Determining aquatic vegetation in an upland sandhill lake:

Applications of hyperspectral image processing

NICK GROKHOWSKY & GRETCHEN STOKES

SUR 5386

University of Florida

Outline

I. Introduction

II. Methods
   A. Study Area
   B. Field Data Collection
   C. Hyperspectral processing

III. Results

IV. Discussion

Introduction: Threats to Freshwater Biodiversity

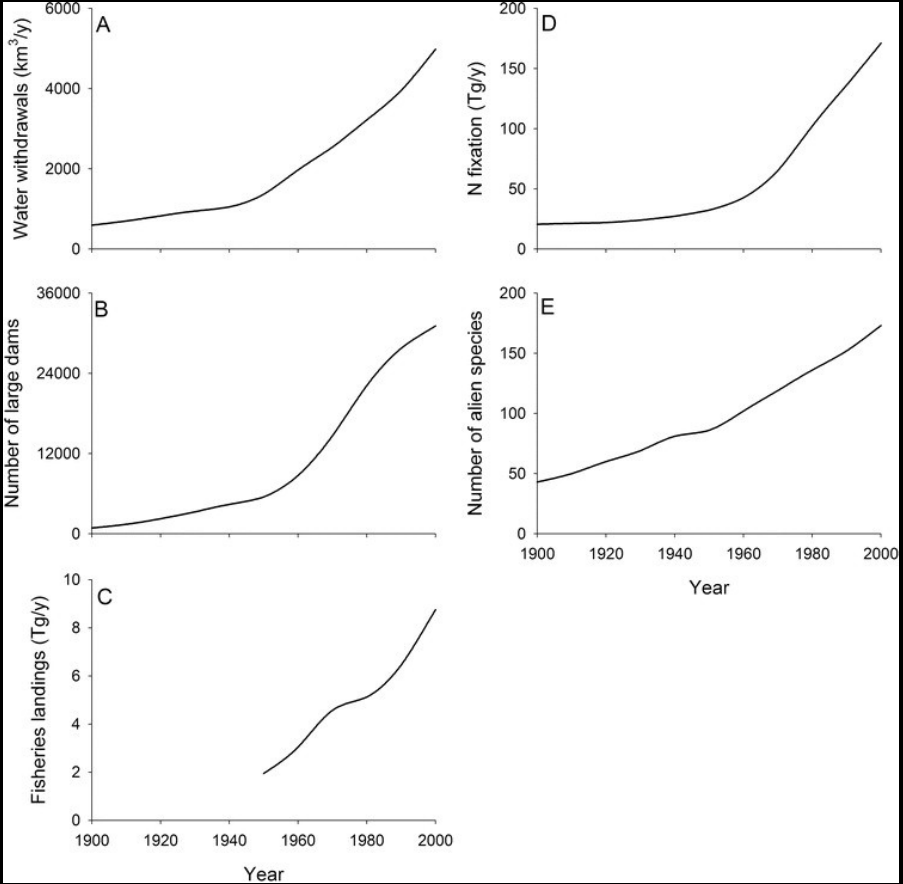
Increased Withdrawals

Large Dams

Overfishing

Increased Nitrogen

Invasive Species

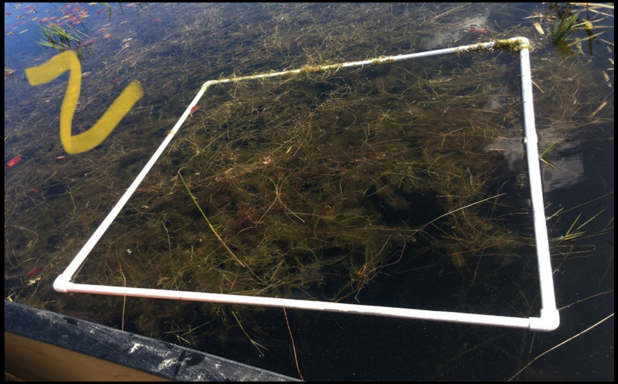


Introduction: Freshwater Monitoring

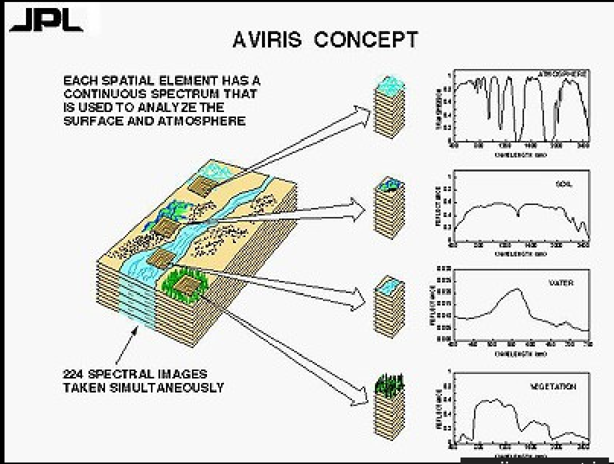
- Traditionally, water monitoring done manually = time consuming, costly, low efficiency

- Hyperspectral remote sensing (image spectroscopy) allows for identification of spectral signatures (e.g. oil, water)

- Reflectance can be used to assess water quality (e.g. phosphorus), species distributions (terrestrial vegetation), etc.



Manual field data collection takes considerable time and effort



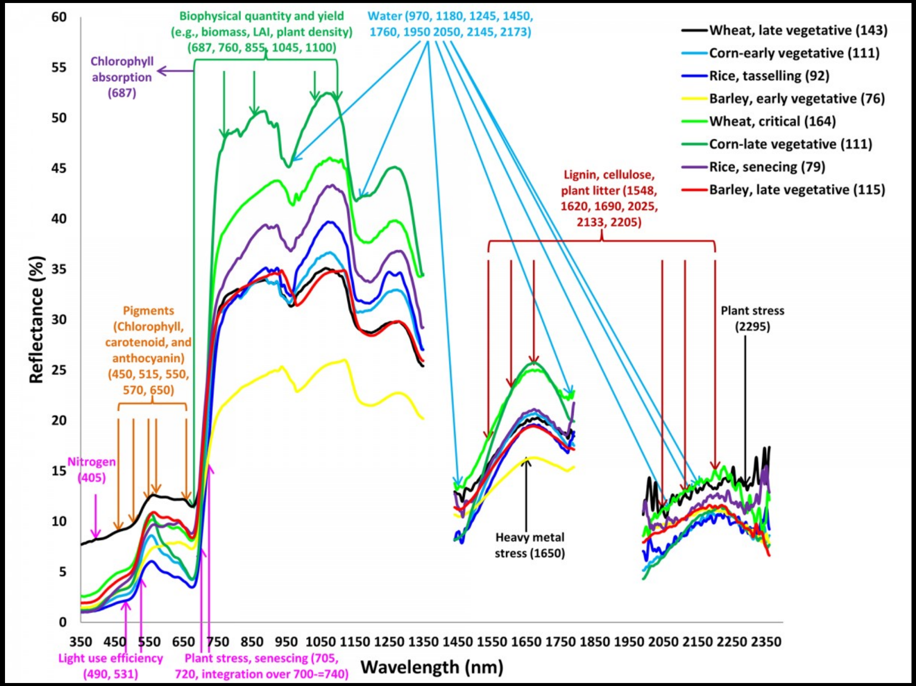
ces.iisc.ernet.in

Introduction: Drones and Remote Sensing

- Terrestrial: Sensors helpful for species identification

- Aquatic: Challenges in differentiating spectral reflectance of submerged, floating and emergent plants

- Drones equipped with hyperspectral sensors can generate finer spectral resolution at a larger scale



USGS

Objectives

Research question: How effectively can hyperspectral images be used to:
a) distinguish submerged aquatic vegetation (SAV) from water, and
b) identify different categories of SAV?

Research Objectives:

1. Learn and demonstrate the process of converting hyperspectral image file to use in ENVI
2. Apply corrections learned in class to improve images
3. Apply supervised and unsupervised classifications and determine the best method for determining vegetation

Study Area

<u>Ordway Swisher Biological Station</u>

Putnam County, FL

Lake McCloud

- Upland sandhill lake
- 200 m x 150 m
- Surrounded by hardwood-pine forest
- Seasonally flooded, drains to Lower St. Johns River basin

Methods

<u>GatorEye Unmanned Flying Laboratory (GE-UFL)</u>

- Headwall Photonics VNIR 270 spectral band hyperspectral sensor

- Geolocated to +/- 2.5 cm using dual frequency GNSS and post-processed kinematic algorithms to a base station

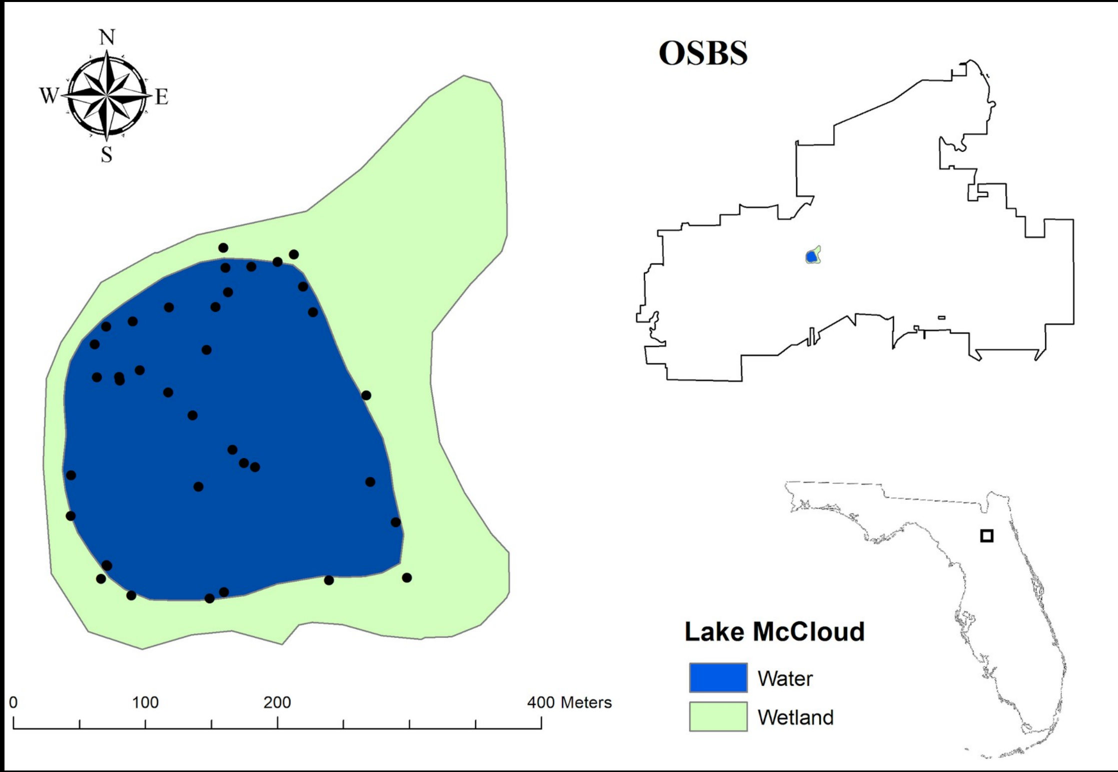
Methods

GatorEye Unmanned Flying Laboratory (GE-UFL)

January 30, 2018 Flight Parameters:

- 2 missions flown
- 15 minutes each
- Total of 9 transect lines, 25 m apart
- 60 m above ground level
- 125 points per second (6m/s)

Field Surveys



Pre-Processing:

Geometric Correction

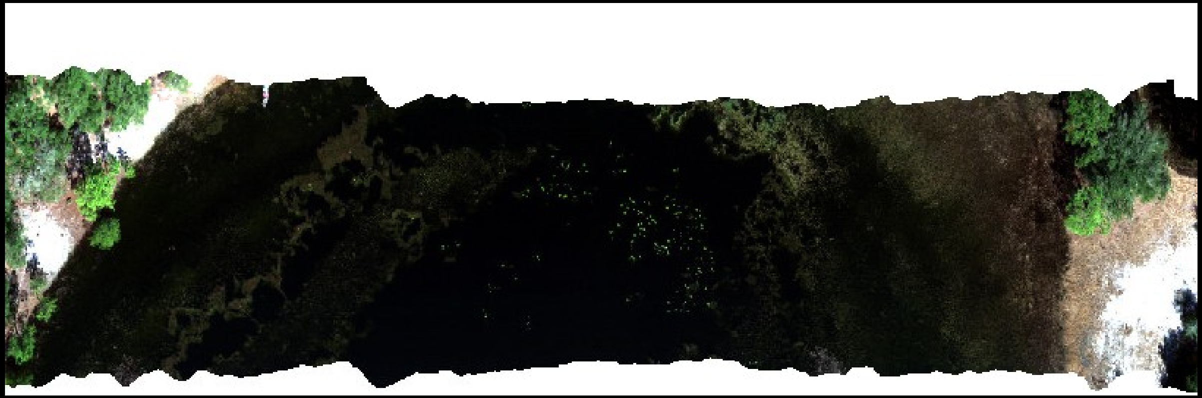
1. Distortion correction

2. Sensor altitude effect: roll, pitch, yaw

3. Spatial interpolation: Ortho rectification

4. Export corrected radiance files

Image Processing



Methodology:

- Spatial Subset - geographic points, ROI mask
- Spectral Subset - PCA
- Training Data
- Neural Network Classification
- Accuracy Assessment
- IsoData Classification
- Band Ratioing - NDVI
- Unmixing
- Interpretation and Comparison

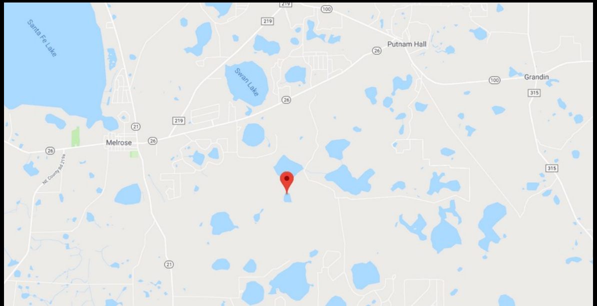
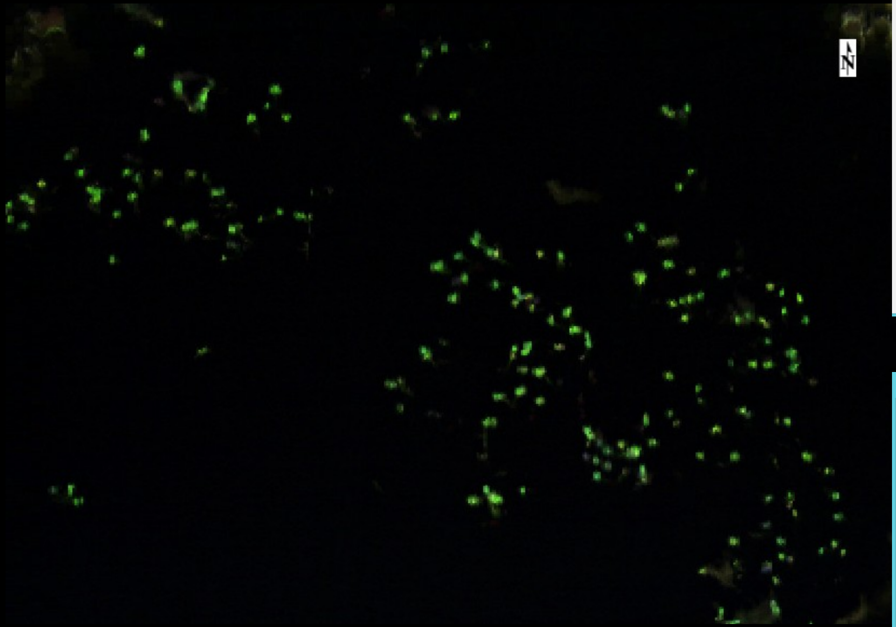
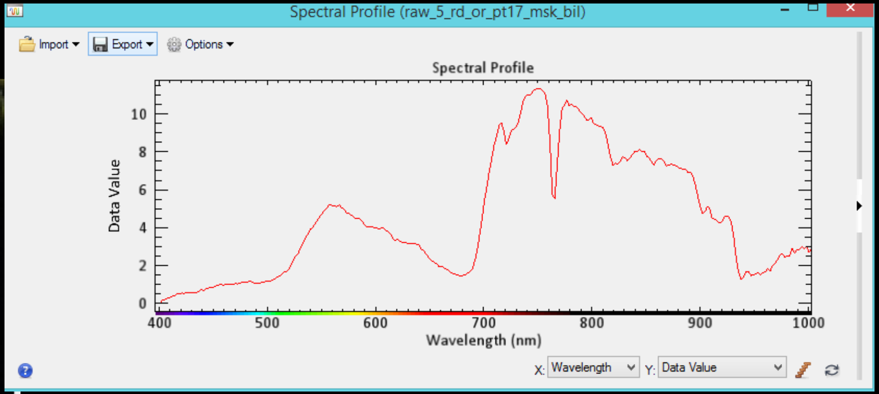


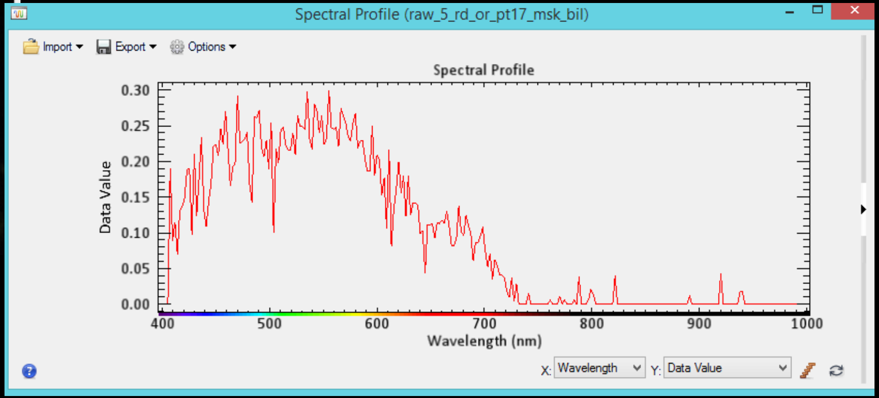
Image Exploration



Green = vegetation
Black = water

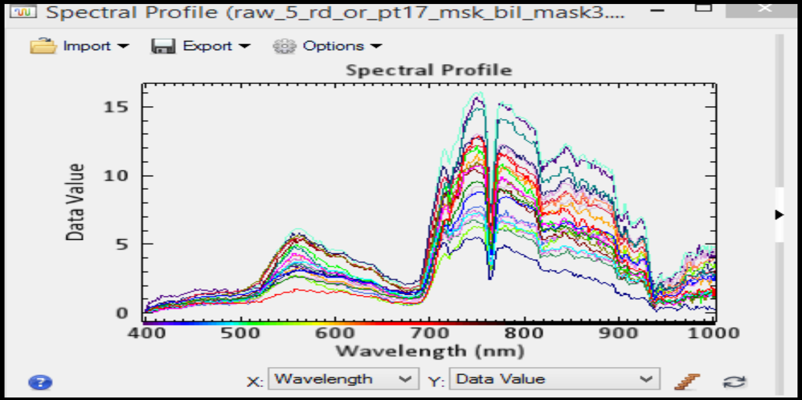


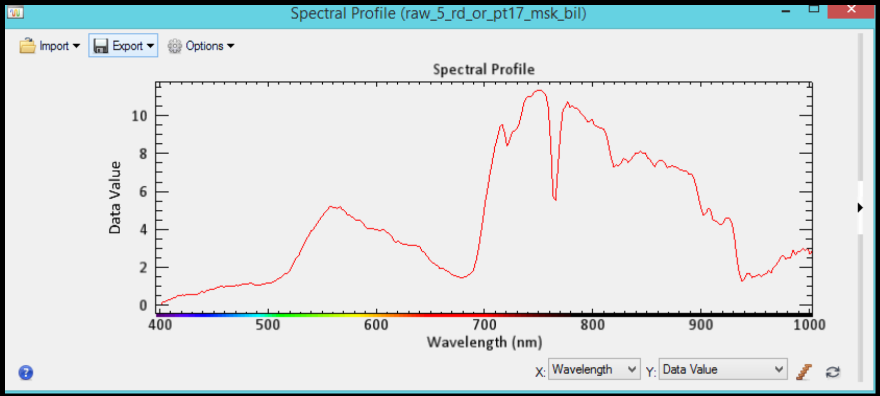
Vegetation



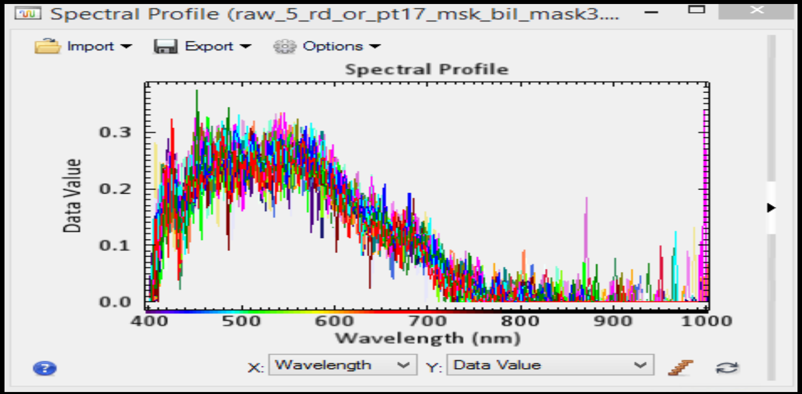
Water

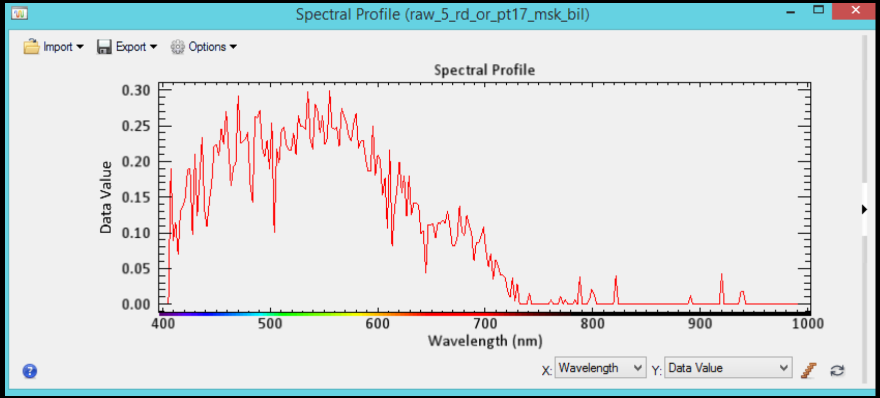
Image Exploration





Vegetation





Water

Spatial Subset

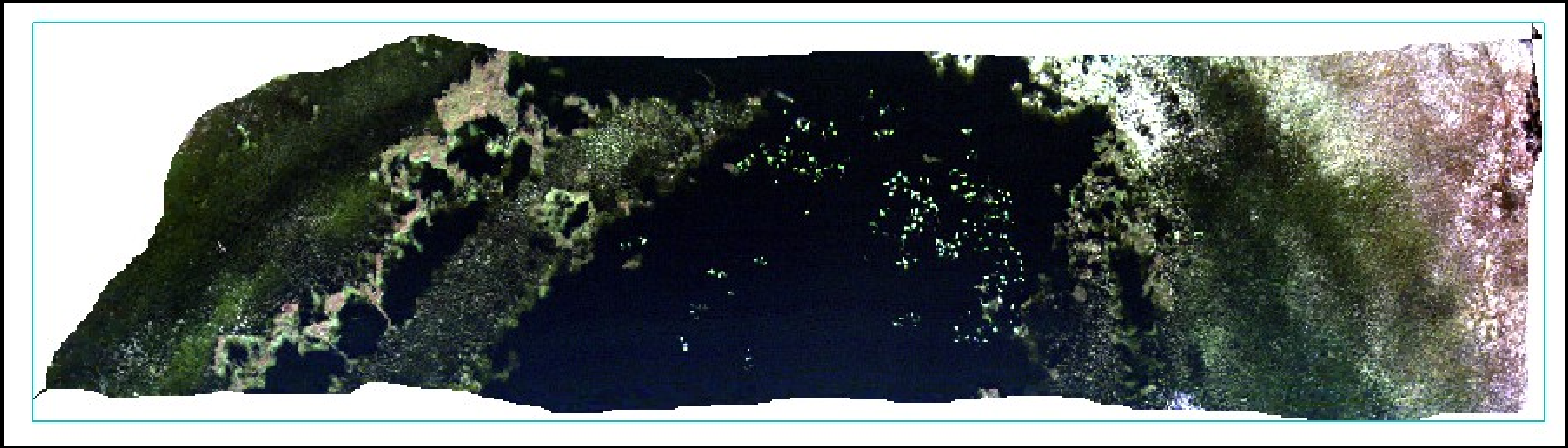
Due to the large image size (spatially and spectrally), we felt it was necessary to reduce the image to a manageable study area.

The method we used to complete this spatial subset was to use the file -> save as -> spatial subset function.

Within the spatial subset function we used chosen geographic coordinates to reduce the image to just the lake. This region of interest was the section of the lake we are studying (all forest-only sections excluded).

Spatial Subset - Alternate Option

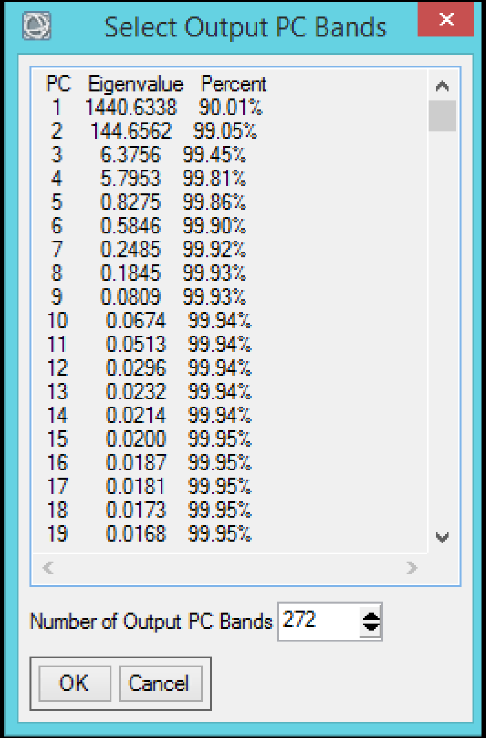
We also experimented with creating an ROI, saving it as a “mask” and then saving the BIL file with a spatial subset using the mask ROI and a mask as the same ROI file. We came up with very similar results as the spatial subset, though slightly less accurate because we drew the ROI by hand.



Principal Component Analysis

The PCA was used because the hyperspectral image contains 272 bands. Performing the PCA would identify which bands contained the majority of the variation compared to the bands that contain less variation, and therefore more noise.

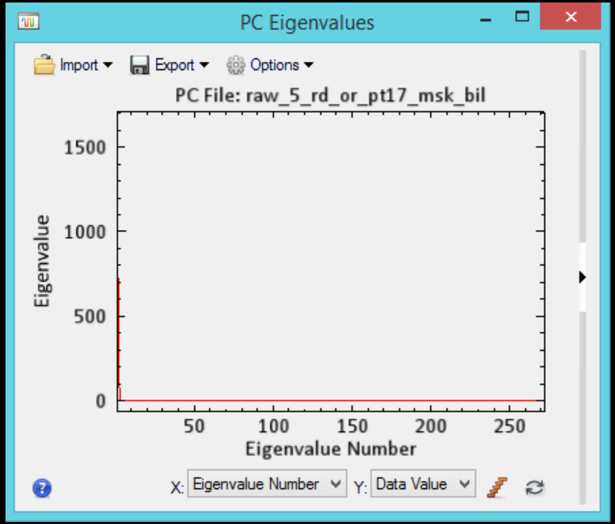
The outcome of the PCA was that the top two bands contain more than 99% of the variation.


Select Output PC Bands

PC	Eigenvalue	Percent
1	1440.6338	90.01%
2	144.6562	99.05%
3	6.3756	99.45%
4	5.7953	99.81%
5	0.8275	99.86%
6	0.5846	99.90%
7	0.2485	99.92%
8	0.1845	99.93%
9	0.0809	99.93%
10	0.0674	99.94%
11	0.0513	99.94%
12	0.0296	99.94%
13	0.0232	99.94%
14	0.0214	99.94%
15	0.0200	99.95%
16	0.0187	99.95%
17	0.0181	99.95%
18	0.0173	99.95%
19	0.0168	99.95%

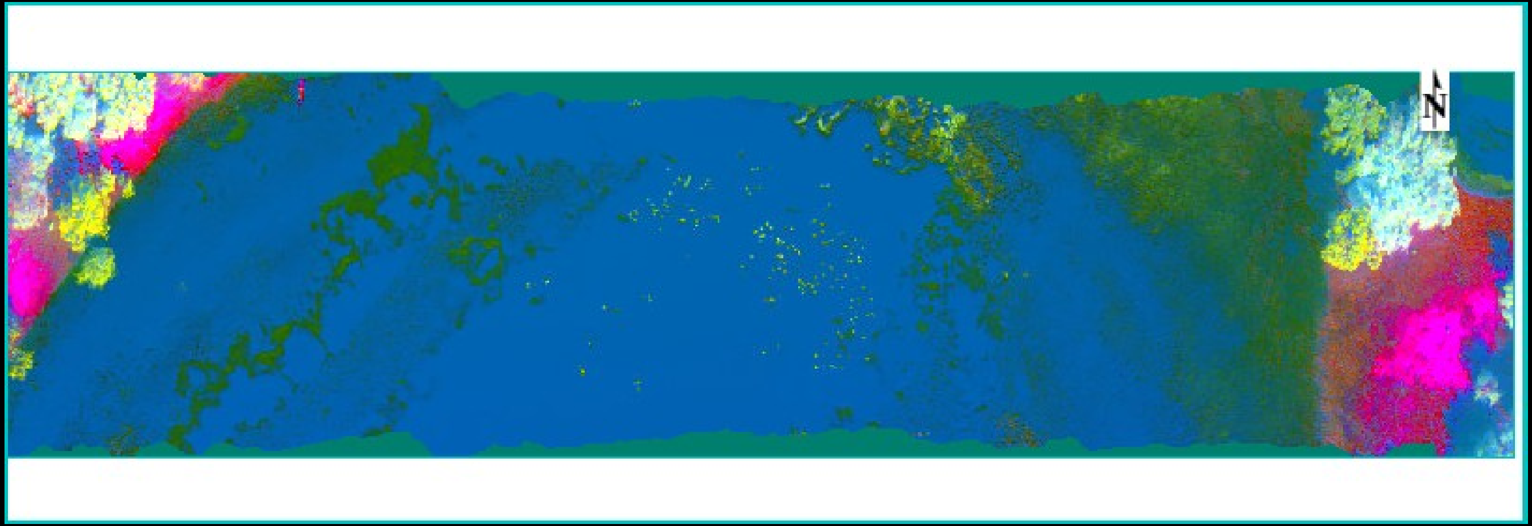
Number of Output PC Bands 272

OK Cancel



Principal Component Analysis

The first three bands were selected because they contain 99.45% of the variance in the entire image. When these three bands are applied to an RGB image it shows a very clear delineation of colors being reflected. Our primary subject of this study is the aquatic vegetation. It is interesting to see that the aquatic vegetation reflects somewhere in the yellow-green spectrum. This is discussed in *Signal Classification of Submerged Aquatic Vegetation Based on the Hemispherical–Conical Reflectance Factor Spectrum Shape in the Yellow and Red Regions*, where it is explained that traditional methods of terrestrial vegetation indexes are not accurate in sensing submerged aquatic vegetation. They went further to discuss that it is in the range of 585 nm - 685 nm that submerged aquatic vegetation and algae are seen (1).

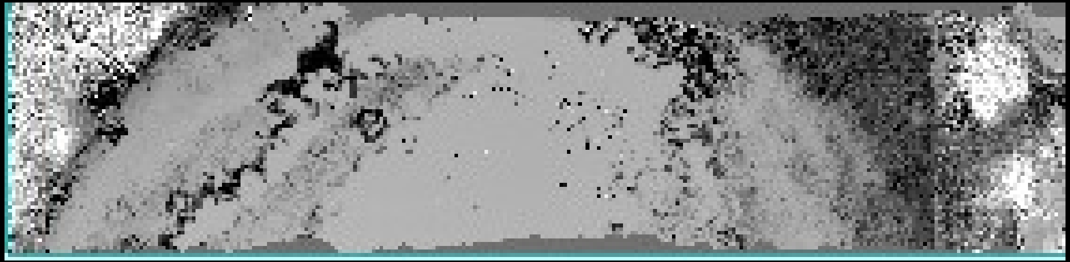


Principal Component Analysis

BAND 1: IR and green spectrum

BAND 2: blue and yellow spectrum

BAND 3: red spectrum



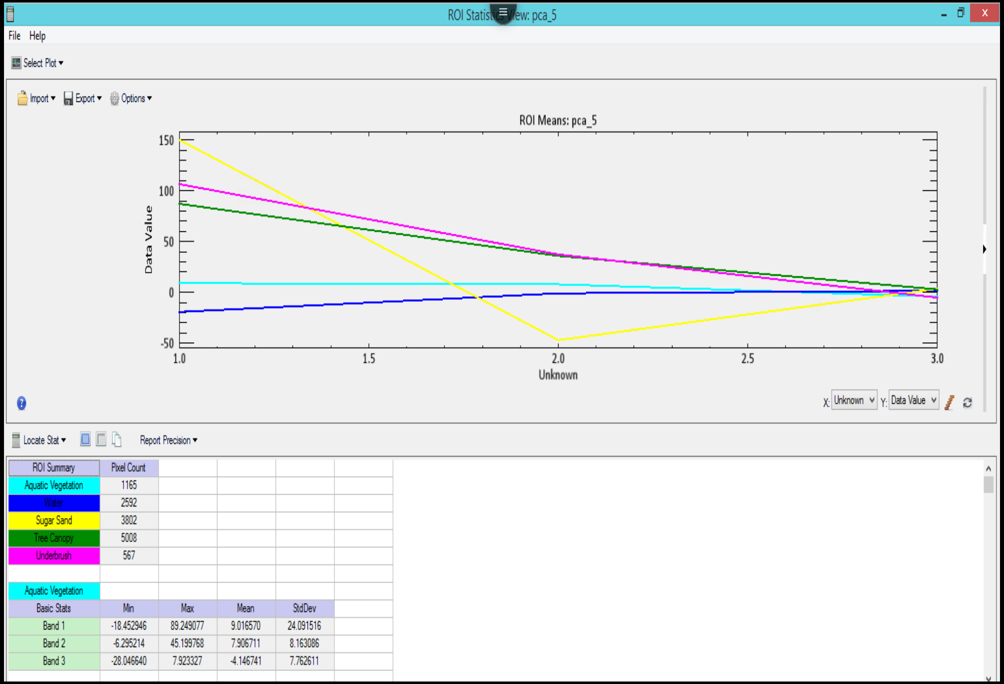
Principal Component Analysis

Although it is clearly seen in the image, the majority of the features the image contain are either water or vegetation according to the outcome of the PCA. This is identified by the first band that the PCA captured, 90.01% of the variation, and reflected the green and infrared bandwidths. This identifies vegetation as the primary source of reflectance in the image. The second band of the PCA captured 9.04% of the variation and reflected the blue and yellow spectrum. This identifies water and possibly soil in the blue spectrum, but through visual inspection we know water is the majority of the two. It is the yellow reflectance that is going to be of interest to us because it potentially will be the key in identifying whether vegetation is aquatic or not.

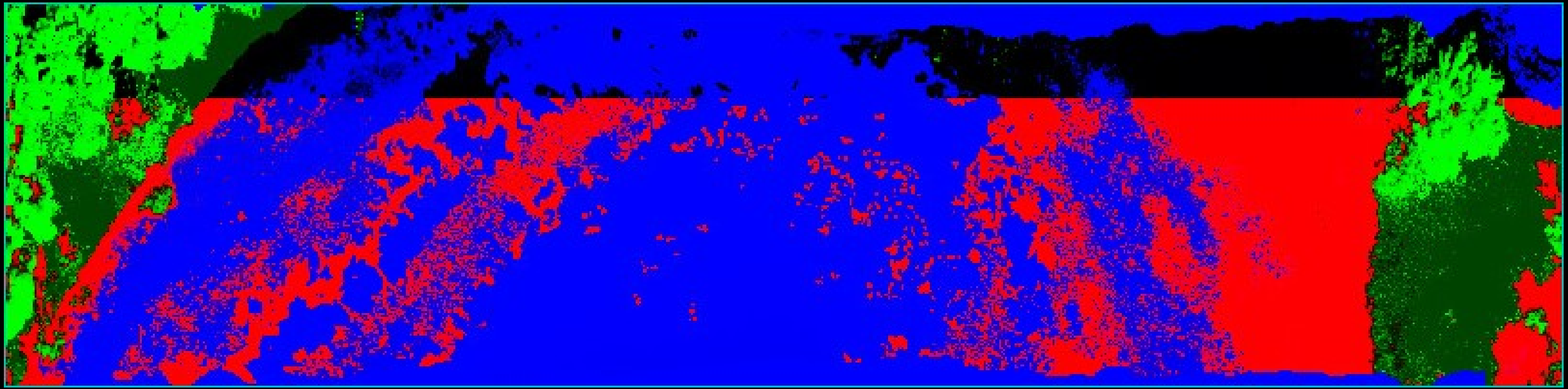
Training Data - Regions of Interest

First Iteration:

- Aquatic Vegetation
- Water
- Terrestrial Vegetation
  - Tree Canopy
  - Underbrush
- Sugar Sand

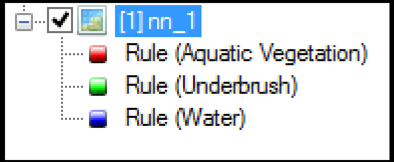


Neural Network Classification

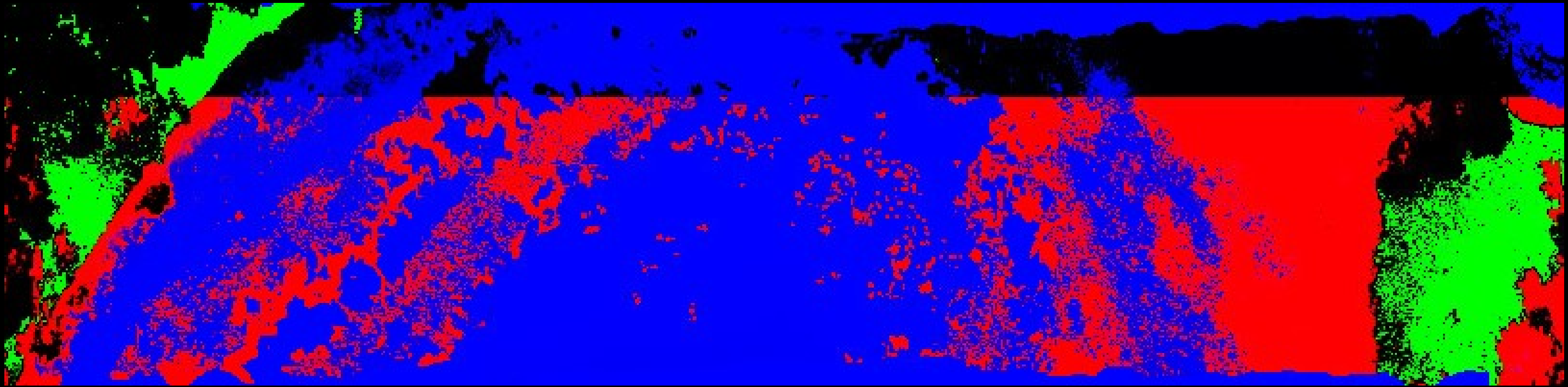


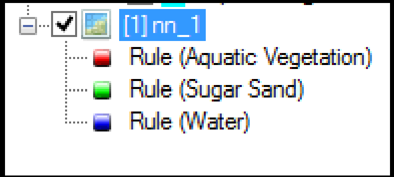
The first iteration of the neural network classification has a dark unclassified section toward the north of the image.

This first representation shows aquatic vegetation, water, and underbrush.



Neural Network Classification

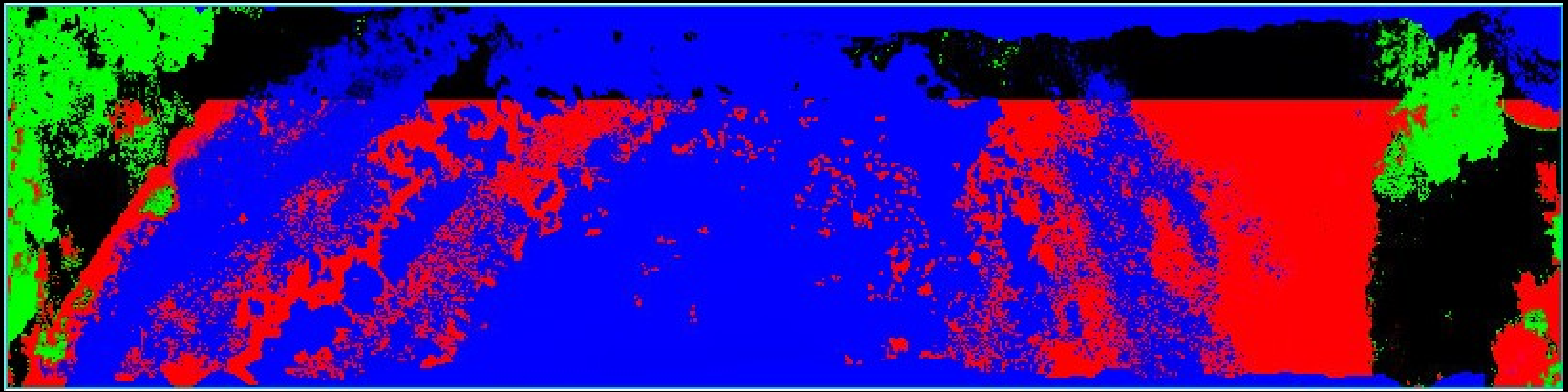




The first iteration of the neural network classification has a dark unclassified section toward the north of the image.

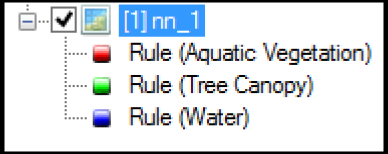
This second representation shows aquatic vegetation, water, and sugar sand.

Neural Network Classification



The first iteration of the neural network classification has a dark unclassified section toward the north of the image.

This third representation shows aquatic vegetation, water, and tree canopy.



Neural Network Classification

The first run of the neural network classification was decent, but some corrections need to be made, and the data needs to be teased a little. The separation between terrestrial vegetation and sugar sand was successful for the most part, but the separation among tee canopy and underbrush was not successful. Also, the aquatic vegetation was successfully identified, but we would like to take this a step further and identify submerged versus unsubmerged aquatic vegetation. Lastly, we need to classify the dark patches along the top of the image.

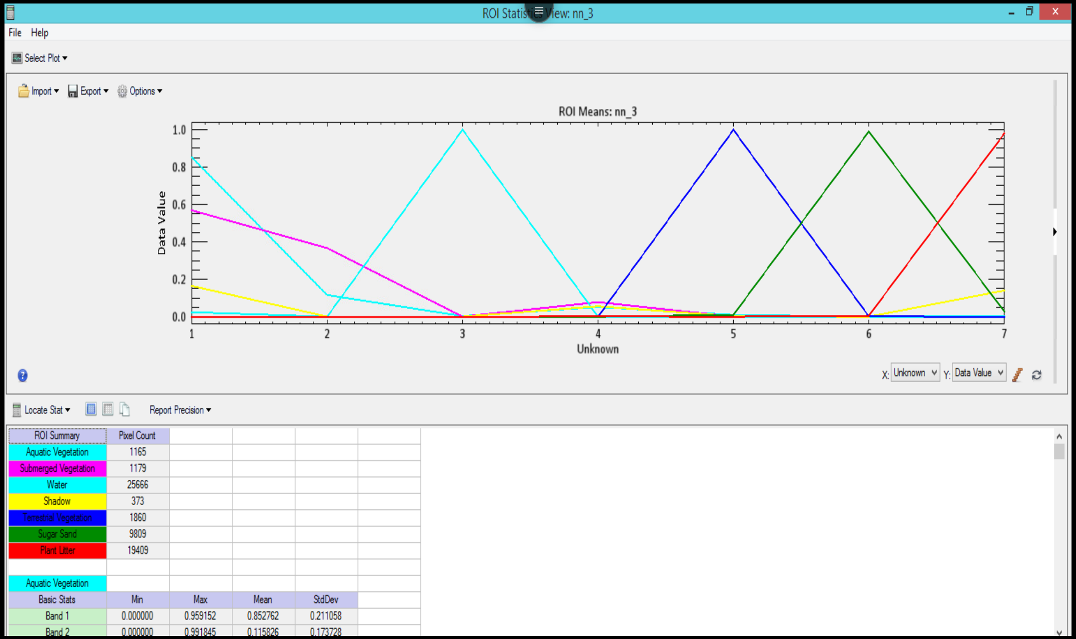
Next Steps:

- Group terrestrial vegetation regions of interest
- Separate submerged and unsubmerged aquatic vegetation

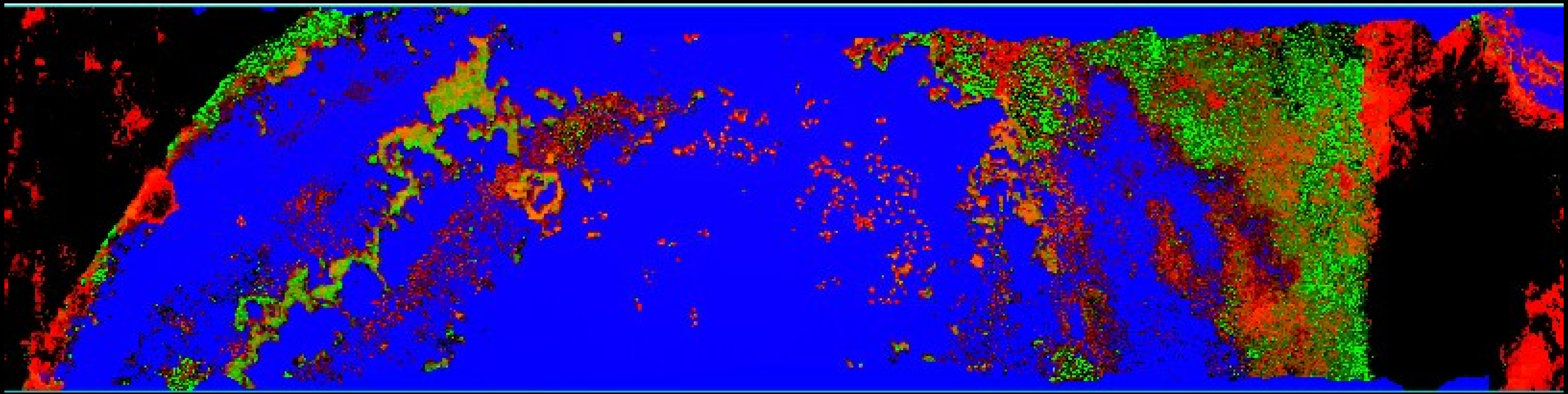
Training Data - Regions of Interest

Second Iteration:

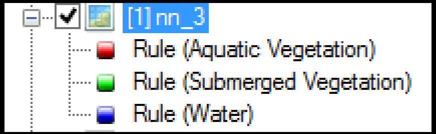
- Aquatic Vegetation
- Submerged Vegetation
- Water
- Terrestrial Vegetation
- Sugar Sand
- Plant Litter



Neural Network Classification



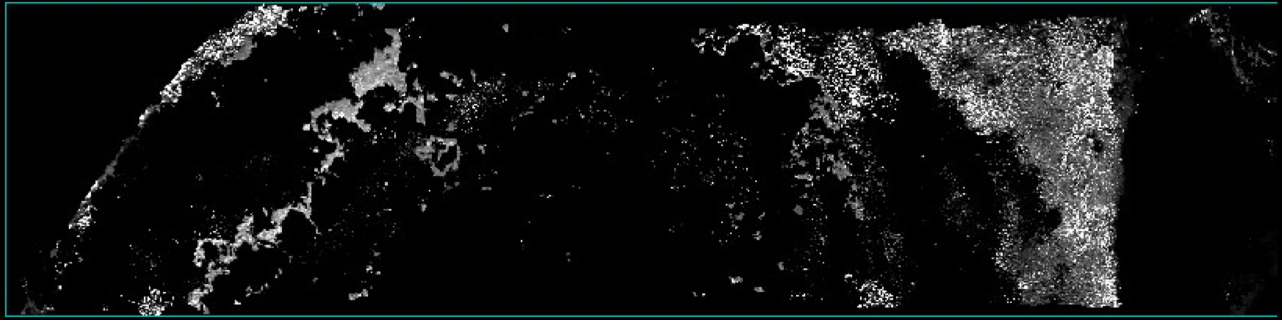
The reclassification made significant changes. Although the red classifier is titled ‘Aquatic Vegetation,’ it could be considered ‘Unsubmerged Vegetation.’ This classification looks very promising considering the large regions of ‘Submerged Vegetation,’ and their proximity to ‘Unsubmerged Vegetation’ in the middle of the ‘Water.’



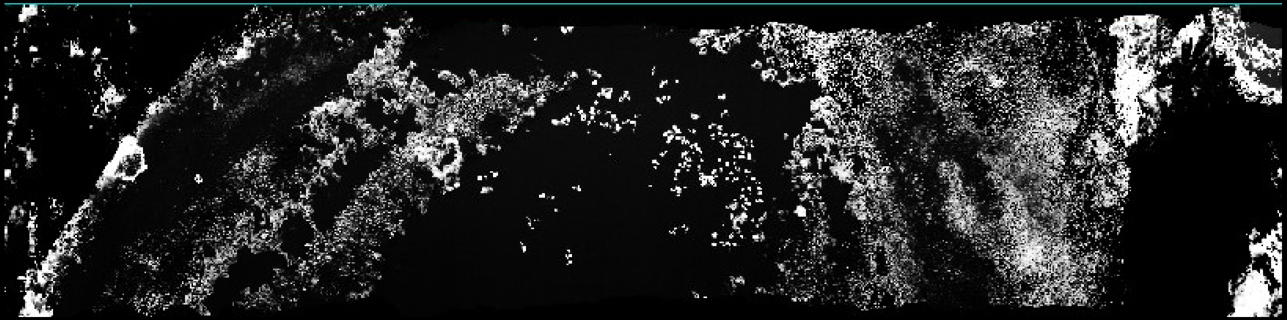
One piece to point out is that the unsubmerged vegetation also captured terrestrial vegetation because it is unsubmerged as well.

Neural Network Classification

Submerged Vegetation



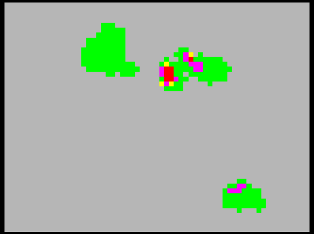
Unsubmerged Vegetation

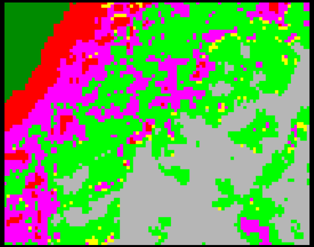


Neural Network Classification

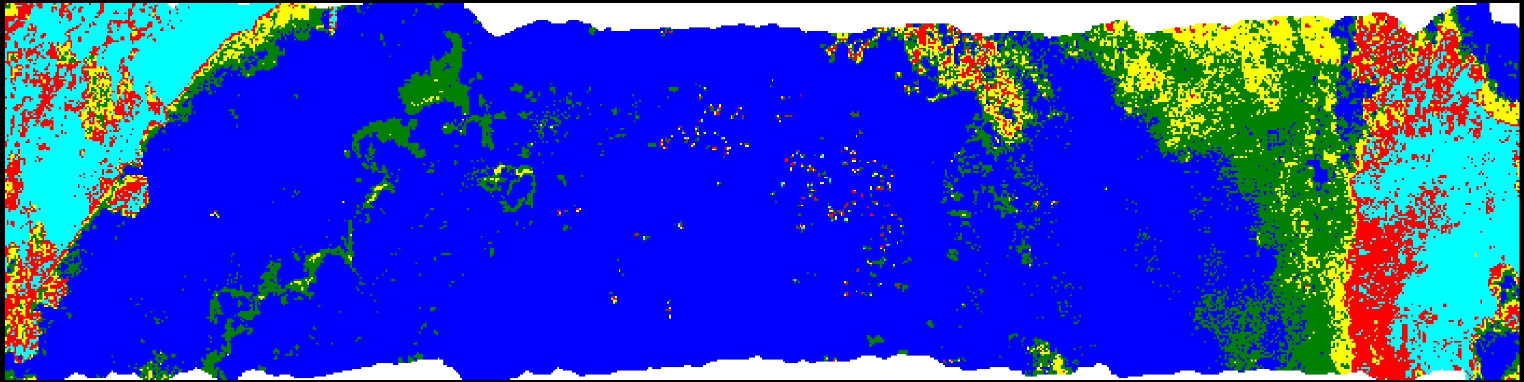
Accuracy Assessment:

Particularly for supervised classifications, it is important to check how accurate our estimates were and compare them to actual field data. We did this using the GPS trimble points collected by canoe during the aquatic vegetation surveys. We compared the percentage vegetation at each point (1m x 1m), plus georeferenced photographs of each point, to determine how well the supervised classification worked. We see that generally it was able to identify vegetation correctly; however there were a few times where accuracy may have been compromised because of dark colored vegetation under the water, or lighter grasses. Overall, we feel confident to say that these results are very accurate though not perfect (perhaps 80% or more accuracy).

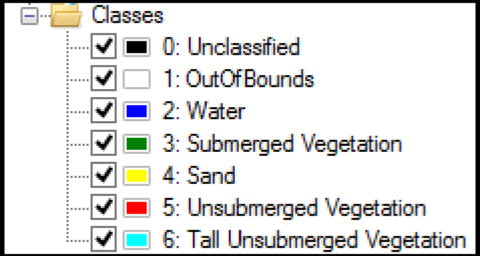




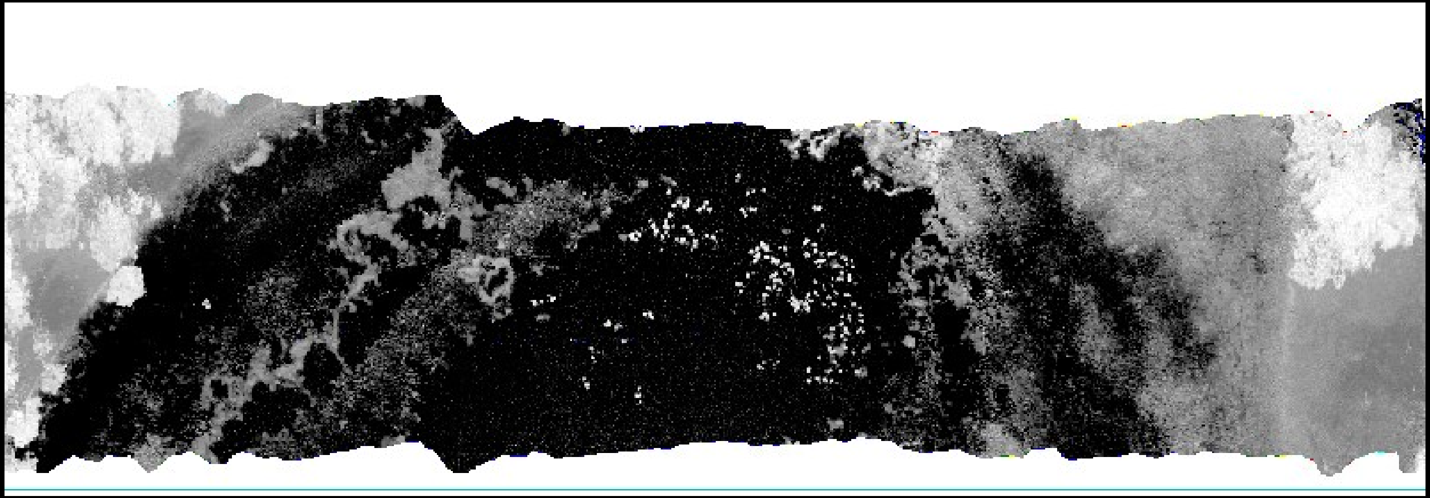
IsoData Classification



The unsupervised method using IsoData Classification appears to be accurate in the first instance of running the method. This was done with 100 iterations and 5-10 potential classifications. It looks like it has identified submerged versus unsubmerged vegetation accurately.



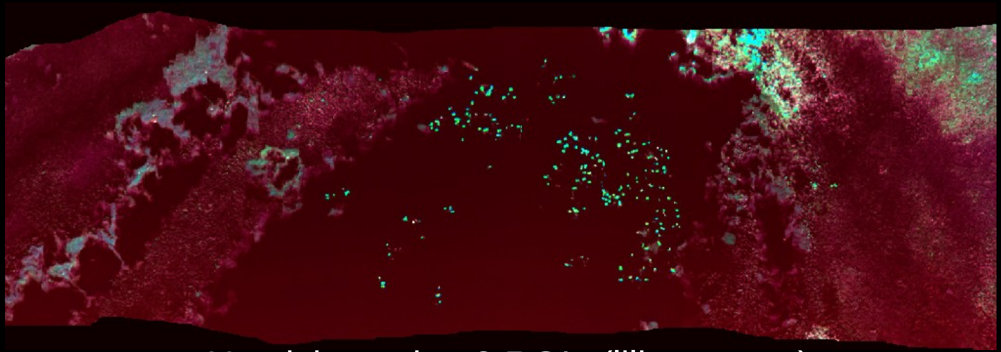
Band Ratio - NDVI



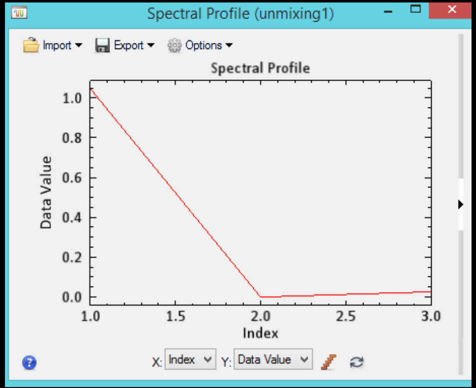
Although it had been mentioned by *Signal Classification of Submerged Aquatic Vegetation Based on the Hemispherical–Conical Reflectance Factor Spectrum Shape in the Yellow and Red Regions* that NDVI is a bad measure for submerged aquatic vegetation, we felt it would be a good idea to compare the NDVI to test the accuracy of the statement.

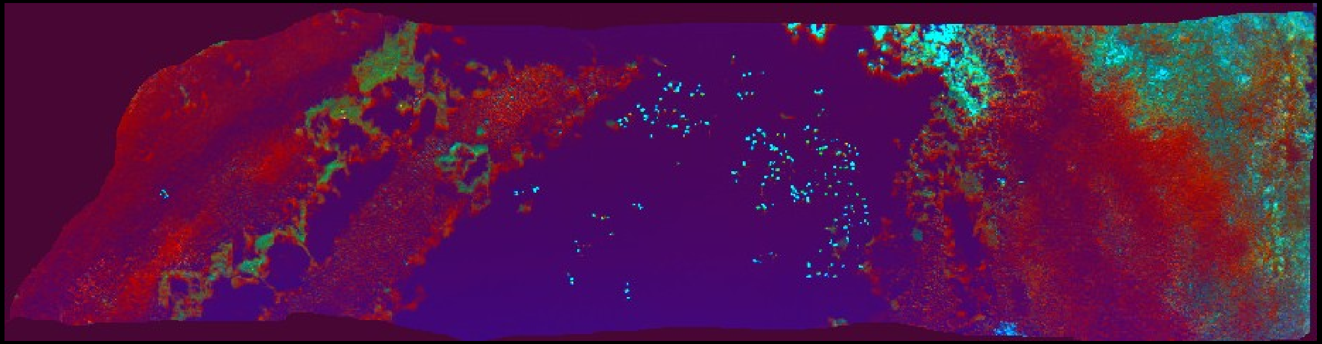
It is clearly seen here that the terrestrial vegetation and unsubmerged vegetation are brighter values in comparison to the areas that we found to be submerged vegetation.

Unmixing

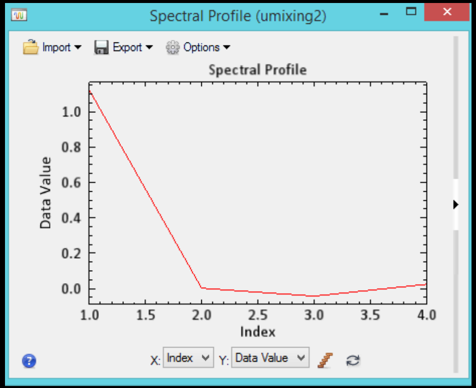


Unmixing using 2 ROIs (lilies, water)

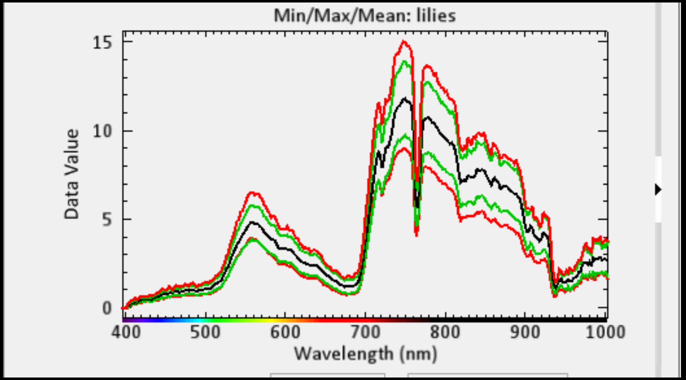




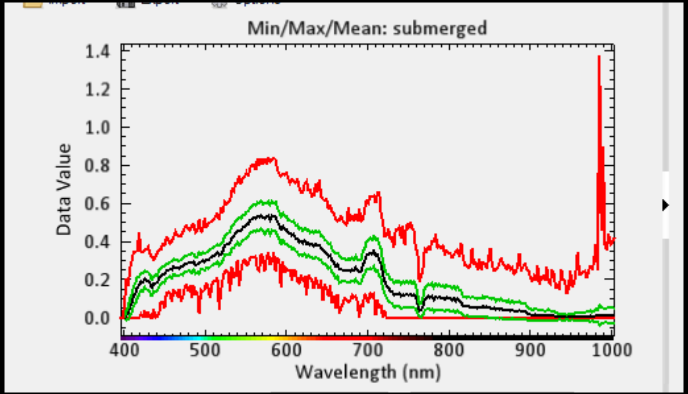
Unmixing using 3 ROIs (submerged aquatic vegetation, floating vegetation, water)



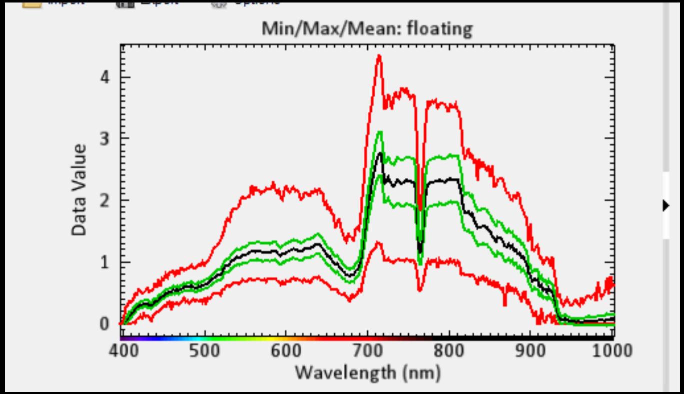
Unmixing



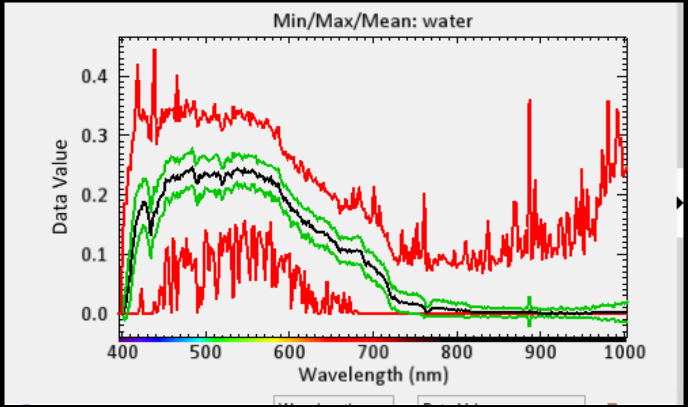
Min/Max/Mean: lilies
Data Value
Wavelength (nm)



Min/Max/Mean: submerged
Data Value
Wavelength (nm)



Min/Max/Mean: floating
Data Value
Wavelength (nm)



Min/Max/Mean: water
Data Value
Wavelength (nm)

Discussion

- We observed what we expected: that we would be able to identify aquatic vegetation from water and even down to different types of vegetation (submerged and floating).

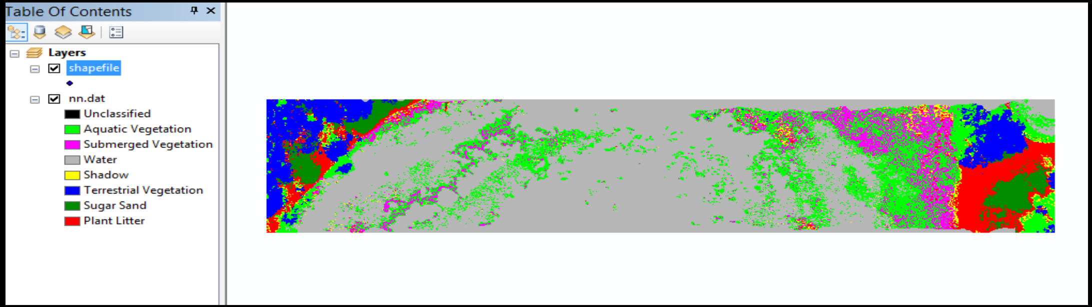
- However, we are not able to get it down to species, due to highly similar profiles of different species with similar appearances

- This tells us that hyperspectral data may be appropriate for some types of questions and data collection, while if the specific focus is on plant species, we may not yet have the ability to distinguish that.

- Overall, we see broad applications of the use of hyperspectral drone data for aquatic plant classification

Discussion

- We sometimes think that supervised classifications may be more effective because we have control over training the program to correctly identify points.
- However, that is not the case here. We actually see improved (and regularly with sub-meter accuracy) identification of vegetation types versus water in the IsoData classification
- This leads us to believes that in fact the unsupervised classification techniques may be not only more efficient but also more accurate in our case
- This does not, however, mean that IsoData is always best, but we believe it to be best in our study due to constraints of training for supervised classifications



Discussion

- Applications of these results include using drones for aquatic vegetation classification and detection in remote, difficult-to-access, or potentially areas that would be unsafe on the ground. In a total of 30 minutes of flying, we collected highly accurate data that provides even more information, with equal or greater accuracy, than four hours by canoe.

- This may also prove useful in very large expanses of wetlands that would be extremely time-costly to access by water (e.g. Pantanal, Everglades, etc).

- Furthermore, it would be interesting to see invasive species could be accurately distinguished from all others?

- This would also set a high-quality standard and baseline in questions related to climate change and invasive species that may greatly affect watersheds or aquatic ecosystems over time.

Conclusion

- Aquatic vegetation <u>can be</u> distinguished from water
- It is also possible to distinguish between types of aquatic vegetation - in this case, submerged and floating
- However, getting down to the species level remains difficult for submerged aquatic vegetation
- In our study, IsoData was the most successful and efficient process leading to very accurate results; however in other circumstances, Neural Network classification can also be highly useful
- Future studies should focus on areas with a high diversity of highly distinguishable / identifiably different species

Thank you!!