) A., and P.W. Michael. "Echinochloa Walteri." *Echinochloa Walteri - FNA,* beta.floranorthamerica.org/Echinochloa_walteri.

Hellings, S., Gallagher, J. 1992. The effects of salinity and flooding on *Phragmites australis*. *Journal of Applied Ecology* 29(1): 41-49

Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D. 2013. Change detection from remotely sensed images: from pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing.* 80: 91-106

Jensen, J. 2015. Introductory digital image processing: a remote sensing perspective: a remote sensing perspective, 4th edition. Prentice-Hall Inc. Upper Saddle River, NJ

Juel, A., Groom, G.B., Svenning, J.C. and Ejrnaes, R., 2015. Spatial application of Random Forest models for fine-scale coastal vegetation classification using object based analysis Of aerial orthophoto and DEM data. *International Journal of Applied Earth Observation And Geoinformation*, 42: 106-114

Kaneko, K., Nohara, S. 2014. Review of effective vegetation mapping using the UAV (Unmanned Aerial Vehicle) Method. *Journal of Geographic Information System* 6: 733-742

Luscier, J., Thompson, W., Wilson, J., Gorham, B. 2006. Using digital photographs and object-based image analysis to estimate percent ground cover in vegetation plots. *Frontiers in Ecology and the Environment* 4(8): 408-413

Marks, M., Lapin, B., Randall, J. 1993. Element stewardship abstract for Phragmites australis. *The Nature Conservancy, Arlington* https://soilcropandmore.info/crops/Grasses/Common_reed/phraaus.pdf

Marks, M., Lapin, B., Randall, J. 1994. *Phragmites australis (P. communis):* Threats, Management, and Monitoring. *Natural Areas Journal* 14(4): 285-294

Martin, L. and Blossey, B. 2013. The runaway weed: Costs and failures of *Phragmites australis* management in the USA. *Estuaries and Coasts* 36: 626-632

Meyerson, L., Saltonstall, K., Windham, L., Kiviat, E., Findlay, S. 2000. A comparison of *Phragmites australis* in freshwater and brackish marsh environments in North America. *Wetlands Ecology and Management* 8: 89-103

Mishra, N. and Crews, K. 2014. Mapping vegetation morphology types in a dry savanna ecosystem: Integrating hierarchical object-based image analysis with Random Forest. *International Journal of Remote Sensing*, *35*(3), pp.1175-1198

"Monitoring Techniques." *Great Lakes Phragmites Collaborative*, www.greatlakesphragmites.net/management/monitoring-programs/.

Muchoney, D, Borak, J, Chi, H, Friedl, M, Gopal, S, Hodges, J, Morrow, N and Strahler, A. 2000. Application of MODIS global supervised classification model to vegetation and land cover mapping of Central America.. *International Journal of Remote Sensing*, 21: 1115–1138

North Carolina Coastal Federation, and US Fish and Wildlife Service. 2018. *Lake Mattamuskeet Watershed Restoration Plan*. https://files.nc.gov/ncdeq/Coastal%20Management/documents/PDF/Land%20Use%2Plans/pmg-projects/LMWRP_Final_Appendices_Web.pdf

Osterbrock, A. J. 1984. Phragmites australis. The problem and potential solutions. Ohio Field Office, Stewardship. 8 pp. https://soilcropandmore.info/crops/Grasses/Common_reed/phraaus.pdf

Pande-Chhetri, R., Abd-Elrahman, A., Liu, T., Morton, J., Wilhelm, V. 2017. Object-based classification of wetland vegetation using very high-resolution unmanned air system imagery. *European Journal of Remote Sensing.* 50(1): 564-576

Rojas, O., Zamorano, M., Saez, K., Rojas, C., Vega, C., Arriagada, L. and Basnou, C., 2017. Social perception of ecosystem services in a coastal wetland post-earthquake: A case study in Chile. *Sustainability*, 9(11): 1983

Roman, C., Niering, W., Warren, R. 1984. Salt marsh vegetation change in response to tidal restriction. *Environmental Management* 8(2):141-150

Sahoo, G., & Kumar, Y. (2012). Analysis of parametric & non parametric classifiers for Classification technique using WEKA. *International Journal of Information Technology and Computer Science (IJITCS)*, *4*(7): 43

Schneider, K. 1990. Coordinator, New York Heritage Program. Telephone conversation with Beth Lapin, October 1990. http://www.naturalareas.org/docs/66NAJ1404_285-294.pdf

Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A., Dzeroski, S. 2010. Estimating vegetation height and canopy cover from remotely sensed data with machine learning. *Ecological Informatics* 5(4):256-266

Styers, D., Moskal, L., Richardson, J., Halabisky, M. 2014. Evaluation of the contribution of LiDAR data and postclassification procedures to object-based classification accuracy. *Journal of Applied Remote Sensing.* 8(1):083529

Suchrow, S., Jensen, K. (2010). Plant Species Responses to an Elevational Gradient in German North Sea Salt Marshes. Wetlands, 30, 735-746

Thompson, D., Shay, J. 1985. The effects of fire on *Phragmites australis* in the Delta Marsh, Manitoba. *Canadian Journal of Botany* 63(10): 1864-1869

Tommervik, H., Hogda, K.A., Solheim, I. 2003. Monitoring vegetation changes in Pasvik (Norway) and Pechenga in Kola Peninsula (Russia) using multitemporal Landsat MSS/TM data. *Remote Sensing of Environment.* 85: 370-388

Trimble. 2017. SP80 GNSS Receiver. Available online at: http://trl.trimble.com/docushare/dsweb/Get/Document-844535/SG-SP80-Br-v2.pdf

Trueman, M., Hobbs, R.J., Van Niel, K. 2013. Interdisciplinary historical vegetation mapping for ecological restoration in Galapagos. *Landscape Ecology* 28: 519-532

Tscharntke, T. 1988. Variability of the grass *Phragmites australis* in relation to the behavior and mortality of the gall-inducing midge *Giraudiella inclusa* (Diptera, Cecidomyiidae). *Oecologia* 76: 504-512

U.S. Department of the Interior, Fish and Wildlife Service. 2008. Draft comprehensive conservation plan and environmental assessment. Mattamuskeet National Wildlife Refuge. Hyde County, North Carolina https://www.fws.gov/southeast/planning/PDFdocuments/Mattamuskeet/Draft%20Mattamuskeet%20CCP%20edited.pdf

U.S. Fish and Wildlife Service. 2007. Bear River Migratory Bird Refuge Phragmites Control Plan. https://www.fws.gov/uploadedFiles/Phragmites%20Control%20Plan.pdf

U. S. Fish and Wildlife Service. 2018. Mattamuskeet National Wildlife Refuge Habitat Management Plan. U.S. Fish and Wildlife Service, Swan Quarter, North Carolina. 86 pp. https://files.nc.gov/ncdeq/Coastal%20Management/documents/PDF/Land%20Use%20Plans/pmg-projects/LMWRP_Final_Appendices_Web.pdf

Xie, Z., Xu, X. and Yan, L., 2010. Analyzing qualitative and quantitative changes in coastal wetland associated to the effects of natural and anthropogenic factors in a part of Tianjin, China. *Estuarine, Coastal and Shelf Science*, 86(3): 379-386

Zhang, C., Xie, Z. 2013. Object-based vegetation mapping in the Kissimmee River Watershed using HyMap data and machine learning techniques. *Wetlands* 33: 233-244

Zhang, C., Smith, M., Fang, C. 2017. Applying time series Landsat data for vegetation change analysis in the Florida Everglades Water Conservation Area 2A during 1996-2016. *International Journal of Applied Earth Observation and Geoinformation* 57: 214-223

Zurqani, H. A., Post, C. J., Mikhailova, E. A., Cope, M. P., Allen, J. S., & Lytle, B. A. 2020. Evaluating the integrity of forested riparian buffers over a large area using LiDAR data and Google Earth Engine. *Scientific Reports* 10(1): 1-16