**Remote Sensing and Mapping of Plant Communities for the Preservation of Natural Systems**

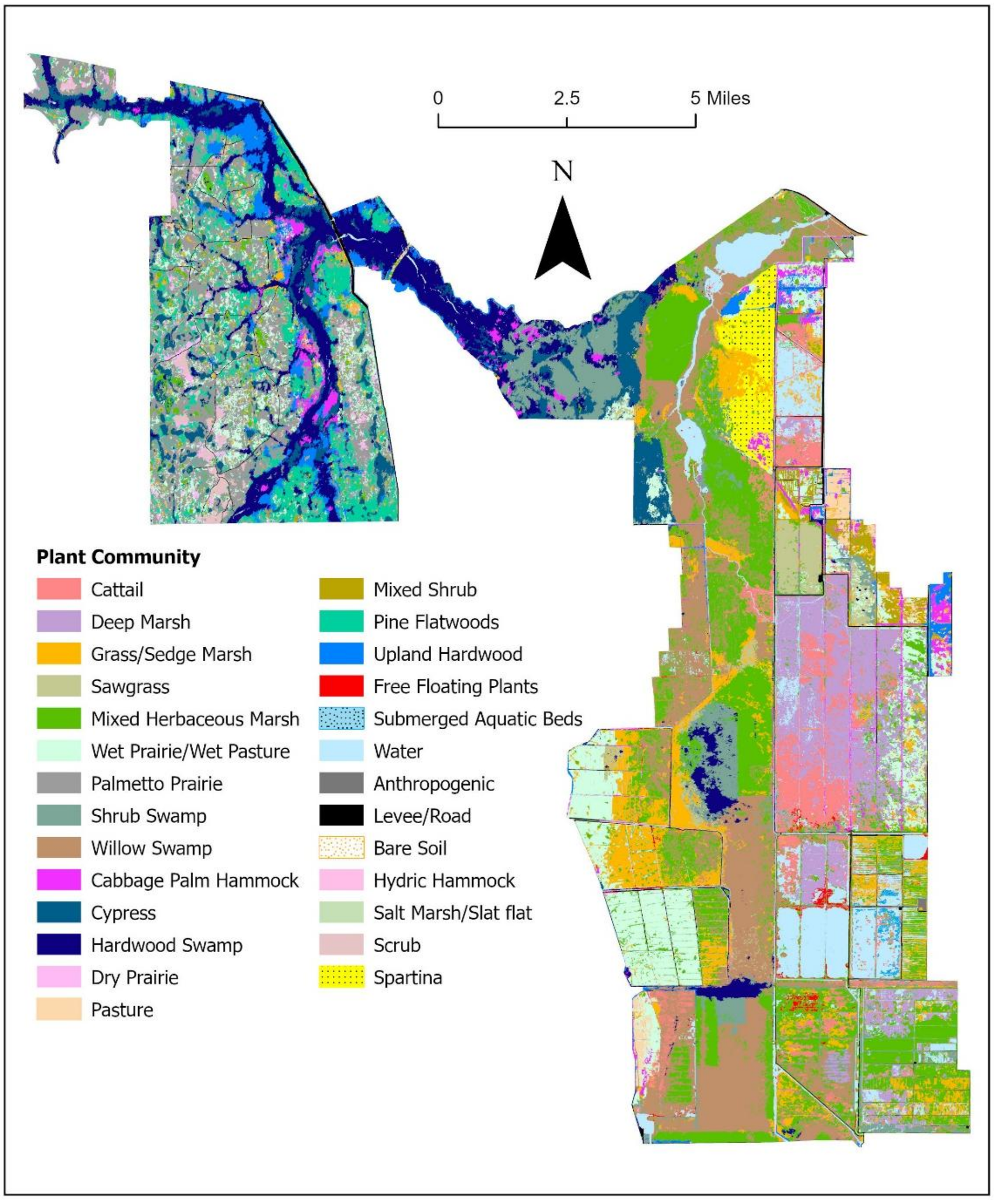
Year 2 (2023-2024) Report (Contract 38076)

Submitted to: St. Johns River Water Management District

Caiyun Zhang, PhD, Professor

Department of Geosciences, Florida Atlantic University

777 Glades Road, Boca Raton, FL 33431, Tel: 561-297-2648; czhang3@fau.edu




Map of wetland communities in 2023 with a coverage of 88,016 acres generated from an ensemble of machine learning classifications and WorldView-2 satellite imagery.

## Contents

## 1. Project background

The St. Johns River Water Management District (District) manages more than 315,000 acres of wetlands which are mapped on a regular basis to monitor the response of natural systems to environmental changes and management activities. In the past, the District collected digital aerial imagery with four spectral bands (Red, Green, Blue, and Near-Infrared (NIR)) to generate maps through a manual interpretation procedure. It takes months to collect aerial imagery, including time spent on flight planning, sensor selection, and final aerial imagery generation and rectification. After the acquisition of aerial imagery, it takes several months to generate a map product through the manual interpretation procedure. This conventional vegetation mapping procedure is time-consuming and labor-intensive and can result in vegetation maps that are inaccurate due to the length of time between imagery acquisition and ground truthing of the mapped environments. Updating the map inventories is challenging due to the high cost and lengthy time of the mapping procedure over such a large area. The previous maps for the current project areas were generated approximately seven years ago.

Satellite remote sensing has been frequently cited as a cost-effective and labor-saving technique for monitoring and mapping wetlands. The operational WorldView-2/3 (WV-2/3) satellites are the most technologically advanced high-resolution optical sensors, collecting 46-centimeter (cm) panchromatic and 1.85-meter (m) multispectral imagery. WV-2/3 sensors have a comparable spatial resolution to aerial imagery, but they add four extra spectral bands: coastal, yellow, red edge and an additional NIR band to the traditional four (Red, Green, Blue, NIR) bands. The revisitation frequency of WV-2/3 (~2 days), in conjunction with their large spatial coverage, make spaceborne sensors more appealing for monitoring wetland response to changing environments than using airborne platforms, especially over large areas.

Contemporary machine learning-based digital imagery classification techniques have been widely used as an alternative to manual interpretation, which is more challenging when delineating natural landscapes because human-generated plant community boundaries are subjectively drawn. In contrast, computer-based digital classification methods are more objective in delineating boundaries based on imagery-based spectral signatures. Integrating the computer-based semi-automated/automated machine learning techniques with WV-2/3 satellite imagery can largely accelerate the wetland mapping process, making it more effective for informing decision making in wetland management than the traditional manual interpretation of aerial imagery.

The project Principal Investigator (PI), Dr. Caiyun Zhang, at Florida Atlantic University (FAU), has over 15 years of experience in satellite remote sensing and digital image analysis. We (Dr. Zhang's team at FAU) previously successfully completed a similar project for the District during 2020-2022 entitled "Development of Automated Methods for Using Satellite Imagery to Monitor Changes in Vegetative Communities" (Contract 35547-1) mapping plant communities over 55,955 acres including the Lake Apopka North Shore, Ocklawaha Prairie Restoration Area, Emerald Marsh Conservation Area, Moccasin Island Restoration Area, and Buck Lake Conservation Area. The developed methods were effective and efficient across these wetland areas. In the current contract project (Contract 38076), we further improved our machine learning procedures, and generated map products covering 81,895 acres of natural wetland areas during project year 1 (2022-2023). In this project year 2 (2023-2024) annual report, we include information on the data we received from the District, data we collected in the field, methods we updated in generating year 2 map products, map products we created, issues we encountered and how they were solved, and suggestions to the District to improve the mapping process, such as timely field visits, supplying accurate project boundaries, and other ancillary data required in the project.

## 2. Project areas and data

*2.1 Project areas*

In project year 2, we mapped nine areas: Bull Creek Wildlife Management Area /Jane Green Flowage Easement (BULLJG - 27,887 acres), St. Johns Marsh Conservation Area (SJMCA – 23,247 acres), C1 Detention Area (C1DETEN – 2,346 acres), C-1 Retention Area / Sawgrass Lake Water Management Area (C1SLWMA – 3,336 acres), Deseret Ranch (DESERET – 6,039 acres), Sartori West (SARTORI – 1,928 acres); Sixmile Creek Marsh Restoration Area (SCMRA -2,729 acres), Three Forks Marsh Conservation Area (TFMCA - 13,716 acres), and C54 Retention Area/Broadmoor Marsh Restoration Area (C54BMRA - 6,788). These areas are displayed as a color infrared composite of the WV-2 imagery collected on March 1 and 7, 2023, in Figure 1.

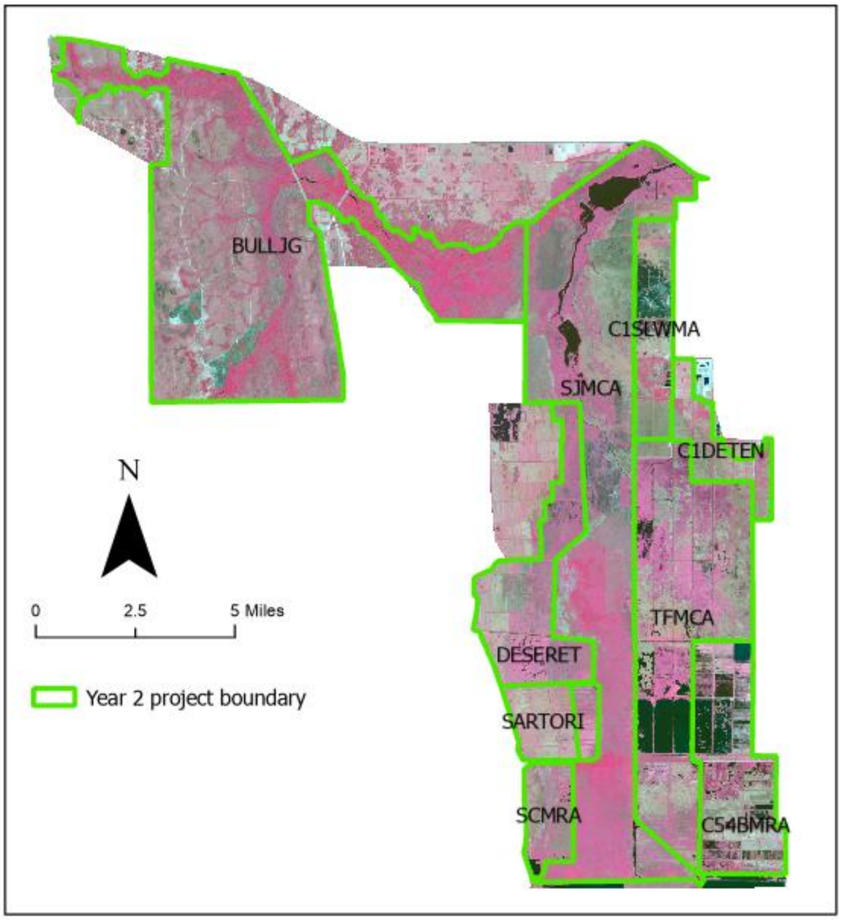



Figure 1 Project areas displayed in a color infrared composite of WorldView-2 imagery collected in March 2023. The imagery is the major input for generating the maps in the project.

*2.2 Data provided by the District*

The District delivered a hard drive with satellite imagery collected in March 2023. Other datasets used in this project year were delivered in 2022 including historical aerial photography collected in 2016, lidar Digital Elevation Models (DEMs), historical vegetation maps created in 2016 by manual interpretation of the aerial photography, road map data, project boundary, elevation contour maps, and a plant community cross walk file to document the modification of plant community names used in the historical maps and the current plant community classification system. These data were provided in different formats including ArcGIS vector Shapefile, raster data (TIF), ArcGIS lpk (map package format), MS Word and Excel. In total, 85 GB satellite imagery data were received from the District during this period including 16-bit, 8-band, 0.5-m WV-2/3 imagery product and mosaic of 8-bit, 4-band, 0.5-m imagery product. Imagery collected on March 1 and 7 were merged for mapping the entire project area to patch cloud cover. A screenshot of the data is shown in Figure 2.

| | | | |
|---|---|---|---|
| Mosaic_PS_8bit | 6/1/2023 10:33 AM | File folder | |
| 23MAR01161847-PS-050146466010_01_P001_8Band_16bit.prj | 5/18/2023 4:44 PM | PRJ File | 1 KB |
| 23MAR01161847-PS-050146466010_01_P001_8Band_16bit.tfw | 5/18/2023 4:44 PM | TFW File | 1 KB |
| 23MAR01161847-PS-050146466010_01_P001_8Band_16bit.tif | 5/18/2023 4:47 PM | TIF File | 13,051,228 ... |
| 23MAR01161847-PS-050146466010_01_P001_8Band_16bit.tif.aux.xml | 3/29/2024 2:48 PM | Microsoft Edge H... | 18 KB |
| 23MAR01161847-PS-050146466010_01_P001_8Band_16bit.tif.xml | 6/2/2023 8:10 AM | Microsoft Edge H... | 7 KB |
| 23MAR01161856-PS-050146466010_01_P006_8Band_16bit.prj | 5/18/2023 4:47 PM | PRJ File | 1 KB |
| 23MAR01161856-PS-050146466010_01_P006_8Band_16bit.tfw | 5/18/2023 4:47 PM | TFW File | 1 KB |
| 23MAR01161856-PS-050146466010_01_P006_8Band_16bit.tif | 5/18/2023 4:53 PM | TIF File | 25,350,687 ... |
| 23MAR01161856-PS-050146466010_01_P006_8Band_16bit.tif.xml | 6/2/2023 8:11 AM | Microsoft Edge H... | 9 KB |
| 23MAR07155827-PS-050146466010_01_P004_8Band_16bit.prj | 5/18/2023 4:53 PM | PRJ File | 1 KB |
| 23MAR07155827-PS-050146466010_01_P004_8Band_16bit.tfw | 5/18/2023 4:53 PM | TFW File | 1 KB |
| 23MAR07155827-PS-050146466010_01_P004_8Band_16bit.tif | 5/18/2023 5:08 PM | TIF File | 29,754,187 ... |
| 23MAR07155827-PS-050146466010_01_P004_8Band_16bit.tif.aux.xml | 6/17/2023 12:38 PM | Microsoft Edge H... | 18 KB |
| 23MAR07155838-PS-050146466010_01_P005_8Band_16bit.prj | 5/18/2023 5:08 PM | PRJ File | 1 KB |
| 23MAR07155838-PS-050146466010_01_P005_8Band_16bit.tfw | 5/18/2023 5:08 PM | TFW File | 1 KB |
| 23MAR07155838-PS-050146466010_01_P005_8Band_16bit.tif | 5/18/2023 5:10 PM | TIF File | 9,169,114 KB |
| 23MAR07155838-PS-050146466010_01_P005_8Band_16bit.tif.aux.xml | 6/15/2023 2:11 PM | Microsoft Edge H... | 18 KB |
| Readme.txt | 4/30/2024 12:14 PM | Text Document | 1 KB |

Figure 2 Satellite imagery products received from the District for the 2023 mapping.

### *2.3 Newly collected data*

To teach a machine learning model to classify different plant communities, training samples from each community are required. This process in remote sensing is known as supervised learning. Training samples can be collected from visual interpretation of high-resolution imagery, as well as from field visits. We combined both methods for training sample collection in this project. These samples are used to develop the spectral signature keys in the classification. The samples can be split into training samples and testing samples to calibrate and then validate the mapping models in the project. The plant communities, corresponding community type, and abbreviations that we mapped for are listed in Table 1. Their definitions are provided in Appendix 1.

Table 1 Plant community mapped in this project.

| Plant Community | Abbreviation | Community Type | Abbreviation |
|---|---|---|---|
| Anthropogenic | AN | Other | O |
| Bare Soil | BS | Other | O |
| Deep Marsh | DM | Herbaceous Wetland | HW |
| Dry Prairie | DP | Herbaceous Upland | HU |
| Hydric Hammock | HH | Forested Wetland | FW |
| Mixed Herbaceous Marsh | HM | Herbaceous Wetland | HW |
| Hardwood Swamp | HS | Forested Wetland | FW |
| Levee/Road | LR | Other | O |
| Water | OW | Water | W |
| Pasture | PA | Herbaceous Upland | HU |
| Shrub Swamp | SS | Shrub Wetland | SW |
| Upland Hardwood | UH | Forested Upland | FU |
| Wet Prairie/Wet Pasture | WP | Herbaceous Wetland | HW |
| Submerged Aquatic Beds | SAB | Water | W |
| Palmetto Prairie | PP | Shrub Wetland | SU |
| Pine Flatwoods | PF | Forested Upland | FU |
| Cypress | CY | Forested Wetland | FW |
| Forested Flatwood Depressions | FD | Forested Wetland | FW |
| Cabbage Palm Hammock | CP | Forested Wetland | FW |
| Cattail | CT | Herbaceous Wetland | HW |
| Grass/Sedge Marsh | GS | Herbaceous Wetland | HW |
| Willow Swamp | WS | Shrub Wetland | SW |
| Spartina | SP | Herbaceous Wetland | HW |
| Sawgrass | SG | Herbaceous Wetland | HW |
| Phragmites | PH | Herbaceous Wetland | HW |
| Pine Plantation | PI | Forested Upland | FU |
| Bayhead/gall | BG | Forested Wetland | FW |
| Sandhill | SH | Forested Upland | FU |
| Salt Marsh/Slat flat | SM | Herbaceous Wetland | HW |
| Free Floating Plants | FF | Water | W |
| Mixed Shrub | MS | Shrub Upland | SU |

We made a detailed field plan, and collected training samples during 25-29 June 2023 with District personnel, including Dr. Dianne Hall (Dr. Hall, Project Manager), Christy Akers, and Jonny Baker. The group considered site accessibility by the District and FAU’s truck, vegetation communities, and the logistics of planned field time schedule (time and cost budget). The District provided a 3-seat airboat with Christy Akers as the driver. We did not collect drone data this year due to constraints from FAU and the State of Florida.

In total, we collected 3429 field points and took 261 pictures during the visit to document the signatures of each vegetation community appeared in project areas (Figure 3). Figures 4-12 show some of the field pictures.

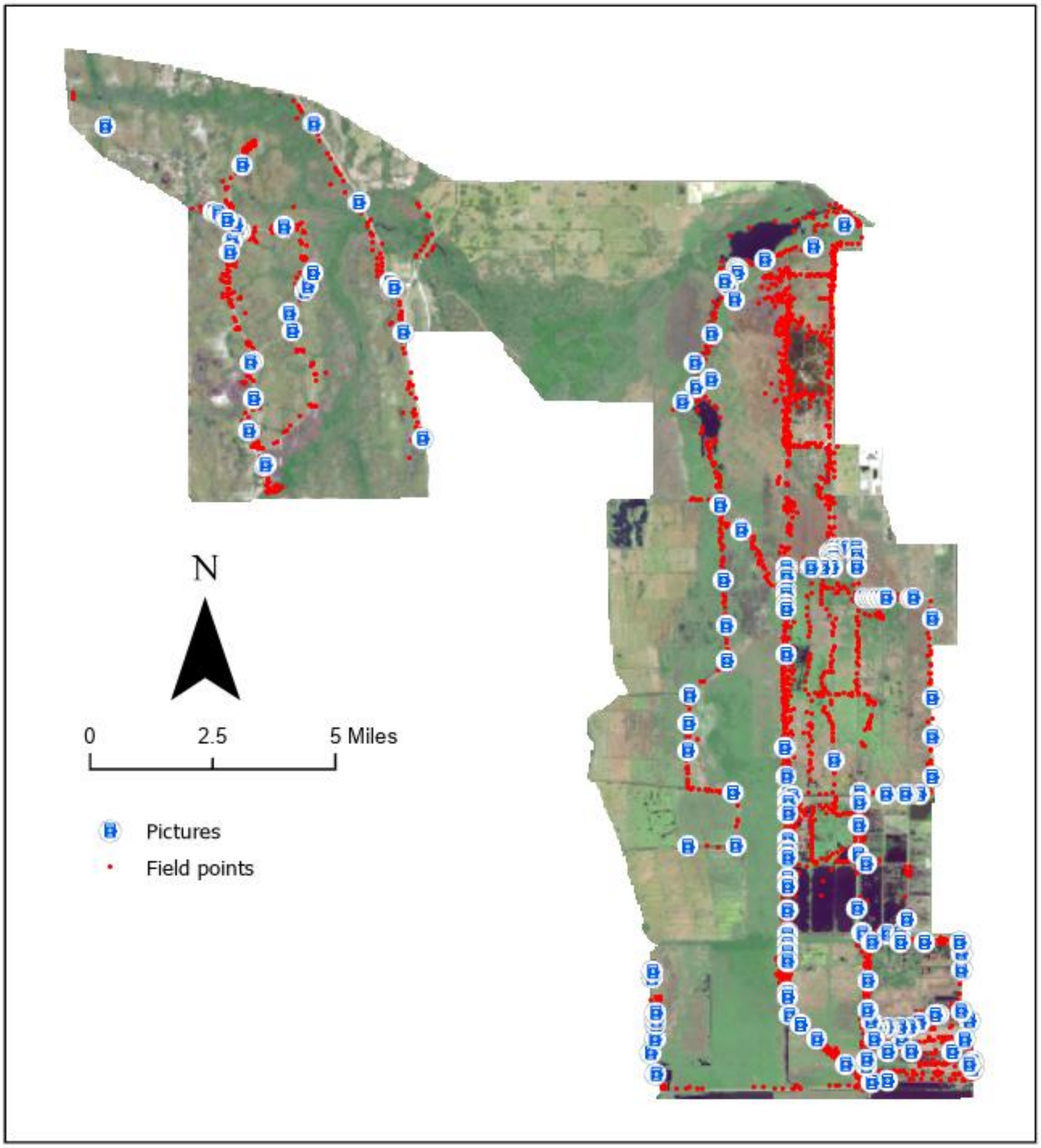



Figure 3 Field point data and pictures collected in project year 2 during June 25-29, 2023.

Figure 4 Palmetto Prairie with sparse Pine Flatwoods and Cypress at BULLJG.

Figure 5 Grass Sedge Marsh invaded by Willow Swamp at SJMCA.

Figure 6 Shrub Swamp, Water, Free Floating plants, and Cattail, Willow Swamp at C54BMRA.

Figure 7 Hydric Hammock and Water at DESERET.

Figure 8 Pasture with Upland Hardwood mixed with Cabbage Palm Hammock at SCMRA.

Figure 9 Mixed Herbaceous marsh with Shrub Swamp at TFMCA.

Figure 10 Deep Marsh, Water, and Shrub Swamp at SJMCA.

Figure 11 FAU crew, Dr. Dianne Hall (Left 3) and Christy Akers (left 1, airboat driver) in the field.


Figure 12 FAU crew at the District's Palm Bay office on June 29, 2023. (Left to right: Caiyun Zhang, Sandip Rijal, Madan Thapa Chhetri, Fiona Benzi, Rabindra Parajuli).

## 3. Methodology

### *3.1 Object-based machine learning ensemble classification*

We have developed an object-based machine learning ensemble protocol to map wetland communities using satellite imagery and lidar Digital Elevation Models (DEMs) provided by the District. This approach proved effective and achieved high accuracy in mapping a large number of diverse wetland vegetation communities in our first contract with the District. It combines modern machine learning, object-based image analysis (OBIA), and ensemble analysis techniques in the mapping procedure. Our previous studies have proven that object-based vegetation mapping is more effective and accurate than traditional pixel-based mapping methods in wetland environments, especially when high resolution imagery like WV-2 data is used, and application of an ensemble analysis which combines the outputs of multiple classifiers makes the final classification more robust (Zhang et al., 2016, 2018; Zhang and Xie, 2014; Zhang, 2014).

Multiple steps are required in the methodology. To conduct OBIA mapping, we first applied the multiresolution segmentation algorithm to the 0.3-m and 8-band pansharpened WV-2 imagery products in eCognition 10 with the scale set to 150. We reduced the scale parameters from 300 (used in project year 1) to 150 to reduce the community mixture issues we found in project year 1. After segmentation, we fused lidar DEM features including mean and standard deviation of elevation with image features to be used for classification. We spatially matched the ground-truth/reference data with the fused dataset, leading to a matched dataset for classification model development. Contemporary machine learning classifiers including Support Vector Machine (SVM), Random Forest (RF), and Artificial Neural Network (ANN) were selected and used to classify the matched dataset, and the classification results were assessed using error matrix techniques and the *k*-fold cross validation approach. SVM is a supervised classifier that aims to find a hyperplane that can separate the input dataset into a discrete predefined number of classes in a fashion consistent with the training samples. RF is a decision-tree-based ensemble classifier, and is a sophisticated machine learning model. ANN is an important technique in image classification that functions by simulating the method by which human brains recognize patterns. All of these classifiers have been widely applied in remote sensing classification with great success. However, each classifier has its pros and cons, and thus ensemble analysis is more valuable than the application of a single classifier to make the plant community classification more robust. We applied an ensemble approach to integrate the outputs of SVM, RF, and ANN for the final classification map generation. In the ensemble procedure, if a majority of the three classifiers determine the segment is the same plant community, then the segment will be assigned to that plant community. If none of the classifiers agree on the plant community for a segment, the segment will be assigned the plant community that has the highest individual accuracy among the three classifiers. We collected both field-based and imagery-based reference samples to train and test the classification models.

*3.2 Selection of scale parameter in image segmentation*

In our first contract with the District, we mapped multiple wetland areas on District land and set the scale parameter in the multiresolution segmentation process at 150. The setting of this parameter is important. A smaller value will lead to a greater number of segments/patches and may lead to "salt-and-pepper" effects in the final maps like the pixel-based mapping method. In addition, the large number of small segments will greatly increase the dataset size requiring more computer memory, resulting in difficulties in further data processing. The small segments are also not reasonable for management activities compared to the minimum mapping unit used in the historical maps. In the current contract, we experimented with different scale parameters (e.g., 150, 300, 500) in a small test area (500 acres) in project year 1 and decided to set the scale to 300. However, when this scale value was applied to the big project area, the segmentation results led to many mixed communities within one segment. For project year 2, we determined to set the scale back to 150 due to the heterogeneity of landscapes.

*3.3 Acceleration of image segmentation using eCognition server*

A big challenge using the multiresolution image segmentation in eCognition Developer 10 for object-based wetland mapping is the time cost. There are no comparable software packages to the state-of-the-art OBIA of eCognition Developer which has been secured by the District. Generating the image objects and exporting relevant image-based features may take over 10 hours for an area of 10,000 acres. Interruption of this procedure will lead to a complete loss of the generated segments and exported features, and thus the procedure needs to restart. In this contract, we found a solution to reduce the segmentation time and save the segments if the procedure is interrupted. We applied the eCognition Server tiling and batch processing functions to solve this issue. Applying the multiple server function is similar to parallel computation using multicore processing, and the batching function segments image task by task. The segmentation and export function can be saved as a batch rule set to be applied in the batch processing, then the server will process and apply the batch rule to analyze the image. Large images can be split into smaller tiles and the server will process the data tile by tile. Though this current contract does not require tutorials for the mapping procedure as in the previous contract, we can provide a detailed manual for the District to use the eCognition Server batching and tiling function using the WV-2 imagery as an example. This approach improved the segmentation procedure and greatly reduced the time cost in the project.

*3.4 Modification of project boundary*

The original project boundary provided by the District was not correct. There were gaps between project areas. This was revealed after dissolving the original project boundary ArcGIS shapefile. The topology of the original boundary file needs to fix first to avoid holes/gaps. The north boundary of the TFMCA was not correct, leading to two incorrect boundaries for TFMCA and C1SLWMA. The incorrect dividing line between TFMCA and C1SLWMA was found after we finished mapping these two areas. To solve this issue, maps of two areas were merged into one big area first, and then clipped out using the updated/correct polygon boundary file. Edge issues may be present due to the incorrect boundary file. A correct delineation of the project boundary is important because it defines where we should collect reference samples as well as other data processing procedures. Thus, we suggest the District refine and improve each project area boundary and check the corresponding imagery before the mapping project starts.

*3.5 Mapping of anthropogenic areas and levees/roads*

Both anthropogenic areas and levees/roads are challenging to classify using automated digital image analysis. Anthropogenic areas can include boat ramps, buildings, crops, cars, etc., while levees/roads can be dirt roads, paved roads, or grassed levees. This causes computer confusion due to the diversity of spectral signatures within these two classes and between these classes and the other plant community classes. For example, grassed levees can be confused with pasture (PA) areas, and boat ramp areas can be confused with paved roads. The District has an existing levee/road dataset as a line feature, and the anthropogenic areas are few. Both levees/roads and anthropogenic features can be clearly identified on the imagery, and thus we manually mapped these two classes. For the levee/road features, we found a couple of them misaligned with roads shown on the imagery, especially over the BULLJG. We refined/updated the road based upon the imagery, and then measured the width of different roads on the imagery. We finally applied a buffer with different width (50 feet, 100 feet, 150 feet, and 200 feet) to generate a polygon feature of the levee/road map layer. For the anthropogenic areas, we manually digitized anthropogenic features if they were not completely covered by the levee/road map layer. The wetland community map was erased for areas covered by levee/road or anthropogenic classes first and then merged with the levee/road and anthropogenic layers to generate the final map product. Therefore, the error matrix of each project area did not include these two classes.

*3.6 Automation of map product generation*

In our previous contract, we developed a detailed tutorial to implement the developed object-based machine learning mapping procedure using ArcMap, eCognition, open-source WEKA, and alternative *R* packages. In this contract, we increased the automation of the major steps to speed up the entire mapping procedure. We designed data preparation modules to generate inputs for our *R* script and created map product generation modules using ArcGIS model builder. We implemented the entire machine learning and classification process using *R* script to reduce the manual input procedure of WEKA. In this way, the inputs and outputs are consistent in data format, and product attributes, making the process more efficient. We updated the modules in ArcGIS Pro and updated our R code developed in project year 1 to make the entire mapping procedure more consistent and automated.

*3.7 Post-classification refinement*

The automated digital procedure generated encouraging results for delineating communities for each project area. However, misidentifications could appear for classes within a community type, which is the upper level of the plant community used by the District's classification system (Appendix 1). For example, within the Forest Wetland community type, Cypress (CY) can be confused with Hardwood Swamp (HS), and within the Shrub Wetland community type, Shrub Swamp (SS) can be confused with Willow Swamp (WS) because their definitions are not mutually exclusive. Cypress trees occur in HS and willows occur in SS. The definition of the communities is dependent on the percent cover of cypress or willow, respectively, which is open to interpretation by the personnel selecting the training samples.

Another issue is the misclassification of shadows which often appear beside tall shrubs or forests. Because we did not include shadow in the classification system and it appears in fine resolution imagery such as WV-2 products used in the project, we need to correct these misclassifications. Shadows are often misclassified as water because they both appear dark on an image in a natural color composite. We manually assigned the shadows to shrubs or forests adjacent to them. Small patches such as shrubs that appeared in a big patch such as Upland Hardwood (UH) were suspected of being misclassified. By examining the historical map to gain the domain knowledge and the current imagery, these suspicious misclassifications were manually revised. The manual post-classification procedure is time consuming but necessary to improve the integrity of the map products. However, this refinement is not captured in the accuracy assessment error matrices. Consequently, the accuracy of the maps is greater than the error matrices indicate.

Refinement of a large volume of the classified segments is challenging in ArcGIS Pro. In project year 2, the classified segment shapefile was dissolved first to reduce the data volume, then the misclassified/suspicious spots were identified using a newly created point feature layer. This point feature layer will be corrected and then spatially joined to the classified segment shapefile to correct the misclassified polygon segments. This point feature file is very small and easy to handle in ArcGIS Pro. This is a successful solution when refining a large volume of classified segment shapefile.

## 4. Results

### *4.1 Three testing plots*

As required by the District, the developed procedures need to be tested over selected plots to meet the accuracy requirement before going further to map the entire project area. The selected three 500-acre plots were within BULLJG, SJMCA, and TFMCA with a diverse plant community (in total of 22 communities). The mapping results are provided in Figure 13. In general, the mapping products, and accuracy was acceptable after discussions with Dr. Hall, but also raised concerns for some communities such as mixed herbaceous marsh (HM), CY, and grass/sedge (GS) for each individual testing area. This was mainly caused by the limited training samples in the selected areas, but the low accuracy of mixed herbaceous marsh in these test areas was mainly caused by the difficulties of the training sample selection on the imagery because of the overlap of species between different plant communities. The testing results also suggested the necessity of post-classification manual refinement. After refinement of each test plot, the maps were considered very accurate and were used to train us to select training samples across the entire project area for a broad area mapping.

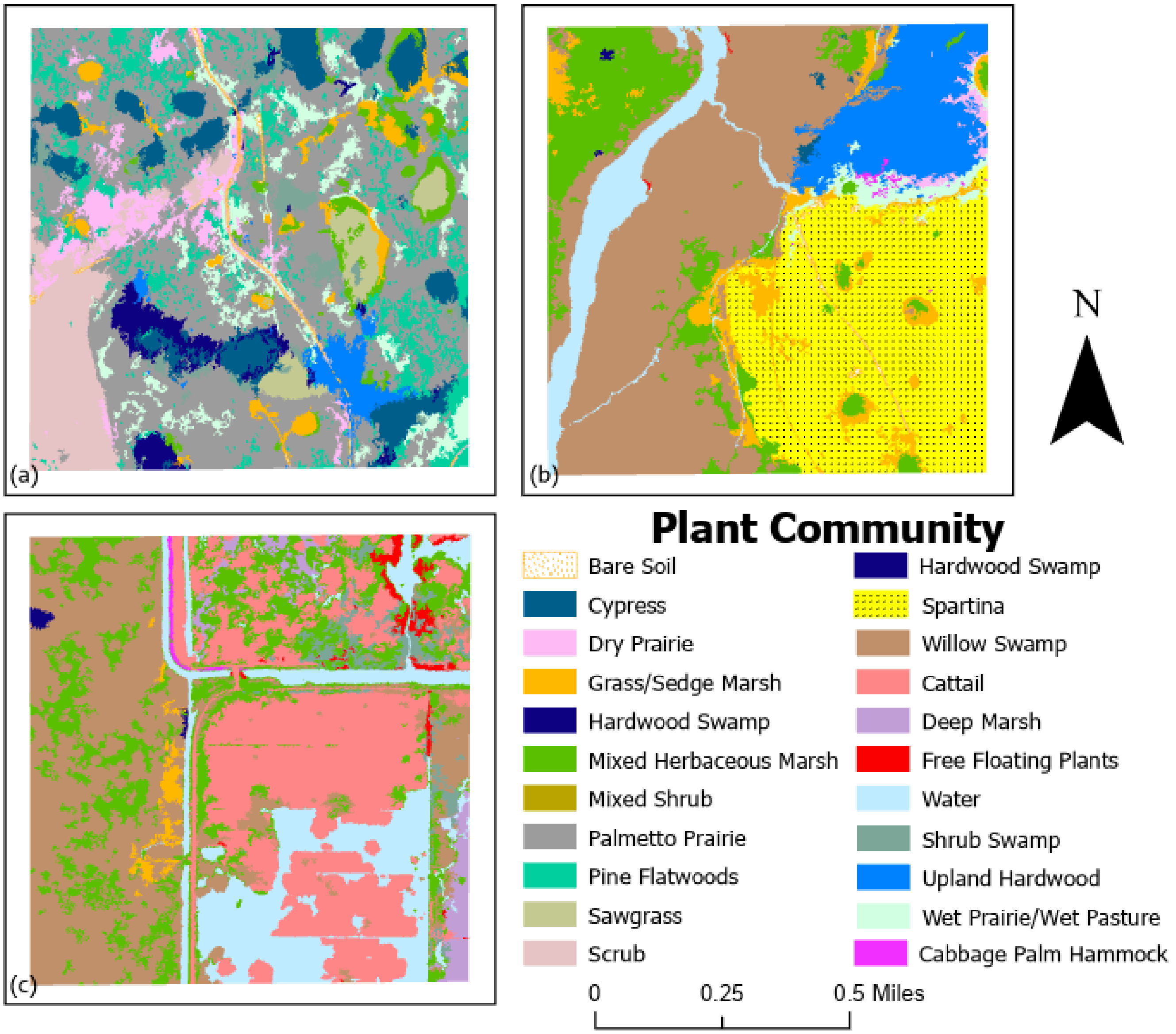



Figure 13 Maps of three 500-acre testing plots within (a) BULLJG, (b) SJMCA, and (c) between SJMCA and TFMCA.

*4.2 BULLJG*

The classified map for BULLJG is displayed in Figure 14, and the error matrix for this area is displayed in Table 2. This area is dominated by Palmetto Prairie (5,950 acres) and followed by Hardwood Swamp (5,690 acres). Pine Flatwoods, Cypress, Shrub Swamp, Wet Prairie, Grass Sedge, and Mixed Herbaceous Marsh are also present. In total, there are 17 plant communities in this area plus Levee/Road and Water. About 482 acres of this area is covered by Scrub which often not shown in other areas. Based upon the error matrix, an overall accuracy (OA) of 0.89 was achieved for this area with most communities having an accuracy larger than 80%. Minor communities include Willow Shrub (WS) (6 acres), and Free Floating plants (FF) (3 acres) were delineated manually and not included in the error matrix. PA and UH had accuracy less than 80% due to both being rare communities in this area. BULLJG has burned areas appearing dark on the imagery and was observed in the field too (Figure 15). Most of these burned areas were covered by Palmetto Prairie with sparse pines. After talking with Dr. Hall, these burned areas were manually labelled as Palmetto Prairie if Palmetto still showed up or Mixed Herbaceous Marsh if Palmetto was gone and became wet. If the area looked dry, it was mapped as pasture or prairie.

Table 2 Error matrix for BULLJG from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. OA=Overall Accuracy. Plant community abbreviations are listed in Table 1.

| | DP | PA | HM | WP | GS | SG | PP | SC | MS | SS | OW | UH | PF | HS | CY | CP | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DP | 19 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 21 | 0.91 |
| PA | 0 | 6 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 8 | 0.75 |
| HM | 0 | 0 | 30 | 0 | 2 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 36 | 0.83 |
| WP | 0 | 0 | 0 | 30 | 0 | 1 | 3 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 35 | 0.86 |
| GS | 0 | 0 | 3 | 1 | 30 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 36 | 0.83 |
| SG | 0 | 0 | 1 | 0 | 0 | 34 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 35 | 0.97 |
| PP | 0 | 0 | 0 | 2 | 0 | 0 | 35 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 38 | 0.92 |
| SC | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 1.00 |
| MS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 11 | 0.91 |
| SS | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 2 | 0 | 28 | 0.86 |
| OW | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4 | 1 | 0 | 0 | 0 | 0 | 5 | 0.80 |
| UH | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 3 | 0 | 21 | 1 | 1 | 0 | 0 | 27 | 0.78 |
| PF | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 26 | 0 | 0 | 0 | 30 | 0.87 |
| HS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 3 | 1 | 23 | 0 | 0 | 27 | 0.85 |
| CY | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 51 | 0 | 52 | 0.98 |
| CP | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 10 | 11 | 0.91 |
| Column Total | 20 | 6 | 34 | 33 | 33 | 39 | 42 | 24 | 11 | 28 | 4 | 28 | 31 | 24 | 55 | 10 | 422 | |
| PA | 0.95 | 1.00 | 0.88 | 0.91 | 0.91 | 0.87 | 0.83 | 0.92 | 0.91 | 0.86 | 1.00 | 0.75 | 0.84 | 0.96 | 0.93 | 1.00 | | |
| OA | 0.89 | | | | | | | | | | | | | | | | | |
| Kappa | 0.88 | | | | | | | | | | | | | | | | | |

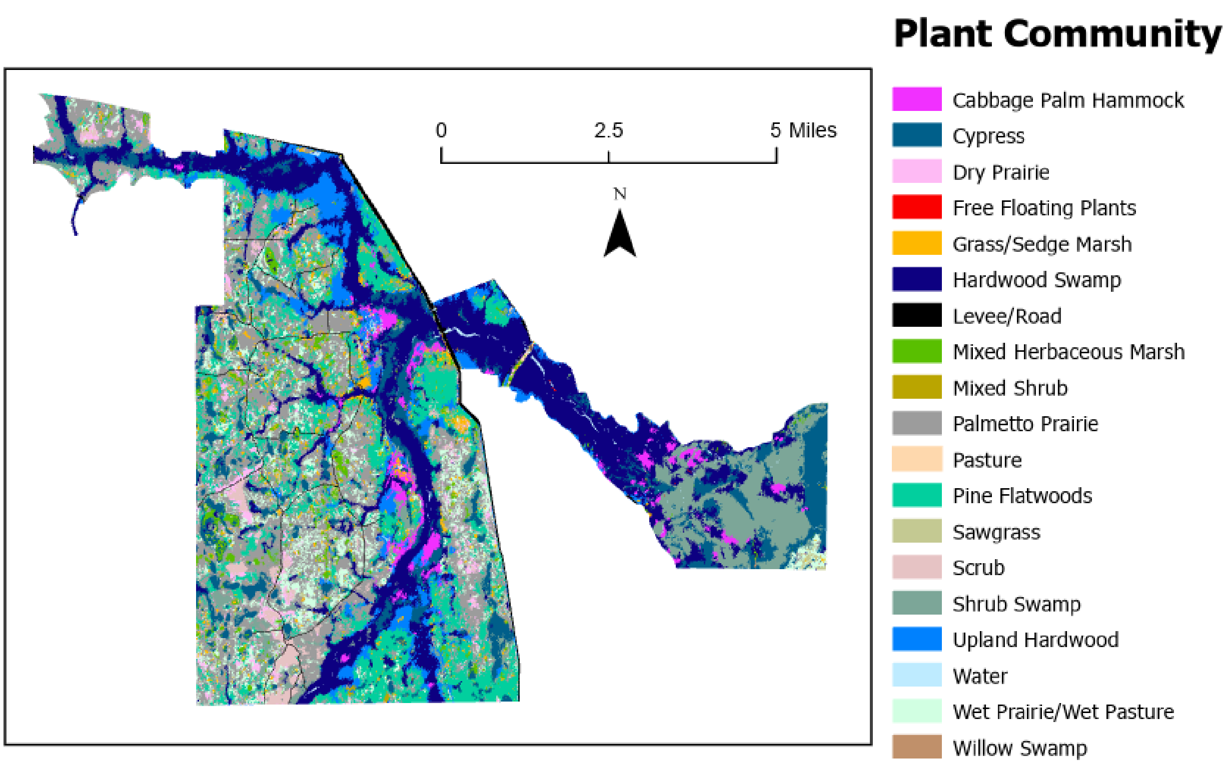



Figure 14 Classified map for BULLJG.

Figure 15 Burned Palmetto Prairie at BULLJG.

*4.3 SJMCA*

The classified map for SJMCA is displayed in Figure 16, and the error matrix for this area is displayed in Table 3. This area is the second largest area to map with a total of 12 plant communities, Water and Levee/Road. A small coverage of Free Floating plant (34 acres) and Dry Prairie (10 acres) were observed on the imagery and were labeled for corresponding segments manually due to limited training samples. An overall accuracy of 85% and Kappa value of 0.84 were obtained based upon a total of 781 reference segments. Mixed Herbaceous Marsh (6,923 acres) and Willow Swamp (6,756 acres) dominated this area. Grass Sedge, Spartina, and Shrub Swamp occupied between 1000-2000 acres. Hardwood Swamp, Cypress, Cabbage Palm Hammock, and other herbaceous marshes were less than 1000 acres. Spartina was the only plant community that only appeared in SJMCA among the nine mapped project areas in project year 2. For this area, Sawgrass (317 acres) had the lowest accuracy (UA=71%) and was confused with Mixed Herbaceous Marsh and Wet Prairie. This is not surprising because both Mixed Herbaceous Marsh and Wet Prairie can have part of Sawgrass within it, leading to confused spectral signatures among them. Mixed Shrub (MS, 73%) and Upland Hardwood (UH, 77%) also had an accuracy less than 80% due to confusion each other. This is also not surprising because some shrubs are bigger like tree canopies while some small trees within UH look like shrubs. Similar situation occurs in CP which can be confused with UH and MS, leading to a lower accuracy (77%). Three lakes including Lake Hellen Blazes, Little Sawgrass Lake, and Sawgrass Lake are within this area and are connected by St. Johns River. The water bodies are delineated well through the machine learning identification.

Table 3 Error matrix for the SJMCA from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | HM | WS | GS | MS | SP | OW | CY | SG | HS | WP | CT | UH | CP | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HM | 39 | 0 | 4 | 0 | 0 | 0 | 1 | 5 | 0 | 0 | 1 | 1 | 0 | 51 | 0.77 |
| WS | 0 | 90 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 90 | 1.00 |
| GS | 1 | 0 | 46 | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 1 | 0 | 0 | 51 | 0.90 |
| MS | 0 | 0 | 0 | 56 | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 2 | 5 | 77 | 0.73 |
| SP | 0 | 0 | 0 | 0 | 49 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 50 | 0.98 |
| OW | 0 | 0 | 0 | 0 | 0 | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 25 | 1.00 |
| CY | 0 | 0 | 0 | 0 | 0 | 0 | 29 | 0 | 0 | 2 | 0 | 0 | 0 | 31 | 0.94 |
| SG | 6 | 1 | 1 | 0 | 1 | 0 | 0 | 29 | 0 | 3 | 0 | 0 | 0 | 41 | 0.71 |
| HS | 0 | 2 | 0 | 3 | 0 | 0 | 0 | 0 | 110 | 0 | 0 | 0 | 2 | 117 | 0.94 |
| WP | 3 | 0 | 2 | 0 | 11 | 0 | 2 | 1 | 0 | 61 | 0 | 0 | 0 | 80 | 0.76 |
| CT | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 26 | 0 | 0 | 29 | 0.90 |
| UH | 0 | 1 | 0 | 4 | 0 | 0 | 0 | 0 | 3 | 0 | 1 | 57 | 8 | 74 | 0.77 |
| CP | 0 | 0 | 0 | 6 | 0 | 0 | 0 | 0 | 3 | 0 | 0 | 6 | 50 | 65 | 0.77 |
| Column Total | 50 | 94 | 53 | 69 | 62 | 25 | 32 | 35 | 130 | 71 | 29 | 66 | 65 | 781 | |
| PA | 0.78 | 0.96 | 0.87 | 0.81 | 0.79 | 1.00 | 0.91 | 0.83 | 0.85 | 0.86 | 0.90 | 0.86 | 0.77 | | |
| OA | 0.85 | | | | | | | | | | | | | | |
| Kappa | 0.84 | | | | | | | | | | | | | | |

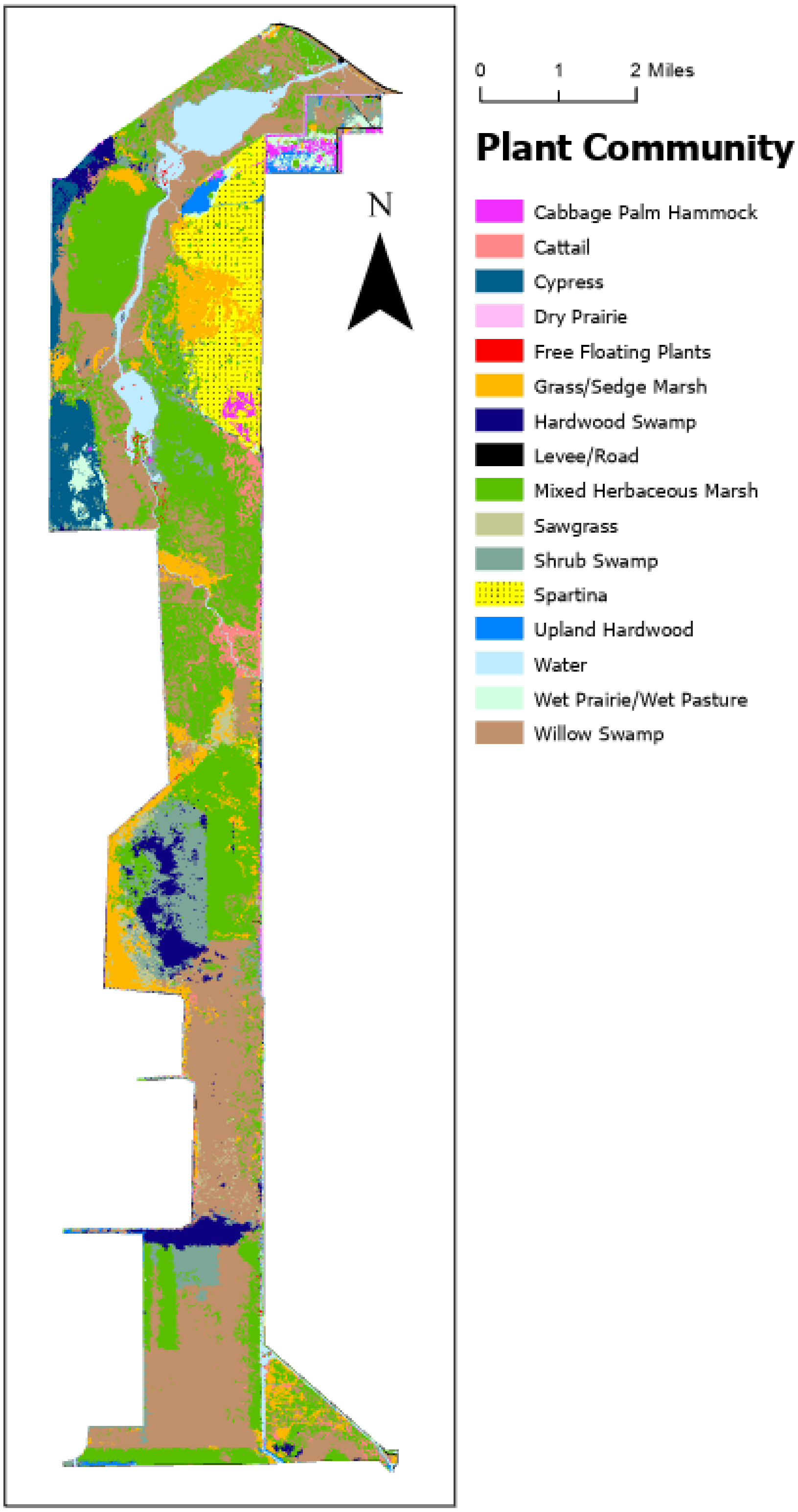



Figure 16 Classified map for SJMCA.

*4.4 C1DETEN and TFMCA*

C1DETEN and TFMCA were mapped together because these two areas are adjacent, and plant spectral signatures are similar due to similar environments. Also note that some communities are rare communities in TFMCA (e.g., UH and CP) but are major communities in C1DETEN, mapping them together can help training sample selection to develop a more robust classification model. Thus, these two areas shared the same error matrix as shown in Table 4. A higher OA of 0.87 and Kappa value of 0.86 were obtained based upon a total of 215 reference segments. HM, PA, CP, FF and MS have an accuracy less than 80% due to overlap of these plant community compositions.

For C1DETEN, Mixed Shrub is the major community with a coverage of 491 acres, followed by Pasture with a coverage of 391 acres (Figure 17). The east sector is dominated by Upland Hardwood (UH) and Cabbage Palm Hammock (CP). Sparse Hardwood Swamp is present over large areas of herbaceous wetlands in the west.

For TFMCA, the south section is dominated by Willow and Mixed Herbaceous Marsh (HM), middle sector is mainly covered by Water and Deep Marsh with Free Floating plants and Cattail dispersed, and the north sector is dominated by Deep Marsh and Cattail (Figure 18). Mixed Herbaceous Marsh dominates the northeast sector. The original boundary file from District showed a wrong dividing line between TFMCA and C1SLWMA where a large piece of Sawgrass should belong to C1SLWMA, and the boundary line between these two areas should follow the levee. Final map products for these two areas were revised by merging the two maps first and then split based upon the correct boundary.

Table 4 Error matrix for C1DETEN and TFMCA from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | DM | HM | UH | OW | PA | CP | CT | GS | WS | SG | FF | MS | WP | Row Total | User's Accuracy |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| DM | 32 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 33 | 0.97 |
| HM | 1 | 7 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 10 | 0.70 |
| UH | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 9 | 0.89 |
| OW | 1 | 0 | 0 | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 9 | 0.89 |
| PA | 1 | 0 | 0 | 0 | 10 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 1 | 14 | 0.71 |
| CP | 1 | 0 | 1 | 0 | 0 | 8 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 0.73 |
| CT | 0 | 1 | 0 | 0 | 0 | 1 | 29 | 0 | 0 | 0 | 0 | 0 | 0 | 31 | 0.94 |
| GS | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 12 | 0 | 0 | 0 | 0 | 0 | 13 | 0.92 |
| WS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 0 | 0 | 2 | 0 | 16 | 0.88 |
| SG | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 11 | 1.00 |
| FF | 2 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 0 | 13 | 0.77 |
| MS | 2 | 0 | 2 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 14 | 0 | 21 | 0.67 |
| WP | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 | 24 | 1.00 |
| Column Total | 40 | 8 | 11 | 9 | 13 | 10 | 31 | 13 | 15 | 11 | 11 | 18 | 25 | 215 | |
| Producer's Accuracy | 0.80 | 0.88 | 0.73 | 0.89 | 0.77 | 0.80 | 0.94 | 0.92 | 0.93 | 1.00 | 0.91 | 0.78 | 0.96 | | |
| Overall Accuracy | 0.87 | | | | | | | | | | | | | | |
| Kappa | 0.86 | | | | | | | | | | | | | | |

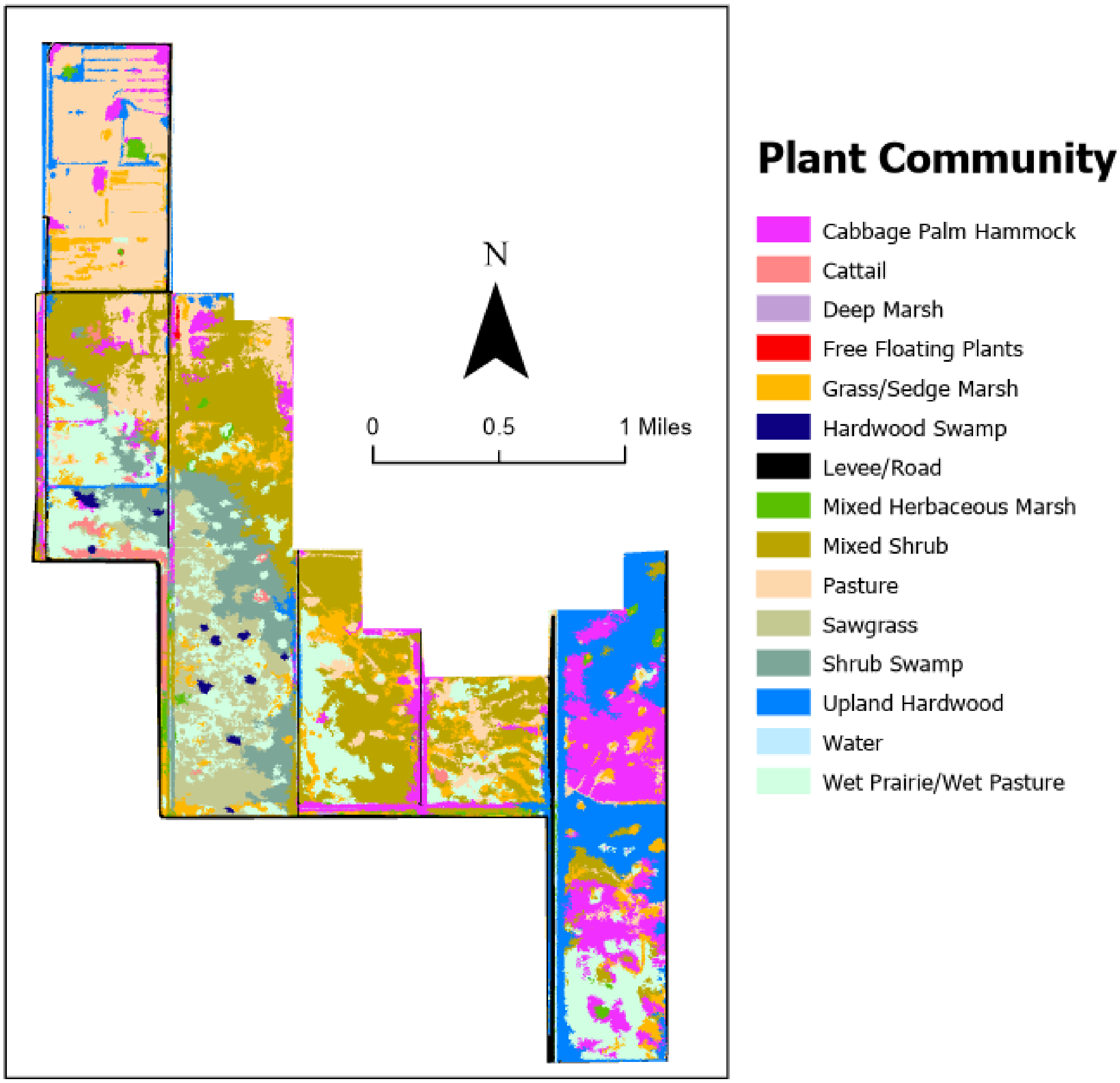



Figure 17 Classified map for C1DETEN.

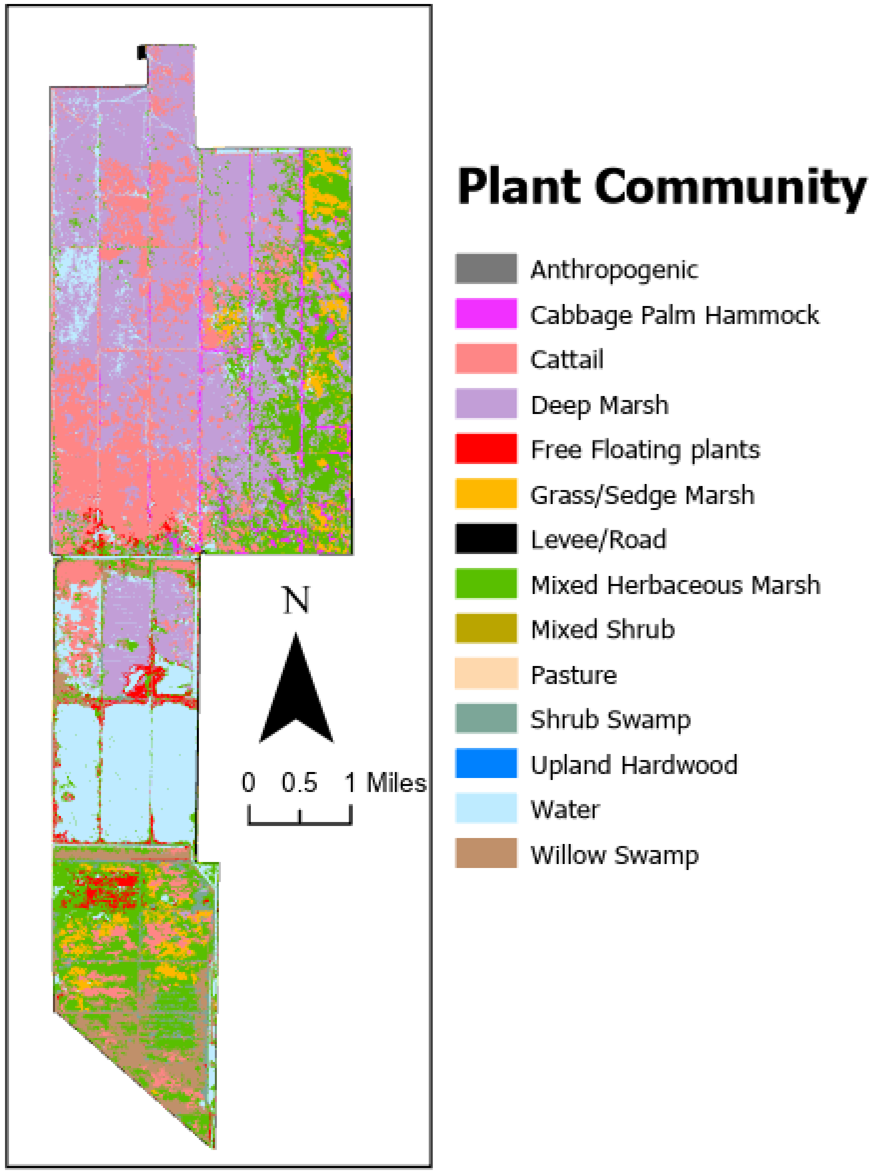



Figure 18 Classified map for TFMCA.

*4.5 DESERET*

Figure 19 shows the classified map for DESERET and Table 5 displays the error matrix result. Willow Swamp (1732 acres), Grass Sedge (1413 acres), Mixed Herbaceous Marsh (1298 acres) and Wet Prairie (1021 acres) are the four dominant communities. The north sector is mainly occupied by Willow Swamp while the south sector is mainly covered by Wet Prairie and Grass Sedge marsh. Cattail is a minor community (5 acres) in this area and was thus manually labeled. Sawgrass, Cypress, Hardwood Swamp, and Cabbage Palm Hammock are also present with a coverage less than 100 acres. A high accuracy was achieved for this area with an OA of 92% and Kappa value of 0.91 based upon 434 reference segments. Upland Hardwood appearing at the boundary of this area has a lower accuracy (70%) due to confusion with Cabbage Palm Hammock and Willow Swamp. Shrubs were labeled as Mixed Shrub (MS) in the classification and then manually further refined as Shrub Swamp if they were present in the wet areas. For this area, west section is drier than the east section. Mixed Shrub had a low classification accuracy of 59%, again, due to confusion with UH and CP. Similarly, UH was confused with CP, HS and shrubs, leading to an accuracy of 70%.

Table 5 Error matrix for DESERET from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | HM | WP | CP | GS | SG | MS | WS | FF | OW | UH | HS | CY | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| HM | 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 31 | 1.00 |
| WP | 0 | 70 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 70 | 1.00 |
| CP | 0 | 0 | 54 | 0 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 1 | 58 | 0.93 |
| GS | 0 | 0 | 0 | 31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 31 | 1.00 |
| SG | 0 | 0 | 0 | 0 | 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 19 | 1.00 |
| MS | 0 | 0 | 3 | 0 | 0 | 16 | 2 | 0 | 0 | 5 | 0 | 1 | 27 | 0.59 |
| WS | 0 | 0 | 0 | 0 | 0 | 0 | 43 | 0 | 0 | 2 | 1 | 0 | 46 | 0.94 |
| FF | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 20 | 0 | 0 | 0 | 0 | 20 | 1.00 |
| OW | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 47 | 0 | 0 | 0 | 48 | 0.98 |
| UH | 0 | 0 | 4 | 0 | 0 | 1 | 3 | 1 | 0 | 26 | 2 | 0 | 37 | 0.70 |
| HS | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 19 | 0 | 22 | 0.86 |
| CY | 0 | 0 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22 | 25 | 0.88 |
| Column Total | 31 | 70 | 65 | 31 | 19 | 18 | 49 | 21 | 47 | 37 | 22 | 24 | 434 | |
| PA | 1.00 | 1.00 | 0.83 | 1.00 | 1.00 | 0.89 | 0.88 | 0.95 | 1.00 | 0.70 | 0.86 | 0.92 | | |
| OA | 0.92 | | | | | | | | | | | | | |
| Kappa | 0.91 | | | | | | | | | | | | | |

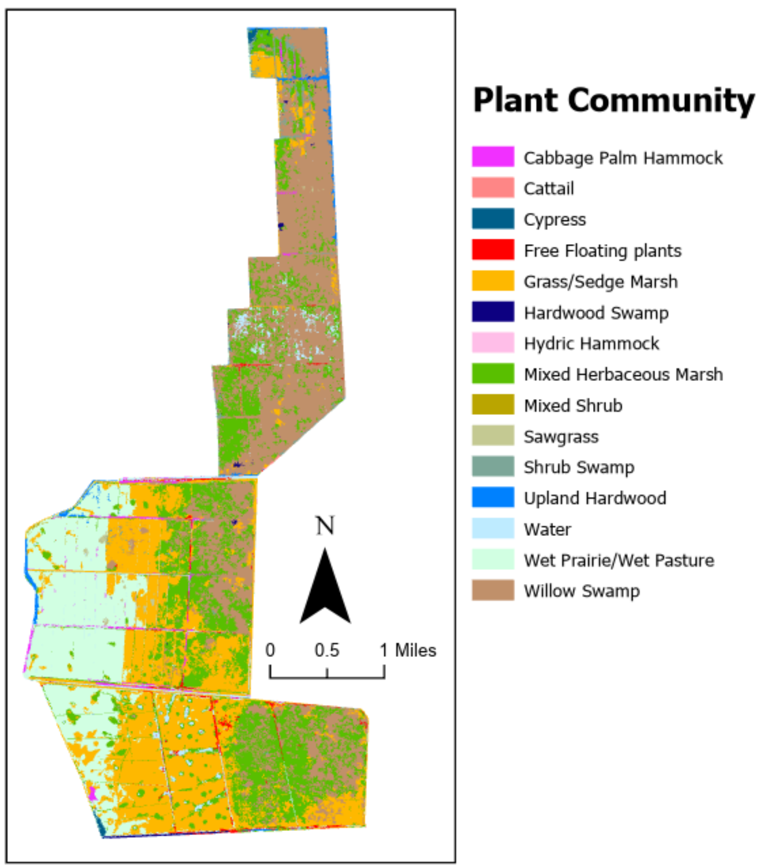



Figure 19 Classified map for DESERET.

*4.6 SCMRC*

At the beginning, SCMRC and SARTORI were mapped together because two areas are adjacent with a relatively small coverage. However, we found many incorrect labels of wetlands in SARTORI and their manual correction would be time consuming. Thus, SARTORI was separated, and training samples were recollected, and two areas were mapped separately. The error matrix result for the SCMRC is provided in Table 6, and the classified map is shown in Figure 20. For SCMRA, the north sector is dominated by Mixed Herbaceous Marsh, Grass Sedge, and the south sector is mainly covered by Willow Swamp on the east side and Wet Prairie and Pasture on the west side. Shrub Swamp and Mixed Shrub were manually labeled to avoid misclassification of Willow Swamp which is the dominant community over this area. Hardwood Swamp had the lowest accuracy (33%) and was mainly misclassified as Willow Swamp. Hardwood Swamp is a minor community occupying about 16 acres and was manually corrected based upon domain knowledge and the historic map, so the accuracy of this community is much higher than indicated by the error matrix.

Table 6 Error matrix for SCMRC from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | PA | HM | WP | CT | GS | WS | FF | OW | UH | CP | HS | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| PA | 25 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 26 | 0.96 |
| HM | 0 | 32 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 33 | 0.97 |
| WP | 1 | 0 | 63 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 66 | 0.96 |
| CT | 0 | 0 | 0 | 11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11 | 1.00 |
| GS | 1 | 1 | 5 | 0 | 20 | 0 | 0 | 0 | 0 | 0 | 0 | 27 | 0.74 |
| WS | 0 | 0 | 0 | 0 | 0 | 52 | 0 | 0 | 0 | 0 | 1 | 53 | 0.98 |
| FF | 0 | 0 | 0 | 0 | 0 | 1 | 17 | 0 | 0 | 1 | 0 | 19 | 0.90 |
| OW | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58 | 0 | 1 | 0 | 59 | 0.98 |
| UH | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 10 | 0 | 0 | 11 | 0.91 |
| CP | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 8 | 1 | 11 | 0.73 |
| HS | 0 | 0 | 0 | 0 | 0 | 9 | 0 | 0 | 0 | 1 | 5 | 15 | 0.33 |
| Column Total | 27 | 33 | 68 | 11 | 23 | 64 | 17 | 58 | 12 | 11 | 7 | 331 | |
| PA | 0.93 | 0.97 | 0.93 | 1.00 | 0.87 | 0.81 | 1.00 | 1.00 | 0.83 | 0.73 | 0.71 | | |
| OA | 0.91 | | | | | | | | | | | | |
| Kappa | 0.90 | | | | | | | | | | | | |

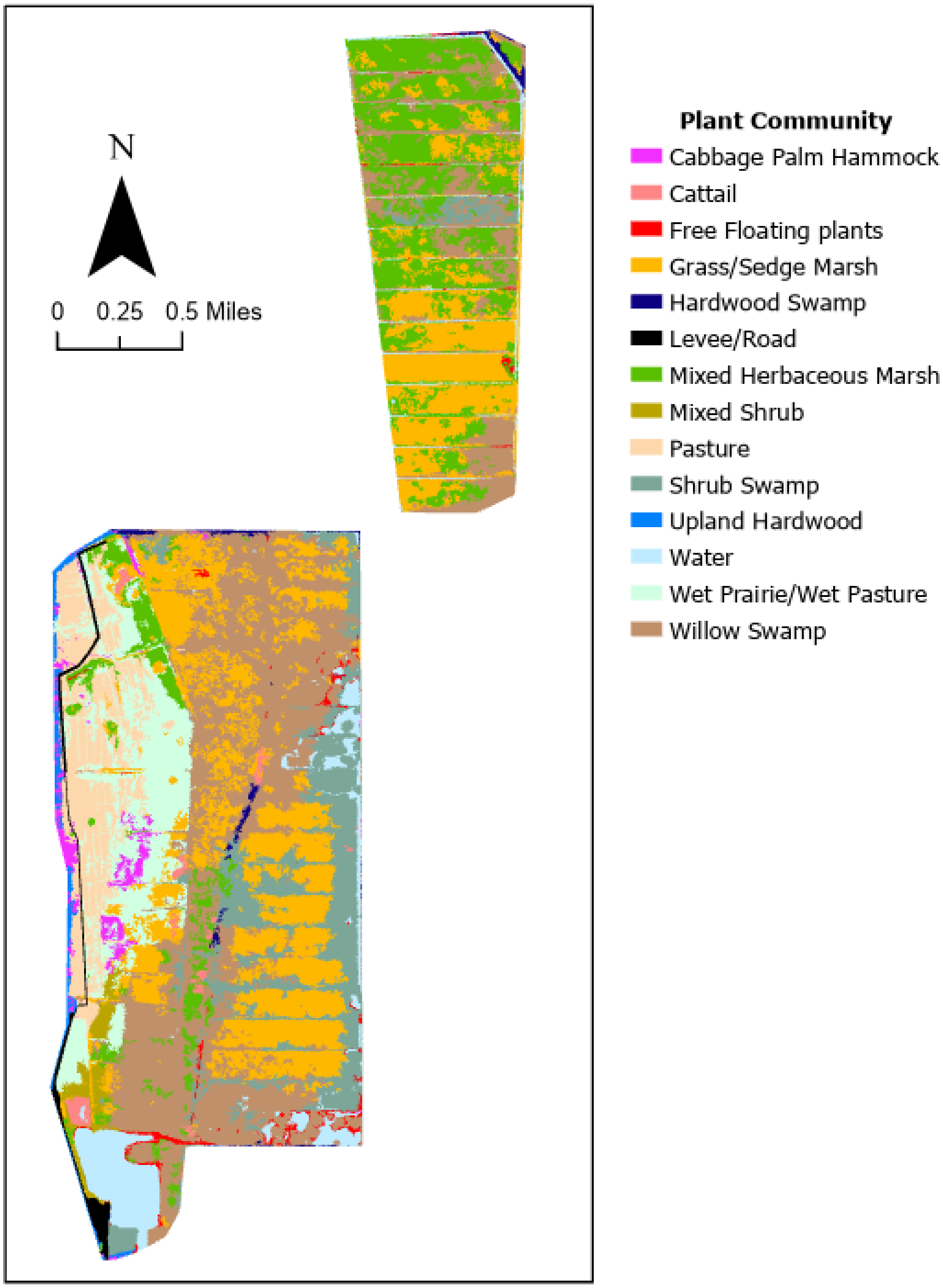



Figure 20 Classified map for SCMRC.

*4.7 SARTORI*

The classification accuracy for SARTORI is displayed in Table 7 and the classified map is displayed in Figure 21. SARTORI is mainly covered by Wet Prairie with Grass Sedge and Mixed Herbaceous Marsh interspersed, and trees or shrubs lined along levees. Mixed Herbaceous Marsh is present within the prairie. An OA of 88% was achieved in classify 8 communities here. DM and CP were manually labeled due to being rare over this area. A small portion of the FF is present with a lower accuracy in the classification (54%). GS also had a lower accuracy (61%) due to overlap composition with HM and WP.

Table 7 Error matrix for SARTORI from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | WP | HM | GS | CT | UH | FF | OW | SS | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|
| WP | 211 | 0 | 2 | 0 | 1 | 0 | 0 | 0 | 214 | 0.99 |
| HM | 1 | 51 | 12 | 0 | 0 | 0 | 1 | 0 | 65 | 0.79 |
| GS | 7 | 17 | 37 | 0 | 0 | 0 | 0 | 0 | 61 | 0.61 |
| CT | 0 | 0 | 0 | 5 | 0 | 0 | 0 | 0 | 5 | 1.00 |
| UH | 1 | 0 | 0 | 0 | 12 | 0 | 0 | 2 | 15 | 0.80 |
| FF | 0 | 0 | 2 | 0 | 0 | 7 | 0 | 4 | 13 | 0.54 |
| OW | 0 | 0 | 0 | 0 | 0 | 0 | 13 | 0 | 13 | 1.00 |
| SS | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 37 | 37 | 1.00 |
| Column Total | 220 | 68 | 53 | 5 | 13 | 7 | 14 | 43 | 423 | |
| PA | 0.96 | 0.75 | 0.70 | 1.00 | 0.92 | 1.00 | 0.93 | 0.86 | | |
| OA | 0.88 | | | | | | | | | |
| Kappa | 0.83 | | | | | | | | | |

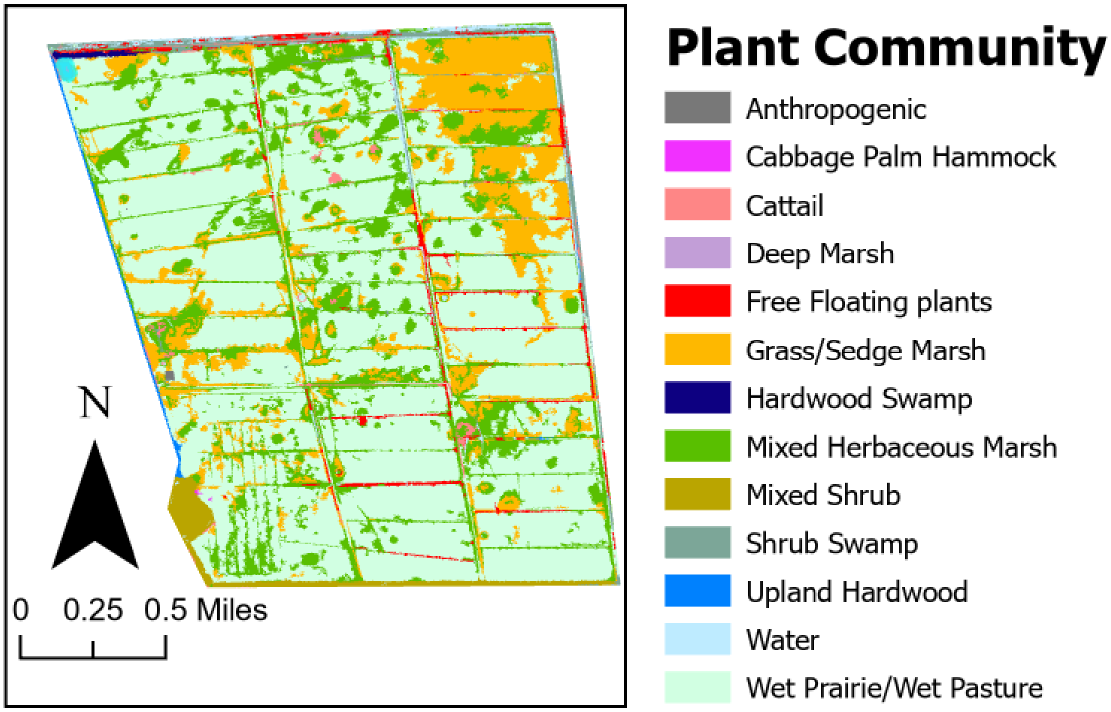



Figure 21 Classified map for SARTORI.

*4.8 C54BMRA*

The classified map for C54BMRA is displayed in Figure 22 and the error matrix is displayed in Table 8. An OA of 88% and Kappa value of 0.85 were obtained for this area in terms of 1006 reference segments. As shown in Table 8, major communities all achieved a high accuracy (>80%) in the identification such as Deep Marsh, Mixed Herbaceous Marsh, Grass Sedge, and Submerged Aquatic Beds. Free Floating plants are difficult to identify for this area because they appeared in different shapes and textures. Some are smooth patches, while some are dispersed with a look like flooded grasses. Confusion of floating plants with grasses is not surprising. Intensive manual refinement work was conducted after machine learning classification for this area due to large spectral variations for each individual community. For example, grasses can look very different across the sectors, leading to confusion among Grass Sedge, Free Floating plants, Deep Marsh, and Mixed Herbaceous Marsh. After manual refinement, major communities were delineated well. Small Cattail patches dispersed within large water bodies were also well delineated. Submerged Aquatic Beds was also present for this area and mapped (75 acres).

This area was the most challenging area to map because it undergoes significant land management activities, including herbiciding and flooding on a yearly basis. It was difficult to collect spectral signatures from the imagery because the 2016 map provided less information in this area relative to other areas which only saw small changes. Flooded vegetation was mixed with Submerged Aquatic Beds and floating plants, and Deep Marsh was confused with Grass Sedge which looked different from the smooth signature of grasses seen over other areas. We carefully selected spectral signatures and training samples during multiple collaboration events with Dr. Hall to confirm each community on the imagery and reviewed data taken during field visits before developing the model. We also met with Joel Andreas, a biologist with the Florida Fish and Wildlife Conservation Commission (FWC), who is very familiar with the plant communities and management activities for this area, since it is managed by FWC, rather than the District. He provided information essential to the mapping effort. Accurate spectral signature collection and training/reference sample acquisition is critical in the mapping procedure, since mislabeled training samples will directly cause misclassification of communities.

Table 8 Error matrix for C54BMRA from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | DM | HM | CT | GS | SS | WS | SAB | FF | OW | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|
| DM | 92 | 1 | 0 | 2 | 1 | 0 | 0 | 0 | 0 | 96 | 0.96 |
| HM | 2 | 280 | 6 | 22 | 6 | 0 | 0 | 0 | 0 | 316 | 0.89 |
| CT | 0 | 15 | 69 | 0 | 2 | 0 | 1 | 0 | 0 | 87 | 0.79 |
| GS | 3 | 23 | 0 | 206 | 4 | 0 | 0 | 2 | 1 | 239 | 0.86 |
| SS | 0 | 4 | 2 | 3 | 93 | 4 | 0 | 0 | 0 | 106 | 0.88 |
| WS | 0 | 0 | 0 | 0 | 7 | 29 | 0 | 0 | 0 | 36 | 0.81 |
| SAB | 0 | 3 | 0 | 0 | 0 | 0 | 20 | 0 | 1 | 24 | 0.83 |
| FF | 2 | 0 | 0 | 5 | 1 | 1 | 0 | 11 | 0 | 20 | 0.55 |
| OW | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 81 | 82 | 0.99 |
| Ccolumn Total | 99 | 327 | 77 | 238 | 114 | 34 | 21 | 13 | 83 | 1006 | |
| PA | 0.93 | 0.86 | 0.90 | 0.87 | 0.82 | 0.85 | 0.95 | 0.85 | 0.98 | | |
| OA | 0.88 | | | | | | | | | | |
| Kappa | 0.85 | | | | | | | | | | |

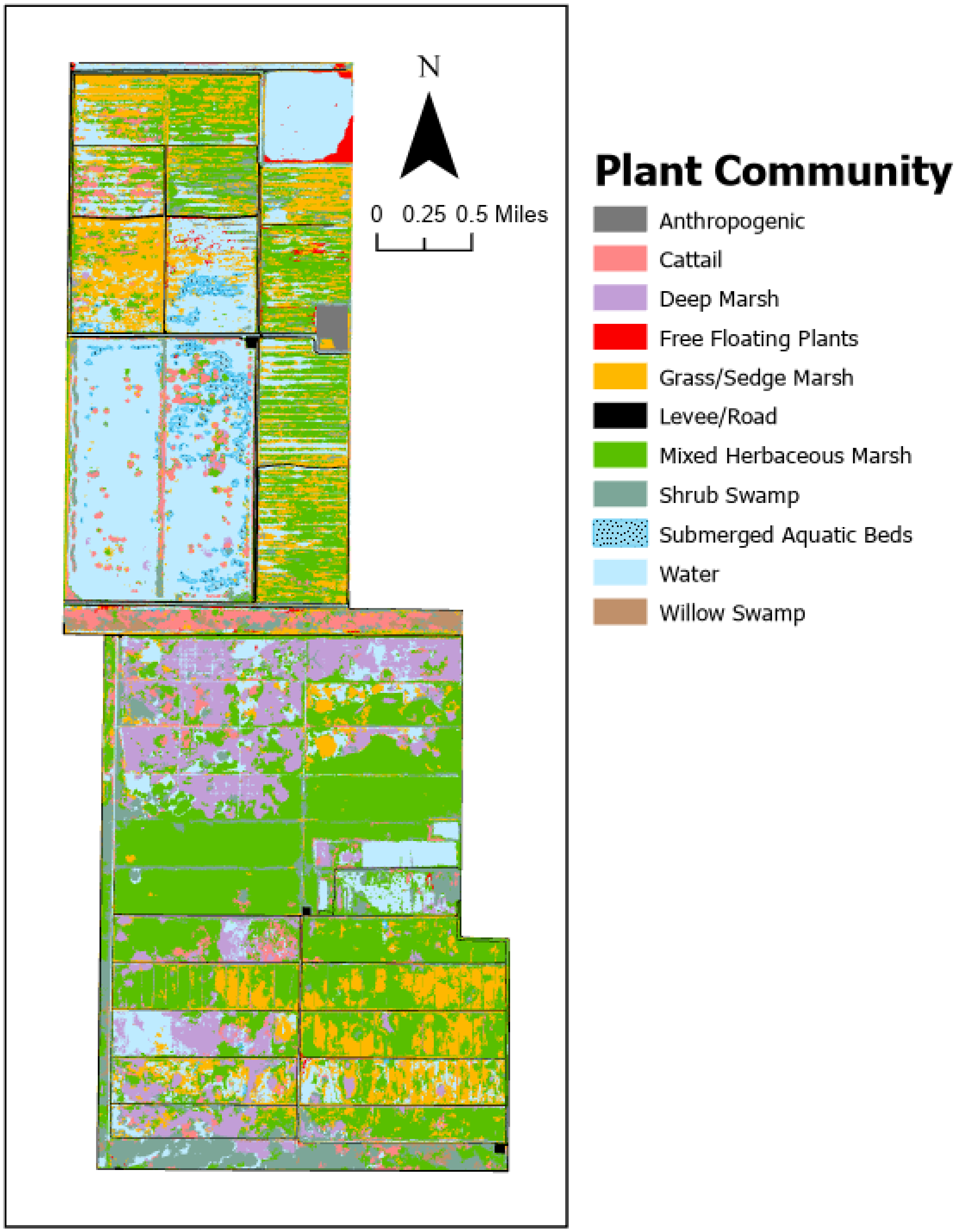



Figure 22 Classified map for C54BMRA.

*4.9 C1SLWMA*

Table 9 and Figure 23 show the mapping results for the C1SLWMA. The south part is dominated by Sawgrass, Cattail and open water in the middle sector and Wet Prairie in the north part. Relatively large homogeneous Cabbage Palm Hammock patches are well present over this area, which offers good signature collection for this community. Narrow features, and small and large patches of Cabbage Palm Hammock were delineated well via the classification. Shrubs within the big patch of Wet Prairie were also well depicted. An OA of 86% and Kappa value of 0.85 were obtained with major communities having an accuracy larger than 80% in terms of 692 reference samples. A lower accuracy of UH (78%) was caused by confusion with CP. A lower accuracy of Spartina (SP, 78%) due to mixture with Wet Prairie (WP).

Table 9 Error matrix for C1SLWMA from the ensemble classification. UA=User's Accuracy, PA=Producer's Accuracy. Plant community abbreviations are listed in Table 1.

| | GS | SG | HM | WP | CT | CP | OW | MS | SP | UH | PA | DM | Row Total | UA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| GS | 38 | 1 | 1 | 3 | 3 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 47 | 0.81 |
| SG | 1 | 51 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 53 | 0.96 |
| HM | 0 | 2 | 38 | 0 | 3 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 44 | 0.86 |
| WP | 1 | 1 | 0 | 57 | 2 | 2 | 0 | 2 | 1 | 2 | 1 | 0 | 69 | 0.83 |
| CT | 0 | 0 | 2 | 2 | 88 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 92 | 0.96 |
| CP | 0 | 0 | 0 | 1 | 1 | 85 | 0 | 9 | 0 | 7 | 0 | 1 | 104 | 0.82 |
| OW | 0 | 0 | 0 | 0 | 0 | 2 | 35 | 1 | 0 | 0 | 0 | 1 | 39 | 0.90 |
| MS | 0 | 0 | 0 | 3 | 0 | 8 | 0 | 61 | 0 | 2 | 0 | 0 | 74 | 0.82 |
| SP | 0 | 0 | 0 | 3 | 1 | 0 | 0 | 0 | 14 | 0 | 0 | 0 | 18 | 0.78 |
| UH | 0 | 0 | 2 | 1 | 0 | 14 | 0 | 1 | 0 | 62 | 0 | 0 | 80 | 0.78 |
| PA | 0 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 10 | 0 | 12 | 0.83 |
| DM | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 58 | 60 | 0.97 |
| Column Total | 41 | 55 | 45 | 72 | 98 | 111 | 35 | 74 | 16 | 73 | 11 | 61 | 692 | |
| PA | 0.93 | 0.93 | 0.84 | 0.79 | 0.90 | 0.77 | 1.00 | 0.82 | 0.88 | 0.85 | 0.91 | 0.95 | | |
| OA | 0.86 | | | | | | | | | | | | | |
| Kappa | 0.85 | | | | | | | | | | | | | |

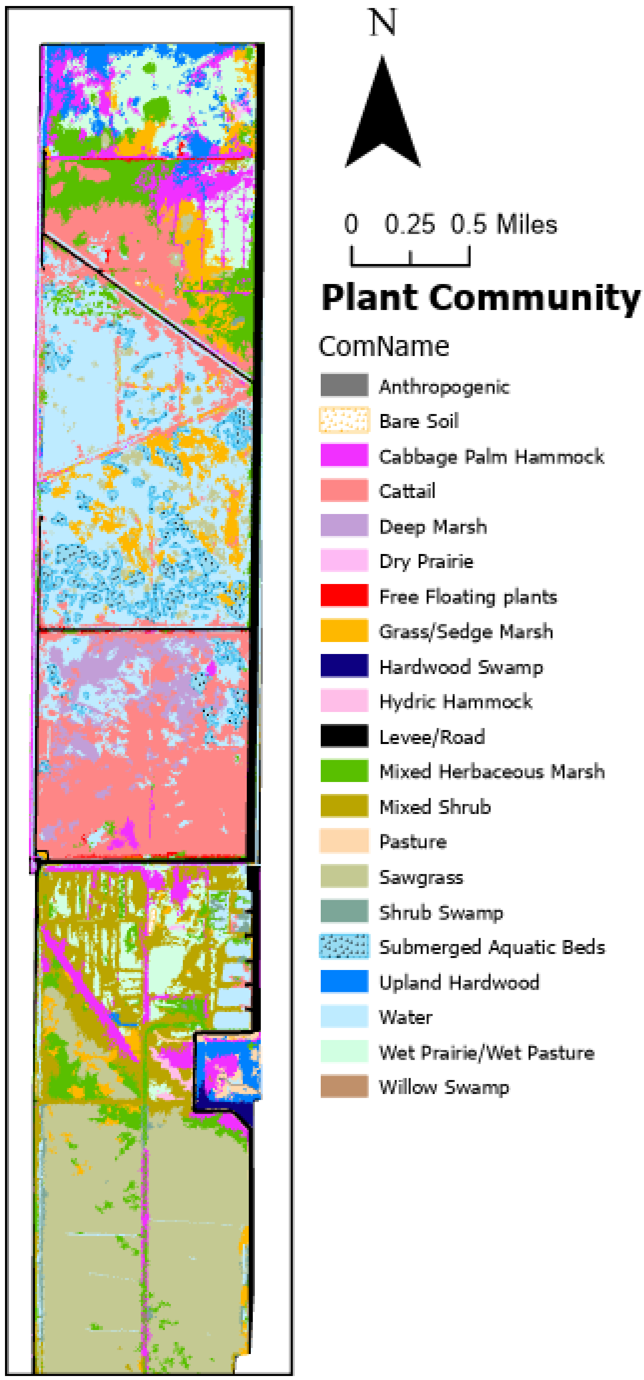



Figure 23 Classified map for C1SLWMA.

## 5. Discussion

*5.1 Factors impacting map quality*

For each project area, we achieved an overall accuracy larger than 80% based on the classification model, but we believe that the integrity of the final maps was higher after the post-classification manual refinement. A lower accuracy for an individual community for a specific area occurred when there were fewer training samples due to difficulties in collecting samples caused by the small size coverage or narrow linear features, or the degree of mixture levels with other communities. These communities often cover less than 5% of each individual project area and thus their effects are minor considering the entire area being mapped. A qualitative examination of each map laid over the imagery and using the swipe function in ArcGIS Pro was also conducted in addition to the quantitative error matrix techniques. The general pattern of each map was consistent with the visual observation of the imagery.

The most critical factor impacting the mapping accuracy is the degree of complexity of the area of interest. In general, more classes, more heterogeneity within a class, and more mixture between classes will be more difficult to map. From a methodology perspective, a major step impacting the map accuracy and results is the training sample selection. Ideally, training samples should be selected based on a stratified sample selection strategy through field visits, but, in practice, field sample collection is difficult especially in inaccessible areas. Misalignment between field samples and imagery can cause issues too. The time of field visit should also be considered, and a large time gap between field visit and acquisition of WV-2 imagery will result in more uncertainties in both field training sample selection and post-mapping ground-truth accuracy assessment. For example, in project year 2, WV-2 imagery was acquired in March 2023, but field spectral keys were collected in June. During the field visit, change was observed between the imagery and field observations over some areas. Therefore, a timely field visit is recommended. The WV-2/3 imagery products have a very fine resolution (0.5 m in this project), which allows reference sample collection by a visual interpretation procedure based upon old maps and current imagery. However, we must be careful not to perpetuate errors from the old map into the new one. A combination of this visual interpretation with samples collected by field visits is more effective than the application of imagery-based or field-based sample collection alone. Special care should be paid in obtaining training samples from communities containing similar groups of species. For example, Mixed Herbaceous Marsh (HM) communities, may contain cattail and small shrubs, which will be present in the Shrub Swamp and Cattail communities as well. Training points should be selected over homogeneous areas where only one community is present. For instance, training samples for Cattails should be taken in areas where cattail is at least 70% of the community. Likewise, training points for Shrub Swamp should only be taken from areas where shrubs are at least 70% of the community. Mislabeling of any training samples will lead to misclassifications on the map and be reflected in the error matrix.

Selection of the scale parameter in image segmentation also impacts the mapping results. In project year 1, we selected the scale of 300 to avoid the “salt-and-pepper” effect, reduce the size of the dataset for future processing and use, and in consideration of the areal constraints on management activities. The application of this scale created a lot of mixture issues and missed the delineation of small patches or linear features. In project year 2, we decreased the scale to 150 to solve the mixture issue. With the decreased scale, features were segmented better. However, this one-scale for the entire project area may not work for heterogeneous landscapes which can still generate segments with mixtures of plant communities, leading to confusion in the classification. Segmentation refinement is a complicated procedure and, so far, there are no automated methods for this improvement.

Map quality is also influenced by the boundary effects. Ideally, single-date imagery should be used for each individual area. This is easier for an area with a small size but for a large project area such as SJMCA with a narrow shape, imagery often comes from a mosaic of multiple image scenes or images collected on different dates. Inconsistency will occur at the boundary of the mosaic lines. Similarly, when mapping a large area, an image needs to be split so that the computer or software packages can handle the dataset. This can cause boundary inconsistencies and misalignments among the segments across the splitting line, leading to map artifacts. A solution for this is the manual editing/revision of these boundary segments. This is a time-consuming procedure.

*5.2 Challenges in the project*

One challenge in the mapping project was to collect field data within a project area. We have collected data along levees and airboat trails but to better train our own brain to label communities on the imagery, especially for areas we are not sure what they are, a field visit is the best solution. Unfortunately, many of these areas are inaccessible to humans under most conditions. To mitigate this, we met with District's personnel such as Dr. Hall and Kimberly Ponzio online to validate spots where we wanted to put training samples. The District's personnel have rich knowledge across all the project areas and were invaluable in assisting with the labeling of many samples on the imagery to be used as training data to develop the machine learning models. Knowledge of what plant communities should or should not appear in a specific area, and what a community looks like on the imagery and in the field was vital for us to acquire effective training samples for classification.

The other challenge was in training the students. In the PI's team, a total of eight graduate students have been involved in this project including six PhD students (David Brodylo, Mizanur Rahman, Sandip Rijal, Abdullah Ai-Fazari, Fiona Benzi, and Rabindra Parajuli) and two MS students (David Ramirez, and Madan Thapa Chhetri). Exposure of graduate students to a research project and their participation in it is critical for their graduate training and future careers. For this project, students needed to be trained to identify each plant community in the field and on the imagery. They needed to learn satellite imagery processing, data editing, how to run machine learning models, and how to generate the final map products. Although the project PI has developed detailed tutorials and generated a plant inventory with descriptions and pictures, it is still challenging for students to correctly label training samples. In project year 1, none of these students collected effective training samples on the imagery nor generated acceptable map products for their assigned project area. However, during the first phase of the mapping project, most students have become proficient in the entire mapping process and gained field experience. They have increased their skills using ArcGIS Pro, creating, and running *R* scripts, and using ESRI's Field Map software. They have also become proficient at eCognition image segmentation and data preprocessing, and in organizing, sharing, and backing up data. Students have also learned how to work on a project as a team member, how to communicate with others, as well as how to appreciate the beauty of imagery and maps. In project year 2, PhD students Sandip Rijal, Fiona Benzi, and Rabindra Parajuli, and MS student Madan Thapa Chhetri were hired in summer 2023. They all created map products this year though their maps need further revision and editing by the PI. Students may have incorrectly labeled misclassified segments in the map refinement procedure, forgot to include required attributes, or are confused about the label codes used. There should be two sets of label codes with one set to be used in the machine learning classification, and one set to be used for final map generation. The label code for final map generation should add the minor communities which are manually labeled and do not come into the machine classification procedure. I believe that in project year 3 (2024-2025), the students will be able to generate more acceptable map products. However, most of the trained students including Sandip Rijal, Mizanur Rahman, Abdullah Ai-Fazari, Rabindra Parajuli, and Madan Thapa Chhetri cannot work on the project anymore due to graduation or lack of time.

*5.3 Estimated time cost*

The previous mapping method using manual interpretation of aerial photograph by the District would take over a year to generate the map products covering around 80,000 acres. Our estimated time cost is around 10 weeks (2-3 months) after we receive the data from the District, if one person, proficient in plant community identification and the mapping procedure works solely on this project. The largely reduced time length in map generation will guide timely management activities for preserving the District's natural systems. A detailed time cost for major steps in our computer-based machine learning classification procedure is provided below (40 hours/week fully dedicated to the project by the PI).

- Preparation of field plan: 1 week
- Field visit: 1 week
- On-line meeting with District's personnel: 1 week
- Map generation: 1 week per project area, a total of 9 weeks for 9 areas
- Wrap up of deliverables: 1 week including generating levee/road and AN layers

## 6. Summary and deliverables

We have completed the required tasks in project year 2 (2023-2024) for mapping nine project areas with a total coverage of 88,016 acres. We conducted a 5-day field trip to the project area, collected 3429 field points, and took 261 pictures. A summary of the deliverables is provided below:

- Field samples, photos, and notes
- Map products of testing plots and accuracies
- Map products of nine project areas and accuracies
- Map product of a mosaic of all nine project areas
- Annual report (this file)
- Map plots (in preparation, need the Project Manager's confirmation)

## References

Zhang, C., 2014. Combining hyperspectral and lidar data for vegetation mapping in the Florida Everglades. Photogrammetric Engineering & Remote Sensing 80:733–743.

Zhang, C., and Z. Xie. 2014. Data Fusion and classifier ensemble techniques for vegetation mapping in the Everglades. Geocarto International 29: 228–243.

Zhang, C., D. Selch, and H. Cooper. 2016. A framework to combine three remotely sensed data sources for vegetation mapping in the central Florida Everglades. Wetlands 36: 201–213.

Zhang, C., S. Denka, and D. R. Mishra, 2018. Mapping freshwater marsh species in the wetlands of Lake Okeechobee using very high-resolution aerial photography and lidar data. International Journal of Remote Sensing 39: 5600–5618.

## Acknowledgement

We are grateful that the District has provided this project as an opportunity for FAU's graduate students: David Brodylo (graduated Summer 2023), Mizanur Rahman (graduated Summer 2024), Rabindra Parajuli (graduated Summer 2024), David Ramirez (graduated Spring 2024), Madan Thapa Chhetri (graduated Summer 2024), Sandip Rijal, Abdullah Ai-Fazari, and Fiona Benzi to put their classroom learning into a real field project, allowing for the application of geospatial techniques, and generating industry level map products. Some students (Mizanur Rahman and Fiona Benzi) have been working on their thesis/dissertation based upon this project. We also would like to extend our gratitude to the District's Project Manager, Dr. Dianne Hall, for her rich knowledge over the project area, patience to teach the students and PI to identify each community and understand their ecological significance, for providing detailed considerations and logistics for the field visit and improving the final map products and project report. District personnel Christy Akers, Kimberly Ponzio, Jonny Baker, Doug Voltolina and Chris O'Hara, and retired employee Ken Snyder made the field survey a great success.

## Appendix 1 Classification system used in the project

PLANT COMMUNITY CLASSIFICATIONS AND DEFINITIONS

| Community Type | Plant Community | Definition |
|---|---|---|
| **Other (O)** | Anthropogenic (AN) | Areas of agricultural land (including orchards, groves, and row crops, but NOT pasture or pine plantations), and buildings, parking lots, spoil piles, development, etc. Includes bare areas associated with anthropogenic disturbance. |
| | Levee/Road (LR) | Paved or unpaved roads or levees with grassed, gravel, limerock or dirt roads at their apex; levees without roads should be classified based on the vegetation they support. |
| | Bare Soil (BS) | Open areas with > 70% cover of bare soil with no vegetation (e.g., shoreline, beach, mudflats, barrens); not associated with anthropogenic disturbance or salt flats. |
| **Water (W)** | Water (OW) | Areas such as lakes, impoundments, rivers, canals, ditches, and other areas of open water which may support free-floating plant species (e.g., *Eichhornia, Pistia, Salvinia, Lemna*). |
| | Submerged Aquatic Beds (SAB) | Areas containing rooted aquatic plants with photosynthetic tissues below the water surface (e.g., *Vallisneria, Najas, Hydrilla, Potamogeton*). |
| | Free floating plants (FF) | Areas that are flooded and dominated by floating plant species (e.g., *Eichhornia, Pistia, Salvinia, Lemna*) that are not attached or rooted to the substrate of a water body. |
| **Herbaceous Wetland (HW)** | Salt Marsh/Salt flat (SM) | Areas where salty groundwater seeps to the surface and supports salt-tolerant plants (e.g., *Spartina, Sesuvium, Salicornia, Iva, Juncus roemarinus,* and small *Baccharis* shrubs); bare soil is often a significant part of this community and should not be mapped separately. |
| | Deep Marsh (DM) | Areas containing 2: 40% cover of bottom-rooted species with floating leaves (e.g., *Nymphaea, Nuphar, Nelumbo, Brasenia, Nymphoides*) or deep -water emergent species (e.g., *Schoenoplectus*); may also include *Utricularia* spp and may occur as a littoral zone around lakes. |
| | Grass/Sedge Marsh (GS) | Areas containing 2: 70% cover of obligate or facultative-wet grass or sedge spp. (e.g., *Panicum, Sacciolepsis, Eleocharis, Rhynchospora, Cyperus*); other non-graminoid species, such as *Typha* and small shrubs may be present in small amounts. |
| | Mixed Herbaceous Marsh (HM) | Areas containing a mixture of broadleaf emergents, grasses, sedges, and other spp. (e.g., *Pontederia, Sagittaria, Hydrocotyle, Persicaria);* small shrubs may be included as a small part of the community (e.g., *Cephalanthus, Kosteletzkya, Hibiscus*). |
| | Cattail (CT) | Areas containing 2: 70% cover of *Typha* spp. |
| | Phragmites (PH) | Areas containing 2: 70% cover of *Phagmites berlandieri.* |
| | Sawgrass (SG) | Areas containing 2: 70% cover of *Cladium jamaicense.* |
| | Spartina (SP) | Areas containing 2: 70% cover of *Spartina bakeri.* |
| | Wet prairie/wet pasture (WP) | Areas containing a mixture of grasses, sedges, rushes, and forbs (e.g., *Spartina, Juncus, Panicum*) typically classified as facultative-wet or facultative; long-hydroperiod species (e.g., *Typha*, water lilies) should not be present; category should also be used for reflooded pastures and wet, unimproved pastures. |
| **Shrub Wetland (SW)** | Shrub Swamp (SS) | Areas containing 2: 70% cover of mixed shrub species (e.g., *Salix, Morella, Baccharis, Sambucus, Ludwigia, Cephalanthus, Kosteletzkya, Hibiscus*); may include small trees (e.g., *Acer, Gordonia, Persea, Ilex, Sabal, Schinus*). |
| | Willow Swamp (WS) | Areas containing 2: 70% cover of *Salix caroliniana.* |
| **Forested Wetland (FW)** | Cypress (CY) | Areas containing 2: 70% cover of *Taxodium* spp. (bald or pond cypress); includes cypress domes, swamps, strands and lakeshore variants. |
| | Hardwood Swamp (HS) | Areas containing 2: 70% mixed wetland tree species (e.g., *Acer, Taxodium, Fraxinus, Nyssa, Ulmus, Annona*), cypress may be a significant component, but < 70%; includes swamps and communities lying in the floodplain of rivers and streams; understory typically doesn't include *Serenoa repens* (saw palmetto). |

| Community Type | Plant Community | Definition |
|---|---|---|
| **Forested Wetland (FW) cont.** | Bayhead/gall (BG) | Areas containing > 50% cover by one or more evergreen bay trees (e.g., *Gordonia, Persea, Magnolia*); may be in combination with *Pinus*, typically *Pinus serotina.* Ilex cassine may also be present. Subcanopy and understory may be dominated by *Serenoa repens* or bays. Ferns (e.g., *Osmunda cinnamomea, Woodwardia virginica, Thelypteris* spp.) may also be common. |
| | Hydric Hammock (HH) | Areas containing 2: 70% canopy cover of *Quercus laurifolia, Sabal palmetto, Magnolia virginiana, Juniperus, Celtis, Liquidamber*, and/or *Ulmus spp.* and having shorter hydroperiods than Hardwood Swamp; tree canopy may commonly include *Pinus*, but will rarely include *Taxodium* spp.; understory may include *Serenoa repens*, *Callicarpa americana* and *Psychotria nervosa* or ferns |
| | Cabbage Palm Hammock (CP) | Areas similar in species composition to Hydric Hammocks, but containing 2: 70% cover of *Sabal palmetto* |
| | Forested Flatwood Depressions (FD) | Areas containing mixed communities of *Taxodium, Pinus, Sabal*, bays or deciduous hardwoods in shallow depressions; may be located within areas of mesic hardwoods or pine flatwoods. |
| **Herbaceous Upland (HU)** | Dry Prairie (DP) | Areas containing > 50% cover of mixed upland grasses and other herbaceous species (e.g., *Aristida, Andropogon, Xyris, Rhexia*) with < 50% cover of low shrubs (e.g., *Serenoa, Ilex glabra, Lyonia, Vaccinium, Quercus minima*) and few trees. |
| | Pasture (PA) | Areas similar in composition to Dry Prairie, but with recent evidence of pasture/cattle management (e.g., fencing, feeding or water troughs) and the presence of exotic forage grasses (*Paspalum notatum, Hemarthria, Panicum repens, Cynodon*). |
| **Shrub Upland (SU)** | Palmetto Prairie (PP) | Areas containing > 50% cover of *Serenoa repens* and low shrubs (e.g., *Ilex glabra, Lyonia, Quercus minima, Vaccinium*) with few to no trees; < 50% cover of mixed upland grasses and herbs. |
| | Mixed Shrub (MS) | Areas containing 2: 50% mixed low shrub cover (e.g., Serenoa, Lyonia, Tlex glabra, etc.) interspersed with grasses and herbaceous vegetation with few to no trees. Includes overgrown pasture areas. |
| | Scrub (SC) | Areas containing 2: 70% cover of scrub species (e.g., *Ceratiola ericoides, Garberia fruticosa, Sabal etonia*) and scrub oak species (e.g., *Quercus geminata, Q. chapmanii, Q. inopina*), withor without an overstory of *Pinus clausa*; areas of bare white sand may be visible and should be included as part of the community. |
| **Forested Upland (FU)** | Sandhill (SH) | Areas containing < 50% cover of mixed upland grasses and forbs (e.g., *Aristida* spp., *Garberia, Licania, Pityopsis*) with widely spaced *Pinus palustris, Quercus laevis, Q. incana, and/or Q. stellata*; areas of bare sand may be visible and should be included as part of the community. |
| | Pine Flatwoods (PF) | Areas typically containing > 40% cover of *Pinus* spp. with an understory of low shrubs (e.g., *Serenoa repens, Ilex glabra*), grasses (e.g., *Aristida, Andropogon*) and forbs; where intensive burns have occurred, trees may be defoliated and understory vegetation may dominate. |
| | Pine Plantation (PI) | Areas where rows of planted pines (*Pinus elliottii, P. palustris, P. taeda, P. clausa*, etc.) are visible. Understory vegetation is variable. |
| | Upland Hardwood (UH) | Areas containing a mixture of hardwood species with pines < 40% and palms < 70%; may be dominated by mesic oak species (e.g., *Quercus virginiana, Q. laurifolia*). |