Wetland vegetation mapping with Hyperspectral Imagery using OBIA.

Madan Thapa Chhetri[1,2]

[1]Department of Geosciences, Florida Atlantic University
Florida, Boca Raton, 33486, and USA
[2]Department of Geomatics Engineering, Western Regional Campus, Tribhuvan University
Pokhara, Lamachaur, 33700, Nepal
Email: mthapachhetr2023@fau.edu; chhetrim79@gmail.com

Keywords: OBIA, Machine Learning, Wetland Vegetation, ArcGIS Pro, Accuracy Assessment

Abstract:

Hyperspectral Remote Sensing imagery combined with Object-Based Image Analysis (OBIA) offers a robust method for mapping wetland vegetation. This study explores the classification of wetland vegetation focusing on the Fort Drum Marsh Conservation Area (FDMCA) within the St. Johns River Water Management District using OBIA and Machine learning techniques. The research utilizes Hyperion EO-1 imagery acquired in 2003 and reference datasets to classify eight wetland vegetation communities. Utilizing Random Forest (RF) and Support Vector Machine (SVM) algorithms, the study achieves overall accuracies of 83.6% and 80.5%, respectively, with RF demonstrating higher performance. However, challenges such as misclassifications persist, particularly in distinguishing between similar vegetation classes. Despite limitations inherent in coarse-resolution imagery and potential misclassifications, the study highlights the efficacy of OBIA integration with machine-learning techniques for wetland vegetation mapping. It emphasizes the significance of post-classification assessment for improving accuracy and identifies areas for future improvement.

Introduction:

Satellite-based imagery is an essential approach for examining changes and mapping wetlands. Using hyperspectral imagery makes it easy to detect patterns and discern the types of vegetation or land cover categories. It provides high spectral resolution data, allowing it to detect subtle differences in vegetation. This capability is precious in wetland environments, where vegetation plays a significant role in ecological functions. It gives opportunities to obtain more scrutinized information than other multispectral data. This technology has been promising in geospatial research, monitoring, and exploration applications (Shipert, 2004) and wetland vegetation classification (Zhang & Xie,2013).

Classification of wetland vegetation using hyperspectral remote sensing and Object-Based Image Analysis (OBIA) can be a convincing means for monitoring and managing wetlands. OBIA offers a more precise and detailed analysis compared to pixel-based methods. It is commonly known that a pixel-based classification approach leads to an effect known as "salt and pepper" during mapping heterogeneous landcover types. This effect can be mitigated using OBIA and is valuable and promising for vegetation classification (Zhang & Xie,2012). It operates by grouping similar pixels into 'objects,' which can then be classified based on spectral and spatial characteristics. This method enhances the accuracy and reliability of wetland vegetation change detection, facilitating more effective wetland management and monitoring. Wetlands offer numerous ecosystem benefits, but their historical documentation has been inadequate. However, researchers have made significant progress in addressing this gap through remote sensing technology and a method known as change detection (Mahdianpari et al.,2021).

Wetlands contain many ecosystems and applications, including carbon sequestration, water supply and monitoring, wildlife and vegetation, and flood and sedimentation control (Davidson et al., 2019). Vegetation species found in wetland environments are challenging to identify due to low accessibility. Wetlands are crucial in preserving and conserving critical habitats and monitoring water quality. Mapping the vegetation of wetlands has been made easier with recent advancements in remote sensing technology. Using deep learning techniques and machine learning has also helped solve half of the problems associated with vegetation classification (Jafarzadeh et al., 2022). A meta-analysis of 30 years of research has shown that data obtained from remote sensing and machine learning approaches are valuable for wetland observing and multi-representation findings. They may open newfound viewpoints for research studies and advance scientific support for management decisions.

Support Vector Machine (SVM) and Random Forest (RF) are popular machine-learning algorithms for supervised classification because of their high accuracy than other methods for land cover (Nery et al., 2016; Chhetri & Rijal,2023) and vegetation mapping (Sabat-Tomala et al., 2020). SVM is a supervised classification algorithm designed to obtain a hyperplane that effectively divides the input dataset into distinct, pre-defined categories, aligning with the patterns observed in the training data. This approach is widely employed in numerous real-world classification problems due to its strong theoretical foundation and superior performance in terms of generalization (Cercantes et al., 2020). On the other hand, RF is an ensemble classifier based on decision trees (Brieman,2001) representing a sophisticated and powerful machine-learning model.

Generally, vegetation species found in the wetland environment are challenging to reach and identify. Moreover, wetlands are one of the essential components that help in the preservation and conservation of critical habitats and water quality monitoring. With the difficulties in identifying the plant species due to low accessibility, it is hard to map the vegetation of wetlands. However, with the recent advancements of remote sensing technology, mapping of wetland vegetation has become relatively easy. Along with remote sensing imagery, modern machine learning and deep learning have solved half of the problems of vegetation classification. In this study, I aim to analyze the classification of wetland vegetation using OBIA and Machine learning algorithms. This study will utilize Hyperspectral Remote sensing imagery and carry out the supervised classification of wetland vegetation, and areas of classified vegetation will be quantified and analyzed. This process helps to save labor-intensive and manual interpretation, which may take a long time.

Materials and Methods

Study Area

For this research, I utilized the portion of the Fort Drum Marsh Conservation Area (FDMCA) wetland in the St. Johns River Water Management District (SJRWMD) (Figure 1). This conservational area contains a combination of wetland and upland communities jointly acquired by the U.S. Army Corps of Engineers and the Upper St. Johns River Basin project for controlling floods and revitalizing the basin (SJRWMD, 2022). This area has diverse wetland communities, which are vital for improving water quality and enhancing or restoring wetland habitat. Wetland vegetation communities within this area consist of wet prairie, hardwood swamp, cypress, shrub swamp, and many more.

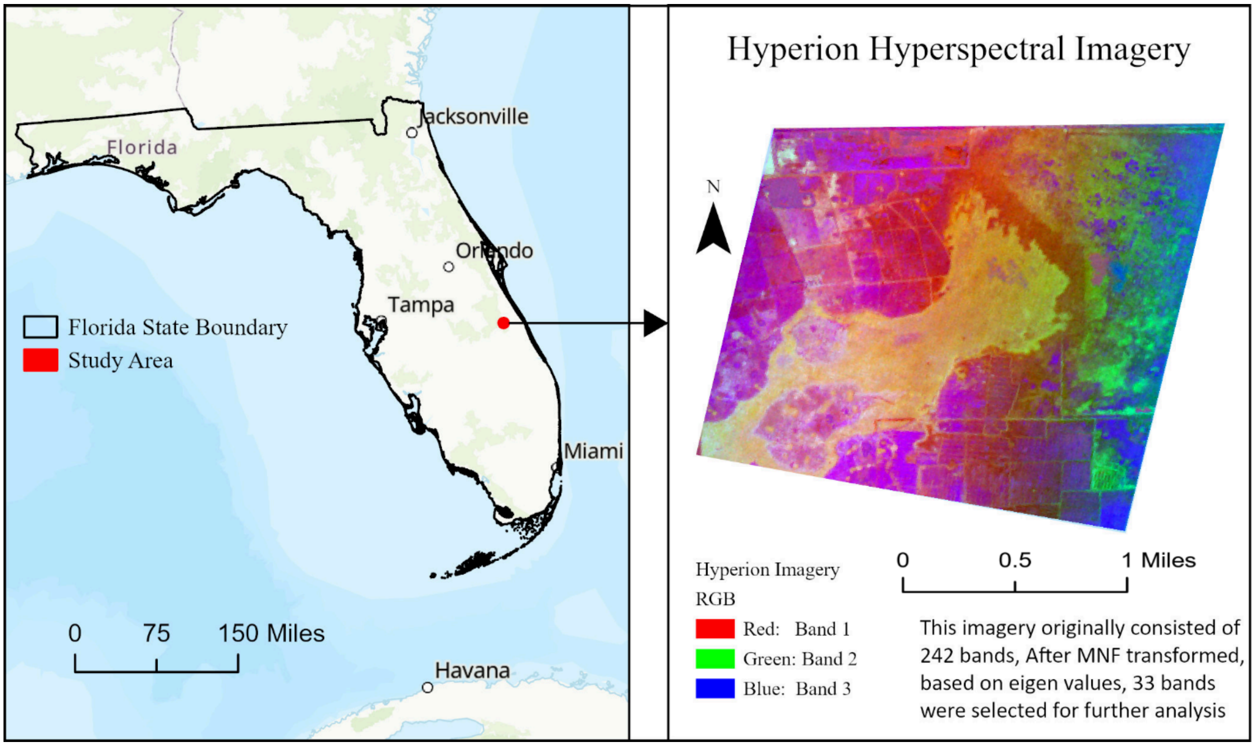

Figure 1 Map Showing the Study area at FDMCA, illustrating Hyperspectral imagery with three MNF transformed bands on it.

Data

Data sources include hyperspectral imagery EO-1/Hyperion and reference datasets. Hyperspectral imagery was collected in 2003/05/04 by the Hyperion Imaging Spectrometer onboarded on the EO-1 spacecraft. The orbit path was 15, the Row was 41, and the target path and row were 15 and 41, respectively. This sensor has 242 contiguous spectral bands with wavelengths of 0.4-2.5 μm and a 30 m spatial resolution. The reference datasets were obtained from the SJRWMD.

Methodology

A detailed flowchart of this research's methodology is given below in Figure 2. After acquiring hyperspectral imagery, radiometric and geometric corrections were performed using ENVI 5.7 software. The imagery was spatially and spectrally subset- by manually visualizing the noisy bands and removing them. A Minimum Noise Fraction (MNF) transformation was harnessed to enhance the accuracy of subsequent classification (Zhang & Xie, 2012). This process decreased noise and eliminated statistically insignificant bands, and 32 bands were selected based on eigenvalues.

Subsequently, the preprocessed imagery was imported into ArcGIS Pro for object-based classification. The initial step involved segmentation using the image classification wizard (ESRI, n.d.). Given the small size of the study area and the diverse vegetation communities, a spectral and spatial size of 15 and a segment size of 10 were chosen to optimize segmentation.

Following segmentation, training samples were selected for eight classes, considering reference datasets to ensure precision. The training datasets were then employed to classify the imagery using SVM and RF methods. The classification results were subjected to accuracy assessment based on the training samples.

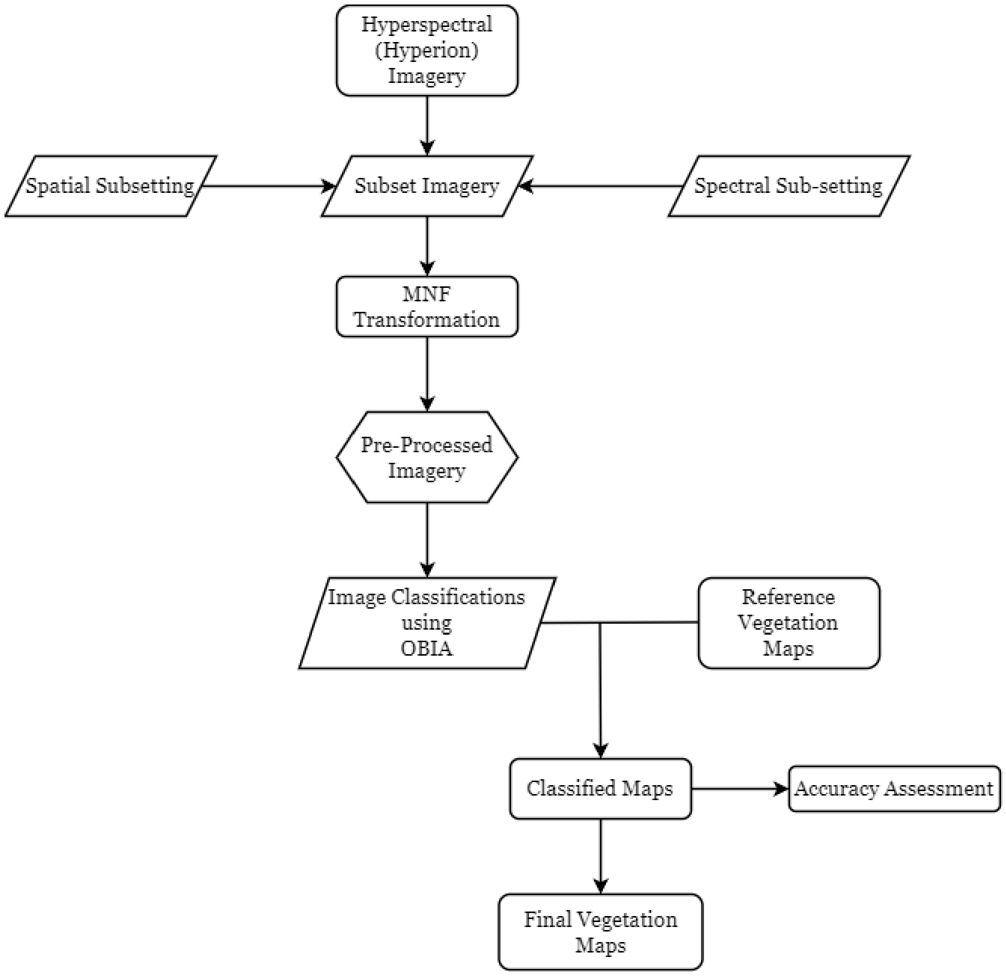

Figure 2 Flowchart for the entire procedure for wetland communities' vegetation mapping.

Results and Discussion

The classified maps using Hyperion Hyperspectral Imagery based on RF and SVM are presented in Figure 3. The vegetation classes mapped in this study are listed in Table 01. Although the reference datasets have more vegetation class, due to the imagery being from 2003, I selected eight communities based on visual inspection analyzing their spectral signatures.

Table 01: Wetland Vegetation Mapped in this project:

Plant Communities	Abbreviation	Community Type	Abbreviation
Mixed Herbaceous Marsh	HM	Herbaceous Wetland	HW
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
Cypress	CY	Forested Upland	FU
Grass/Sedge Marsh	GM	Herbaceous Wetland	HW
Hardwood Swamp	HS	Forested Wetland	FW

The classification accuracies were assessed by creating a Confusion Matrix, and the overall assessment is shown in Table 02. RF showed better overall accuracy and kappa statistics than the SVM during classification. Both classified imageries seem to have good synchronization with reference datasets upon visual inspection and verification.

Table 02: Accuracy assessment summary for the performance of two OBIA classifiers

Methods	RF	SVM
Overall accuracy %	83.6	80.5
Kappa statistics	0.81	0.77

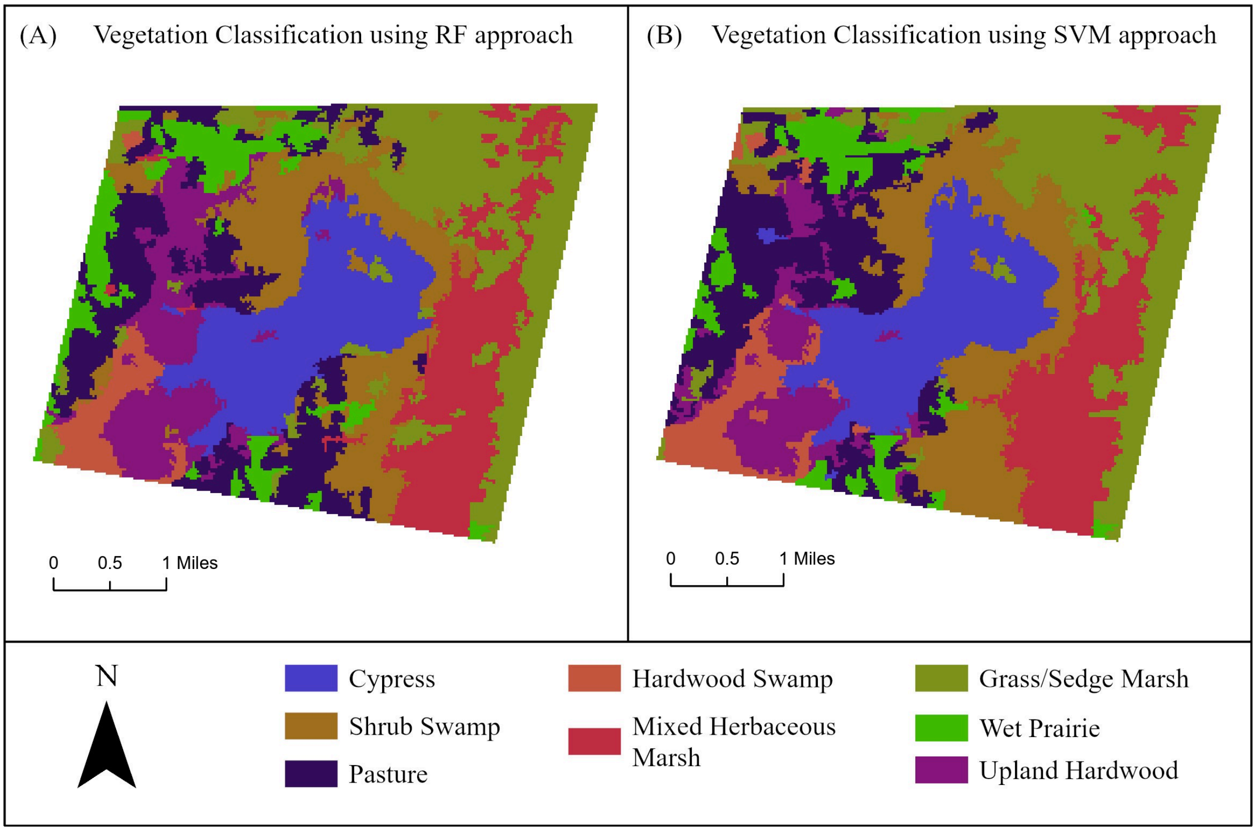

Figure 3 Maps illustrating vegetation classification of a portion of FDMCA wetland, SJRWMD, using hyperspectral remote sensing imagery with OBIA machine learning approach: a) Random Forest and b) Support Vector Machine. Legend and color combinations are common for respective vegetation classes in both maps.

While looking for the RF classification approach, the users' and producers' accuracy was above 70% based on training samples. Mainly, hardwood swamp was misclassified with upland hardwood and cypress, whereas pasture was misclassified to wet prairie and grass/sedge marsh. Mixed Herbaceous Marsh was also misidentified as grass/ sedge marsh due to their similar spatial characteristics. A detailed error matrix showing the accuracy of each class in terms of users and producers is given in **Table 03.**

Table 03: Error matrix of RF classification of the Hyperion imagery

Classes	CY	SS	PA	WP	HS	HM	UH	GM	**Row Total**	UA (%)
CY	18						1		**19**	94.7
SS	2	16							**75**	80
PA		3	14			1			**28**	77.8
WP			1	8				1	**36**	80
HS					7		2	1	**34**	70
HM						15		2	**57**	88.2
UH	1	1				1	8		**22**	80
GM	1		1	1		1		21	**32**	87.5
Column Total	**20**	**20**	**18**	**9**	**7**	**18**	**11**	**25**	**OA = 83.6%**	
PA (%)	90	80	77.8	88.9	100	83.3	72.7	84	Kappa = 0.81	

In the SVM classification approach, user and producer accuracies exceed 60%. However, producer accuracy has notably decreased, primarily due to misclassifications. Wet prairie, for instance, is consistently misclassified as various other classes, except cypress. Similarly, misidentification occurs with grass/sedge marsh, misclassified as pasture, wet prairie, and mixed herbaceous marsh. Challenges persist in correctly classifying hardwood swamp, as it tends to be misclassified as shrub swamp and upland hardwood. These misclassification patterns contribute to the observed decrease in producer accuracy, highlighting areas for improvement in the classification model.

Table 03: Error matrix of SVM classification of the Hyperion imagery

Classes	CY	SS	PA	WP	HS	HM	UH	GM	**Row Total**	UA (%)
CY	16				1		2		**19**	84.2
SS	2	16	1	1		1		1	**75**	80
PA		1	15	1				1	**28**	83.3
WP			1	9					**36**	90
HS		1		1	7		1		**34**	70
HM			1	1		13		2	**57**	76
UH	1				1		8		**22**	80
GM			2	1	1	1		19	**32**	79.2
Column Total	**17**	**18**	**20**	**14**	**10**	**15**	**11**	**23**	**OA = 80.5%**	
PA (%)	94.1	88.9	75	64.3	70	86.7	72.7	82.6	Kappa = 0.77	

Generally, the vegetation of wetlands plays a crucial role in various activities like flood control and treatment of storm and wastewater. Vegetation in the wetland is also controlled systematically such that it helps maintain the ecosystem. Using OBIA, diverse vegetation of communities can be quickly and easily mapped using modern machine learning approaches. The classified vegetation in Figure 3, derived using the methodology presented above for the FDMCA portion, shows alignment upon visual inspection.

Despite using identical training datasets, slight differences in vegetation class classification were noted. Mapping individual vegetation classes is complex due to similarities of spectral signature, leading to occasional misclassifications, such as shadows being identified as water. The analysis, focusing on community-level distinctions, adopts an OBIA approach, favoring better classification for a given area.

Acknowledging inherent limitations, the OBIA approach, as employed in this study, proves highly effective in wetland vegetation mapping, similar to studies on mapping Everglades vegetation (Zhang & Xie, 2012). Similarly, using RF outperforms K-Nearest Neighbors (KNN) in pixel and object-based analyses (Martinez Prentice et al., 2021) when applied to coastal wetland vegetation. The choice of RF underscores its efficacy in accurately classifying wetland vegetation in both this and my studies.

The effectiveness of mapping Everglades vegetation using multispectral imagery, specifically the Landsat imagery series, was demonstrated successfully (Zhang et al.,2017). The approach employed OBIA with the SVM algorithm, and results indicated an acceptable overall accuracy exceeding 87%. The overall classification accuracy was 94%, and the kappa statistics value was 0.94 on average when classifying the 15 vegetation classes (Zhang &Xie.,2012). However, based on these two studies, my study has less overall accuracy; this may have resulted from misclassification and coarse resolution of hyperspectral imagery. In addition, improper labeling and lack of ground truth data make the outcomes less accurate. Despite having access to Ecognition for segmentation, I encountered challenges in exporting classified images based on machine learning due to software issues. Consequently, I shifted the focus of the entire research to ArcGIS Pro for Object-Based Image Analysis (OBIA).

Conclusion

In the context of vegetation classification in wetlands, this study investigates the utilization of coarse-resolution imagery in contrast to the high-resolution imagery trends. It specifically explores the applicability of coarser resolution, exemplified by Hyperion EO-1 imagery, for wetland vegetation classification. Despite the inherent coarser resolution, the classification of images demonstrates a notable resemblance to reference datasets. The study underscores the efficacy of modern machine learning techniques, namely RF and SVM, for such classifications. Both classifiers show overall accuracies exceeding 80%, with kappa statistics surpassing 0.75, affirming their efficiency in vegetation classification. Notably, RF achieves an overall accuracy of 83% and a kappa statistic of 0.8, outperforming SVM, which attains 80% overall accuracy and 0.77 kappa statistics. Despite these high accuracies and resemblances to reference datasets, certain misclassifications and misidentifications across classes were seen. The study suggests that implementing post-classification assessment could enhance accuracy by addressing such discrepancies. It is acknowledged that the overall accuracy in this study falls below that reported in studies involving very high-resolution imagery. In conclusion, the research asserts that OBIA coupled with machine learning approaches proves effective for mapping wetland vegetation cover while recognizing areas for potential improvement.

References:

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L., & Lopez, A. (2020). A comprehensive survey on support vector machine classification: Applications, challenges and trends. *Neurocomputing*, *408*, 189-215.

Chhetri, M. T., & Rijal, S. (2023). Satellite-Based Analysis of Lake Okeechobee's Surface Water: Exploring Machine Learning Classification for Change Detection. *Environmental Sciences Proceedings*, *29*(1), 7.

Davidson, N. C., Van Dam, A. A., Finlayson, C. M., & McInnes, R. J. (2019). Worth of wetlands: revised global monetary values of coastal and inland wetland ecosystem services. *Marine and Freshwater Research*, *70*(8), 1189-1194.

ESRI. (n.d.). *Segmentation*. Segmentation-ArcGIS Pro | Documentation. https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/segmentation.htm

Fort Drum Marsh Conservation Area. SJRWMD. (2022, September 20). https://www.sjrwmd.com/lands/recreation/fort-drum-marsh/

Jafarzadeh, H., Mahdianpari, M., Gill, E. W., Brisco, B., & Mohammadimanesh, F. (2022). Remote Sensing and Machine Learning Tools to Support Wetland Monitoring: A Meta-Analysis of Three Decades of Research. *Remote Sensing*, *14*(23), 6104.

Mahdianpari, M., Jafarzadeh, H., Granger, J. E., Mohammadimanesh, F., Brisco, B., Salehi, B., ... & Weng, Q. (2021, July). Monitoring of 30 years wetland changes in Newfoundland, Canada. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS* (pp. 88-91). IEEE.

Martínez Prentice, R., Villoslada Peciña, M., Ward, R. D., Bergamo, T. F., Joyce, C. B., & Sepp, K. (2021). Machine learning classification and accuracy assessment from high-resolution images of coastal wetlands. *Remote Sensing*, *13*(18), 3669.

Nery, T., Sadler, R., Solis-Aulestia, M., White, B., Polyakov, M., & Chalak, M. (2016, July). Comparing supervised algorithms in Land Use and Land Cover classification of a Landsat time-series. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 5165-5168). IEEE.

Sabat-Tomala, A., Raczko, E., & Zagajewski, B. (2020). Comparison of support vector machine and random forest algorithms for invasive and expansive species classification using airborne hyperspectral data. *Remote Sensing*, *12*(3), 516.

Shippert, P. (2004). Why use hyperspectral imagery?. *Photogrammetric engineering and remote sensing*, *70*(4), 377-396.

Zhang, C., & Xie, Z. (2012). Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. Remote Sensing of Environment, 124, 310-320.

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

Zhang, C., Smith, M., Lv, J., & Fang, C. (2017). Applying time series Landsat data for vegetation change analysis in the Florida Everglades Water Conservation Area 2A during 1996–2016. *International journal of applied earth observation and geoinformation*, *57*, 214-223.