UNIVERSITY OF CALIFORNIA, MERCED

COMPARING MAPPING CAPABILITIES OF SMALL UNMANNED AIRCRAFT AND MANNED AIRCRAFT FOR MONITORING INVASIVE PLANTS IN A WETLAND ENVIRONMENT

A Thesis submitted in partial satisfaction of the
requirements for the degree Master of Science

in

Environmental Systems

by

Erik A. Bolch

Committee in charge:
Professor Erin L. Hestir, Chair
Professor Joshua H Viers
Professor Brandon Stark

Fall, 2020

The thesis of Erik A. Bolch is approved:

Professor Joshua H. Viers Date

Professor Brandon Stark Date

Professor Erin L. Hestir, Chair Date

University of California, Merced

Fall 2020

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Erin Hestir, my advisor, for encouraging me to choose and design a project aligned with my interests and giving me the opportunity to learn many new skills. I would like to thank Brandon Stark for sharing his knowledge of small unmanned aerial systems operations and regulation and advice for obtaining my FAA SUAS license. I would also like to thank my committee chair, Joshua Viers for providing unmatched knowledge on the Delta, letting me tag along with his lab and participate in water quality sampling while I was searching for a study site, as well as lending equipment for completing the fieldwork portion of my project.

I would like to thank Shruti Khanna, and the rest of the CSTARS at UC Davis for coordinating with us so we could collect concurrent data with the 2019 HyMap acquisition and sharing their research.

Several UC Merced students and staff also helped me, and I would like to thank them for their assistance throughout my project, especially Brandon Genco for lending his expertise in boating and helping me collect field data, and Daniel Gomez for helping with fieldwork and always being willing to drive, land or sea. I would like to thank Christiana Ade for remote sensing advice and coding expertise, and Anna Rallings, Zach Yinger, Julia Burmistrova, Jacob Nesslage, Brittany Lopez-Barreto, and Jaycee Martinez for helping with fieldwork operations and data collection.

Lastly, I would like to thank my parents Alan and Pat Bolch, my friends, and family for their love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

ABSTRACT

Invasive plants are non-native species that have detrimental economic, ecological, environmental or environmental effects on their surroundings and can spread rapidly. In aquatic ecosystems, they are particularly harmful because they can affect water quality and availability by disrupting flow patterns. In areas such as the Sacramento-San Joaquin River Delta in California, USA, management programs are in place which use imaging spectroscopy to map and track annual changes in invasive species patterns to inform treatment. These maps are constructed from imagery collected by an airborne imaging spectrometer. Advances in unmanned aircraft systems have enabled small imaging spectrometers such as the Headwall Photonics Nano-Hyperspec (Nano) to provide even higher spatial resolution imagery than manned flights. In this case a 5.1 cm pixel spatial resolution imagery in the VIS-NIR (400-1000 nm) was gathered by the Nano, while HyMap captured 1.7 m VIS-SWIR (400-2500 nm) imagery. For species mapping applications, the higher spatial resolution provided by the Nano allows detection of invasive weeds before they spread over larger areas. To compare the mapping capabilities and utility of unmanned aircraft and manned aircraft for mapping invasive species, aerial imagery was collected concurrently at a wetland study site using a Nano on board a unmanned aircraft (DJI M600P) and the HyMap sensor mounted on board a piloted aircraft (1975 Rockwell International 500-S). Maps were then created from similar remote sensing products using a random forest model. The Nano-Hyperspec was capable of identifying smaller patches of invasive plants, which did not appear in maps generated from the HyMap sensor. Three experiments were used to determine input criteria for the best model for map creation, which included training and test data proportions, flight direction, and acquisition. Results showed that increasing training data proportion resulted in higher median overall accuracies but the decrease in test data caused an increase in distribution width and interquartile range of model accuracy. Acquisition date also impacted model performance, but there were several confounding factors making it difficult to ascertain which, if a single variable was responsible. Flight direction relative to solar position was also significant to model performance; the Nano-generated map performed best when trained on labelled data collected in both flight directions then applied or tested only on data acquired while flying away from the solar plane. Map comparisons between those made from the Nano and those from the HyMap sensor show that the Nano performs as well as the HyMap as a source for spectral data to generate classification maps, with a higher overall accuracy for 2019, and comparable overall accuracies for other years. In addition, the higher resolution of the Nano imagery allowed detection of patches of water hyacinth present in the study site that the HyMap maps could not. However, it would not be feasible to operate the Nano as a replacement to the HyMap despite its improved detection capabilities to due to area coverage limitations. But the Nano could be used to supplement an existing invasive species management program to build an improved map of targeted areas.

1. INTRODUCTION

Invasive plants are a rapidly growing global concern due to their negative ecological and economic impacts. Their introduction often results in extinction of native plants and a reduction biodiversity, either by outcompeting or hybridizing with native species [1]–[3]. Under most scenarios invasive plants arrive without their co-evolved competitors or parasites allowing them to spread rapidly, replacing native plants without assuming their ecological roles. Plant invasions have been shown to modify ecosystem processes such as nutrient availability, nutrient cycling, soil chemistry, water tables, hydrology, food waste, and habitats [4]. Management of these potentially detrimental impacts are complicated by changes in climate [5] and intensified by increases in invasion frequency due to globalization [6], [7]. Human-mediated introductions of invasive plants are most common and tend to be more rapid, increasing propagule pressure and exacerbating the threat of the economic and environmental damages associated with invasive plants [8]–[10].

Over a decade and a half ago, global costs associated with management of invasive plants was estimated at $34.6 billion as a result of approximately 25,000 foreign plants, and costs are bound to be higher now as globalization has increased [2]. To mitigate the economic and environmental damage, governmental cooperation and management on national and international levels is required. The United Nations Sustainable Development Goal 15, *Life on Land*, includes a focus on invasion prevention, control and eradication [11]. Other governmental bodies such as the European Environmental Agency (EEA) and United States National Invasive Species Council (NISC) are also working to improve global management capabilities. The EEA has developed indicators summarizing invasion trends and biodiversity threats and the NISC has been attempting to standardize data formats and protocol [1]. Understanding invasion origins, pathways and processes is a key step in management and requires geographic observations.

Imaging spectroscopy or hyperspectral remote sensing has become a favored tool for invasive plant species mapping due to its synoptic viewpoint and its proven success in providing sufficient spectral data to differentiate between species within complex communities such as wetlands [12]–[17]. The capability to detect plant traits and species using this technology coupled with map creation is invaluable for monitoring and management of invasive plants [18]. For this reason, airborne imaging spectroscopy campaigns have become a common practice because they offer the large volume of spectral information necessary to discriminate between species sharing similar attributes [1], while having moderate to high spatial resolution need to detect species patches.

Improvements in technology of unmanned aircraft systems and sensors have allowed for an increase in their capabilities, offering the potential to fill critical gaps between field spectroscopy and manned flights. Field spectroscopy typically utilizes a researcher with a backpack spectrometer meaning it has a very small geographic coverage due to the difficulty and time associated with navigating field sites while manned flights have a much larger geographic coverage, but loses the extremely high spatial detail associated with field spectroscopy. Unmanned aircraft systems sit between this range of capability, providing

imagery with a geographic footprint smaller than manned flights but larger than the single points provided by field spectrometry. Unmanned aircraft systems also have high operational flexibility and low costs relative to manned flights as well as the extensive field campaigns required for field spectrometry [19], [20]. This gives unmanned aircraft the potential to conduct more frequent, small-footprint acquisitions. The capability to launch on-demand enables better characterization of phenological stages, differences in phenology between species, and allows sampling during specific events such as natural disasters. Multispectral cameras or simply high resolution RGB cameras are the most commonly used sensors with unmanned aircraft systems; however, newer pushbroom imaging spectrometers, such as the Headwall Nano-Hyperspec (Headwall Photonics, Bolton, MA, USA) offer tremendous potential for mapping species. Though there has historically been difficulty in achieving consistent spectral and radiometric quality with from unmanned aircraft mountable line scanners [21], studies have shown success in providing detailed community mapping in notoriously difficult regions such as wetlands [22] and grasslands [23].

The objective of this study is to compare mapping performance of imaging spectrometers mounted on unmanned aircraft and those mounted on manned aircraft spectrometers used routinely for species level monitoring. The unmanned aircraft mounted sensors offer much higher spatial resolution than manned aircraft mounted imaging spectrometers but have lower spectral quality and a smaller spectral range and spatial footprint. Manned airborne campaigns typically have high spectral and spatial coverage, moderate geographic coverage and low revisit time due to the complexity of flight planning and costs. To compare the wetland classification mapping capabilities of these two platforms equipped with imaging spectrometers, classification maps were creating using the Nano-sensor and compared with maps generated from the HyMap sensor using similar methods. To build the best model for classifying the data from the Nano, 3 experimental treatments were performed to determine the effects of: i) training and test data selection and quantities, ii) sun-sensor geometry or flight direction, and iii) phenology or acquisition date on the classification performance of the models developed. The unmanned aircraft mounted Headwall Nano-Hyperspec imaging spectrometer was flown concurrently with the manned aircraft-mounted HyMap imaging spectrometer that was tasked with the purpose of collecting annual invasives-mapping imagery. The data collected by the Nano were then used in multiple models to evaluate the impact of the 3 experimental treatments leading to selection of the best performing model to create a thematic map, which was compared with maps created from manned aircraft data as part of the aquatic weed-mapping program in the Sacramento-San Joaquin River Delta (the Delta) by the University of California Davis.

2. METHODS

2.1 Overview

The unmanned aircraft mounted Nano sensor was used to collect imaging spectroscopy data near Liberty Island, CA, a flooded island in the Delta concurrently with an annual manned flight by the HyMap sensor, flown as part of California's aquatic weed mapping program in the Delta. Two sub-regions of the HyMap imagery collection, each approximately 200 x 200 m were selected for comparison based upon accessibility and presence of invasive species. Data collected by the HyMap sensor was delivered to UC Davis and processed into thematic maps using a random forest model [24]. Taking advantage of on-demand flight capability of unmanned aircraft, two subsequent unmanned flights were conducted on May 3 and June 13, 2019 with the Nano, capturing the same 200 x 200m areas. The imagery collected by the Nano were pre-processed and thematic maps following the target classes (Table 1) were created using a random forest modeling approach. Three experimental treatments were performed with the unmanned aircraft data by limiting the training and test data provided to the models based upon the experimental treatment to determine the impacts of each treatment and select the best model classification performance. These are outlined in Figure 1. After a high performing model was selected, the map created from the HyMap sensor by the UC Davis Center for Spatial Technology and Remote Sensing (CSTARS) and the maps created from the Nano sensor were compared to evaluate their relative species detection capabilities.

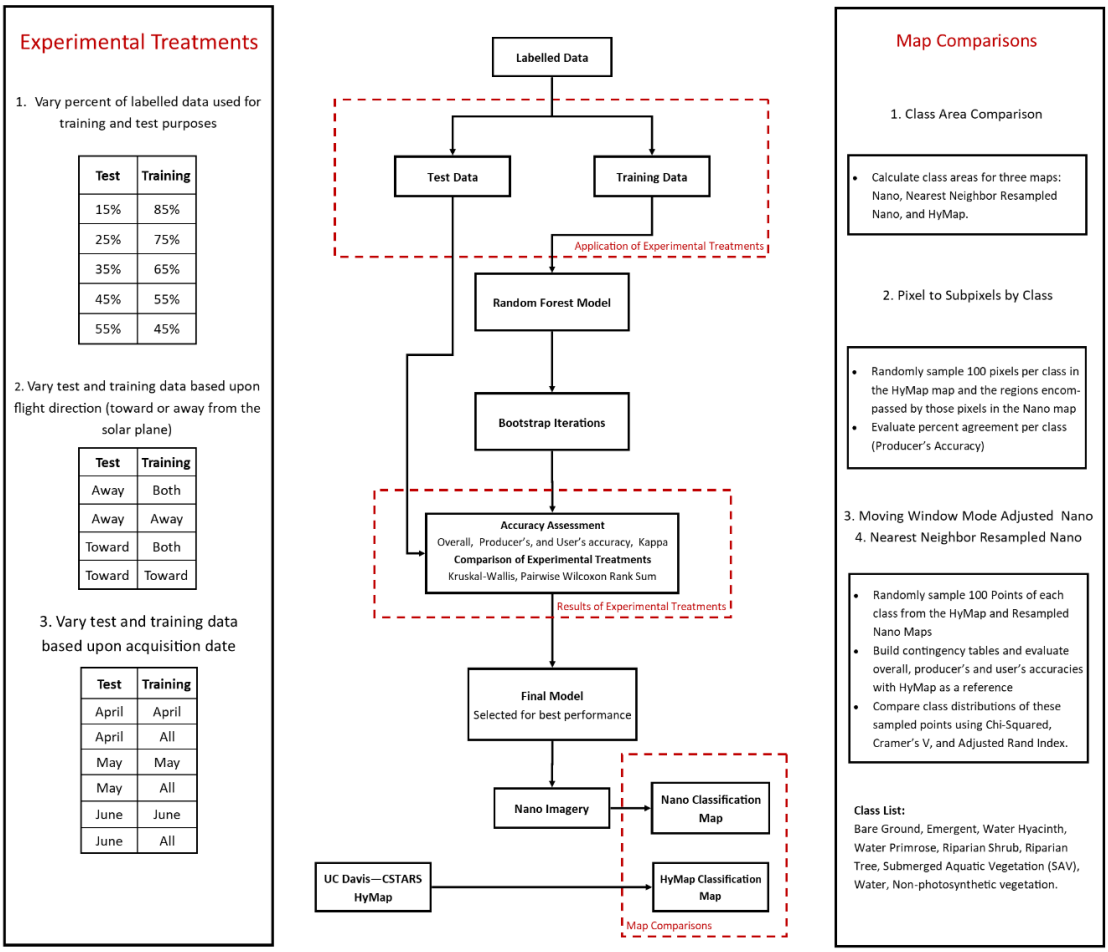


Figure 1. Project flow diagram showing data processing chain, experimental treatments and map comparisons conducted.

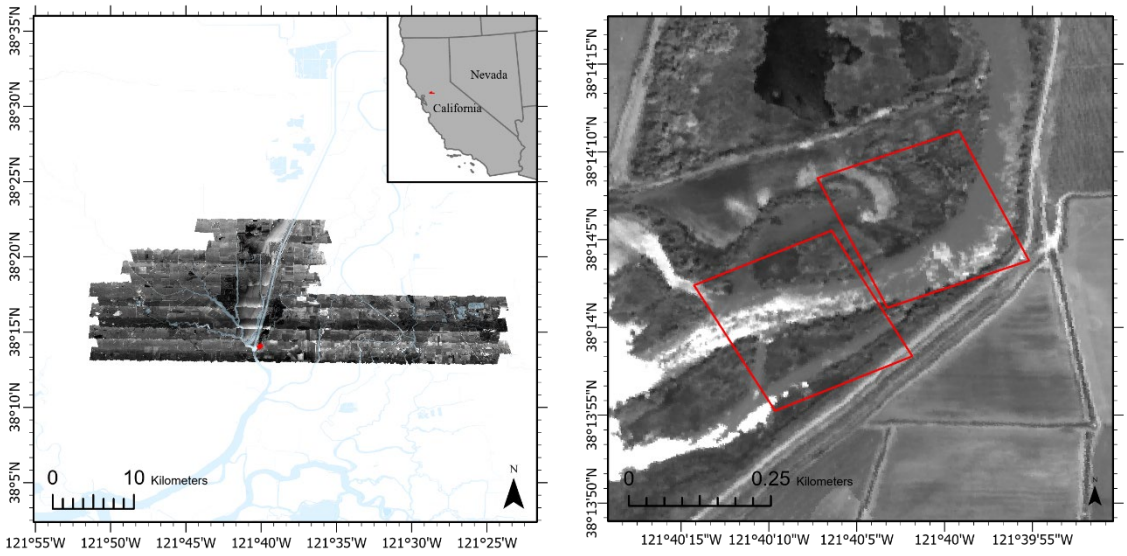
2.2 Study Site

The Sacramento-San Joaquin River Delta, the upstream component of the San Francisco Estuary, is the largest tidal freshwater estuary in the western United States. It is a heavily engineered system consisting of complex waterways, reclaimed islands supporting agriculture, and flooded islands. The Delta provides a habitat for numerous species and has a disproportionately important role in biodiversity and providing vital ecosystem services [25], [26]. It also serves as a major hub for water in the state, providing drinking water to the bay area as well as southern California and the main water source for agriculture in California's Central Valley. Agricultural production in the Central Valley accounts for roughly 80% of the state's $50 billion agricultural industry [27]. For these reasons, maintaining water quality and availability in the Delta region is a priority for California. The state has legislated that Delta water management must meet the co-equal goals of providing reliable water supply while protecting Delta ecosystems. The Delta is one of the most invaded ecosystems in the world. Invasive plants in the region have been

detrimental to water pumping, water quality, commerce, recreation, and have impacted native species [24].

Due to the economic and ecological importance of the Delta, agencies need to know where to spray and remove invasive plants to minimize impacts. Efforts to map and monitor invasive plants using airborne imaging spectroscopy have been under development dating back to 2003, led by CSTARS at UC Davis [12], [15], [16], [28]. A combination of HyMap and AVIRIS imaging spectroscopy data has been collected nearly annually from 2003 – 2008 and from 2014 – to present, and aquatic invasive species mapping is considered near-operational. Currently, images are classified into 5-10 classes that include different species of aquatic macrophytes and other co-occurring vegetation categories using a random forest machine learning algorithm and GPS collected field data through extensive field campaigns. Overall map accuracies range from 85-95% and Kappa coefficients range from 0.83-0.94. These maps have been used to quantify the effectiveness of management interventions including herbicide applications and salinity intrusion barriers [18], [29], investigate drivers and implications of biological invasions [30], [31], and their impact on the water quality and physical habitat (Hestir, Schoellhamer, Greenberg, Morgan-King, & Ustin, 2016).

Although airborne imaging spectroscopy data provide the highest spectral quality and often spatial quality of data for species-level mapping, it is limited to infrequent (e.g., annual, bi-annual) snapshot assessments of species distributions. Furthermore, it has relatively high costs, making it difficult to fund on an annual or seasonal basis. Consequently, there has been recent interest in using unmanned aircraft acquired imagery to increase spatial and temporal quality and reduce costs. Unmanned aircraft operations require much less coordination and planning due to smaller scale, lower flight altitudes, and less overhead. The ease of deployment makes it possible to collect data routinely and on-demand without high costs.



(a) (b)

Figure 2. Maps of study region. (a) Flightlines conducted by HyMap in 2019 with Nano collection area highlighted in red and (b) enlarged Nano acquisition area.

2.3 Target Classes

Two major invasive species of concern in the Delta are water hyacinth (*Eichohornia crassipes)* and water primrose (*Ludwigia* spp.). Water hyacinth is a perennial, mat-forming floating aquatic plant that has been a problematic species in the Delta for decades [33]. Mats of water hyacinth can double biomass within 10 days [34], quickly overtaking areas, decreasing water quality and quantity by decreasing dissolved oxygen content and increasing transpiration [28], [35], and obstructing waterways for commerce and recreation [36]. Water primrose is a problematic amphibious plant which can form floating mats. Its amphibious capability and fast growth rates have made it a threat to the Delta, endangering native species in the region, as well as humans because primrose mats provide habitats for mosquitos transmitting the West Nile virus [4]. As the primary focus of this investigation they are the only species-level classes included. Other land cover classes were chosen to align with existing classes in HyMap-derived maps created by CSTARS the previous year [37]. A complete list of classes for each platform and their descriptions can be found in Table 1.

Table 1. Classes of Interests and Descriptions

Map Class	Description
Unclassified	Unclassified land cover or area outside of analysis
Bare Ground	Asphalt, gravel, levee riprap, and bare soil
Emergent Vegetation (EMR)	Cat tail (*Typha* spp.), common reed (*Phragmites australis*), giant reed (*Arundo donax*), and tule (*Schoenoplectus* spp.)
Water Hyacinth	Water Hyacinth (*Eichhornia crassipes*)
Water Primrose	Water Primrose (*Ludwigia* spp.)
Riparian Tree	Trees in the area including willow species (*Salix* spp.)
Riparian Shrub	Shrubs in the area
Submerged Aquatic Vegetation (SAV)	Algae mats and submerged species: Brazillian waterweed (Egeria densa), Coontail (Ceratophyllum demersum), Curly leaf pondweed (Pomatogedon crispus), Fanwort (Cabomba caroliniana), Sago pondweed (*Stuckenia pectinata*), Watermilfoil (*Myriophyllum spicatum*), and Waterweed (*Elodea canadensis*)
Water	Water
Other Vegetation	Species or cover not obeserved in unmanned aircraft study region including pennywort (*Hydrocotyle* spp.) and mosquito fern (*Azolla* spp.)
Non-Photosynthetic Vegetation (NPV)	Senescent or dead vegetation

2.4 Imaging Spectroscopy Data

The Headwall Nano-Hyperspec is a small, lightweight, pushbroom imaging spectrometer that can be mounted on a unmanned aircraft. It can be purchased bought as a turnkey package including a DJI M600Pro, DJI Ronin gimbal, calibration tarps, and software (Headwall Photonics, Bolton, MA). The Nano records radiance data across 270 bands of visible and near infrared light across the 400 to 1000 nm range of the electromagnetic spectrum (Table 2).

Table 2. Sensor Specifications

	HyMap	Nano-Hyperspec
Type	whiskbroom	pushbroom
Spectral Range	450 - 2450 nm	400 - 1000 nm
Number of bands	128	270
Spectral Resolution	15-18 nm	2.2 nm
Signal to Noise	>500:1	> 15:1 (1000 nm) < 140:1 (550 nm)
Spatial Resolution	1.7 m	0.051 m
Swath Width (FOV°)	61.3	15.85
Operational Altitude	> 458 m *	< 122 m **
Platform	1975 Rockwell International 500-S	DJI-M600P

*Typical lowest safe altitude according to FAA.

**Maximum altitude for unmanned aircraft without FAA approval.

The flight plan was constructed to account for potential effects of solar geometry. Sun-sensor directional effects can result in specular reflectance of water, large amounts of shadow, and variability of across track reflectance [38]. The data from the Nano sensor mounted on the unmanned aircraft were collected to deliberately enable the evaluation of these effects on classifier to performance in experiment 2. The flight plan was created and operations were conducted using Universal Ground Control Software (UgCS, SPH Engineering 2018). The layout consisted of two boxes of 16 flight lines each, to cover two 200 x 200 m areas at 115 m. The flight lines were oriented so the unmanned aircraft would be flying toward (SE direction) or away (NW direction) from the solar plane and each box was designed with a very high, 60% overlap for flight lines to allow a full mosaic of the region to be created using only lines from the NW or SE direction and still have a 10% overlap.

By flying at an altitude of 115 m, the Nano captured imagery with a 5.1 to 5.4 cm spatial resolution, varying slightly due to ground topography and wind conditions. A total of 32 flight lines in the specified study area were collected on each acquisition date, with the May 3 and June 13 acquisitions succeeding in capturing the green-up of several plant species including the target species water hyacinth. Flight conditions for the days of the study can be found in Table 3.

Table 3. Flight conditions for unmanned aircraft flights.

Date	Start Time (PDT)	Stop Time (PDT)	Start Solar Azimuth (°)	Stop Solar Azimuth (°)	Start Tide Height (ft)	Stop Tide Height (ft)	Wind Speed (mph)	Solar Radiation (w/sq.m)
2019-04-09	12:28	13:14	137.04	154.69	1	0.5	7.1	907
2019-05-03	13:06	14:12	146.67	185.31	0.7	1.6	5.1	838
2019-06-13	12:50	13:39	126.98	156.46	2.2	2.7	14.2	974

The raw image cubes collected by the Nano were converted to radiance using a dark calibration of the sensor conducted preflight, then the imagery was orthorectified using an iterative procedure, adjusting offsets of pitch, roll, yaw, and altitude for the onboard GPS and IMU. Each iteration was compared to Google Maps imagery, as well as the previous iteration and known ground location points using Headwall's SpectralView software [40] and ENVI 5.5 [41]. Several iterations adjusting the previously mentioned parameters were performed to achieve a best possible orthocorrection. After this step, the cubes were converted to reflectance in the Headwall SpectralView software using a radiance sample from the 56% reflectance section of the calibration tarp present in the imagery. Sampled healthy target vegetation had close to this reflectance in the NIR region and using similar reflectance to the intended target was recommended (G. Chenevert, personal communication, June 20, 2018). Savitzky-Golay smoothing was considered to reduce the noise present but was omitted from the processing chain because there was still a considerable amount of noise present afterwards, especially in the NIR region.

2.5 Ground Reference Data

Plant species and location data of the target classes (Table 1) were collected during the week following the concurrent flight on April 16-18 for use in training and testing a thematic classification of the Nano data. Using a 2017 HyMap classification map obtained from UC Davis, a stratified random sampling scheme was used to create 200 points, consisting of 10 points per class per flight box, to visit in the field and record [42]. This sampling design was chosen to maintain random sampling as best as possible while still collecting a sufficient number of samples from each class. Of the 200 ground reference points created, the 84 accessible ones were visited, and data were collected using field guides for species identification, a Trimble Geo7X RTK kit with a Zephyr-3 antenna for GNSS location, and a Zodiac Mk. II inflatable boat to the points (Figure 3A).

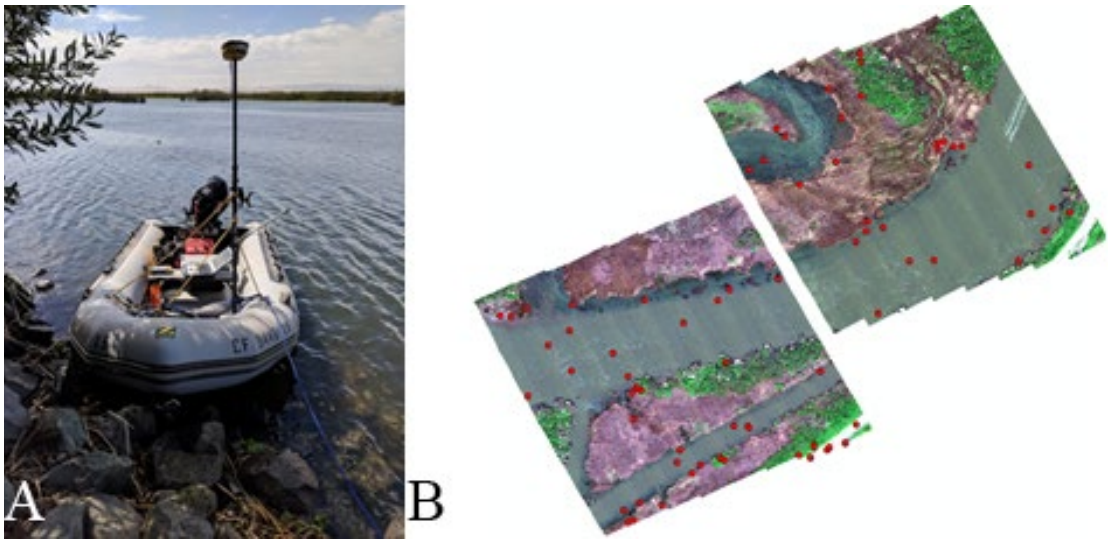


Figure 3. Ground reference point collection equipment and map. a) Zodiac and GPS setup for data collection. b) Ground reference points collected.

At each of the 84 points visited, the target classes present were documented, photos were taken, and additional information regarding patch sizes, and the surrounding area was recorded. After subsequent flights in May and June, 10 randomly selected points from the 84 visited were selected to re-visit and coverage at those locations was checked for consistency with previous month's data. Since the spatial pixel resolution of the Nano-Hyperspec was between 5.1 to 5.4 cm, great effort was taken to maximize positional accuracy of this survey and corrected GNSS coordinates fell within a 10 cm range after conducting a differential correction using GPS Pathfinder Office [43]. Points not accessible via watercraft were labelled using photographic and spectral interpretation of the unmanned aircraft collected imagery aided by Google Earth Imagery and USDA National Agriculture Imagery program (NAIP) aerial imagery. Water hyacinth had very few ground reference points and many patches were senescent during the April survey. Imbalance or low quantities of reference data for classes has been shown to hinder performance of classifiers [44], thus additional datapoints points for that class were added for the water hyacinth class by creating vegetation patch polygons in the unmanned aircraft imagery, then randomly sampling from those polygons to provide additional 72 reference points in total.

2.6 Classification

Random forests (RF) models [45] are a popular choice for remote sensing for classification of species from various types of imagery [1]. RF was chosen as the classification method because it was used by UC Davis on the HyMap data due to high accuracy and proven success for the chosen ecosystem and region [24]. RF works by building a set of classification trees which vote on the most probable class that a datapoint represents. Each tree randomly selects a subset of training data to be used for its construction, reducing the chances of overfitting the model. RF have specifically been

successful at mapping species in complex environments such as wetlands [44] and have performed better than other non-parametric classification methods [24]. Though RF often have improved classification accuracies over other methods, they lack a direct quantification error [46]. To account for the lack of direct error quantification from a random forest, a bootstrapping procedure of building multiple random forests was used to capture the range of accuracies of the RF model based upon the random selection of training and independent test data. This is important to capture this uncertainty because poor model inputs can cause considerable errors in classification [47] and random forests are sensitive to spatial autocorrelation [48]. The bootstrapping method should reveal some of these errors, giving a better idea of the model's performance. After bootstrapping the RF models, accuracies metrics for each model were examined. These consisted of the overall accuracy, producer's accuracy, user's accuracy, and Cohen's Kappa statistic [49]. The overall accuracy specifies the probability that the thematic class identified in the imagery is representative of ground reference data at that location. Producer's accuracy (error of omission) indicates the probability that the land cover class identified at the ground reference has been correctly classified. User's accuracy (error of commission) indicates that probability that a classified point in the map would have the same land cover class if visited in the field. Cohen's Kappa indicates the overall level of agreement between the ground reference data and the classified data taking into account random agreement between reference data and classification results [50], [51]. Kappa was calculated for comparative purposes, but otherwise omitted due to its similar functionality to overall accuracy [52], [53].

The RF models used to classify the Nano data were constructed and evaluated using the caret [54] and randomForest [55] packages in R [56]. A flowchart depicting an outline of the data processing procedure can be found in Figure 1. When constructing the models, selection of training and test data were restricted so that no geographic duplicates were chosen if using labelled data acquired in both flight directions. This was done to prevent inflation of accuracy statistics. To compare experimental results and select the best model for map-making, Kruskal-Wallis and pairwise Wilcoxon rank-sum tests were performed on accuracy metrics to determine if there was a significant difference in distributions between the experimental results with a p-value of 0.05.

2.6.1 Input Variables

The RF model inputs included reflectance data, occurrence texture data, a forward minimum noise fraction (MNF) transformation, 80 spectral indices, and image segments. During the preprocessing stages there was a lot of noise in the reflectance spectra collected. Texture occurrence data contains a kernel derived mean, variance, entropy, and skewness for each pixel and band within an image, effectively measuring how similar a pixel is to the other pixels within the kernel. The MNF transformation is a linear transformation process that consists of using a noise whitening, then uses a principle components analysis (PCA) transformation on the noise-whitened data [57]. This results in an output of bands with an axis of signal to noise ratio, creating a set of images ordered according to noisiness [58]. Occurrence textures metrics were calculated using a 3 x 3 kernel in ENVI 5.5 [41]. The forward MNF rotation was also conducted on all flight lines using ENVI 5.5. The 80

spectral indices included are meaningful indicators associated with plant biophysical properties (Appendix A). They were calculated using R [56], and the formulae from the hsdar package [59]. Image segmentation is the partitioning of an image into sets of pixels or objects based on the similarity of pixels values to one another. At high resolution, object-based image analysis has shown great potential for image classification [60]–[62] and has performed better than pixel-based methods alone [62]. Segments were calculated using a Large-Scale Mean-Shift of the green, red, NIR bands in the segOptim package [63] in R utilizing Orfeo Tool Box [64]. Segmentation parameters are difficult to optimize, often being determined by visual inspection after the parameters relating to size and complexity of targets are chosen [60], [65], a procedure which was followed in this study. Model input data included 987 labelled points with 1702 variables each, and in total there were 96 flight-line images for the Nano, adding up to roughly 8.6 TB of data.

2.6.2 Random Forest Model Construction

All RF models and accuracy statistics were constructed and calculated using R[56]. During construction of the RF models, the number of trees necessary was examined as well as the three experimental treatments hypothesized to affect performance of the models. Each of these factors was explored using a bootstrapping approach consisting of 1000 bootstraps. The 1000 bootstrap iteration approach was chosen to provide a thorough look at the variability of model performance, showing a measure of uncertainty and accounting for the lack of a direct error quantification. The distribution of accuracy statistics across the different factors were analyzed to determine if their role in the model was significant. Due to the extremely high data volume and available processing capabilities, variable importance was used to winnow the number of input variables and final classification models so they could be applied efficiently to classify large imagery.

2.6.3 Data Reduction and Variable Importance

RF are fairly robust to redundant data, there is evidence that reducing data dimensionality does not decrease RF performance [66]. Others suggest that data reduction is necessary when dealing with high dimensional datasets and resulted in improved performance when only important variables are used [67]. Combined with small sample sizes of training data, not reducing redundant high dimensional data may result in misclassification by exacerbating imbalances in class composition of the training data [68]. To reduce the amount of inputs data, variable importance was used to determine the 500 most important variables, reducing the inputs down to 500 from 1702. Variables were selected by calculating an equally weighted combination of their mean decrease in accuracy and mean decrease in GINI impurity. GINI impurity ($\mathbf{G}$, Equation 1) is a measure of how frequently a random selection from the dataset would be labelled incorrectly if it was randomly labelled based upon the distribution of labels in the dataset, where $\mathbf{p}$ is the probability of a selection $\mathbf{i}$ being chosen and $\mathbf{n_c}$ is the number of classes. The mean decrease in GINI score ($\mathbf{MDG}$, Equation 2) can be thought of as the average gain in purity or reduction in groups of data by a splitting a decision tree at that variable [69], where $\mathbf{N_T}$ is the number of trees in the forest, and $\mathbf{T}$ represents all of the nodes where the selected class

is present. The mean decrease in accuracy (**MDA**, Equation 3) can be thought of as the change in out-of-bag error of observations that are incorrectly classified due to excluding a variable from the model in question [70], where $\mathbf{N_T}$ is the total number of trees, $\mathbf{t}$ is the target node, $\boldsymbol{E(t)_{oobr}}$ is the out of bag error with the variable removed and $\boldsymbol{E(t)_{oob}}$ is the out-of-bag error without the variable removed.

$$\boldsymbol{G = \sum_{i=1}^{n_c} p_i\,(1 - p_i)} \tag{1}$$

$$\boldsymbol{MDG = \frac{1}{N_T} \sum_{T} G} \tag{2}$$

$$\boldsymbol{MDA = \frac{1}{N_T} \sum_{t=1}^{N_T} E(t)_{oobr} - E(t)_{oob}} \tag{3}$$

These are both important factors to consider when evaluating variable importance and were therefore were weighted evenly in frequency calculations. After the 500 most important variables were identified for each iteration, they were ranked by the 500 appearing most frequently across all experimental bootstrap iterations. These 500 considered important most frequently were selected to be used in an updated model.

2.6.4 N Trees

Because the number of trees influences the total number of decisions in the model, more trees will typically increase performance up to a threshold, but may result in too rich of a model with unnecessary variance [71]. To assess the optimal number of trees to be used in the model, random forest models with 100, 500, 1000, and 5000 trees were constructed. The performance was assessed by evaluating the mean and variance of the overall, user's, and producer's accuracy of the classes, maintaining a constant of 100 bootstrap iterations with each number of trees, all variables, and a simple 50/50% training and test data split. A Kruskal-Wallis ranked sum test was used to evaluate difference in overall accuracy, and a pairwise Wilcoxon rank sum test was used to assess the differences in user's and producer's accuracy between the differing number of trees.

2.6.5 Experiment 1: Sensitivity to Training and Test Data

With no direct quantification of error when using an RF model, capturing variability in the performance can serve as a proxy for the range of accuracies to be expected from use of a model with the same conditions and different training and test data. The quality of RF models is directly dependent on the quality of training and test data provided to them, and when evaluating a RF model, it's important to maintain separate, independent training and

test data to fairly evaluate the model performance. Using the same dataset to train and test will result in inflated accuracy metrics. With a very low number of labelled data to use in training and testing, the performance of the classification in this study was expected to be even more dependent on the quality of these data. To quantify the variability in performance stemming from the reliability of the labels of the train and test datapoints, a bootstrapping approach was used. Different treatments were applied to the bootstrapping procedures by differing the ratios of training and test data. This enabled an evaluation of the individual training and test points influence the accuracy metrics. By decreasing the amount of training data, model performance was expected to decrease because the model is less prepared to handle the variability of each class. By decreasing the quantity of test data, the range of accuracy was expected to increase if some labelled data points are of poor quality. To evaluate RF sensitivity to labelled data, the independent test dataset was varied from 15 to 55 % of the total data available and the training dataset was varied from 85 to 45 % of the total data available, and 1000 iterations of RF models were run for every test/train data split. One-thousand iterations were chosen to err on the side of caution, because it did not show significant variability from the other quantities of iterations. The variation in results across different splits were assessed using overall, producers', and users' accuracies and Kruskal-Wallis and pairwise Wilcoxon rank sum tests of these metrics.

2.6.6 Experiment 2: Flight Direction Effects

To test the impacts of sun-sensor geometry caused by flight direction, there were four model sets built varying only the flight direction of training and test flights. One set of models was built by training and testing on only on data collected flying away from the solar plane (Away-Away), one was built doing the same toward the solar plane (Toward-Toward), and two were built by training the classifier using data collected from both flight directions, then independently tested only on away from solar plane or toward solar plane collected data (Both-Away and Both-Toward). For each model 1000 RF bootstrap iterations were evaluated, using 1000 trees and a 65/35% ratio of training to test data based on the results of Experiment 1. The resulting accuracies were compared using a Kruskal-Wallis test and pairwise Wilcoxon rank sum tests to determine significance.

2.6.7 Experiment 3: Flight Date Effects

It was hypothesized that image acquisition date would have an impact on the performance of the classification because the quantity of non-photosynthetic vegetation greatly decreased across the spring season, changing their spectral signatures relative to the initial acquisition date. For this reason, it was necessary to determine whether the individual snapshots of each date could be better classified using data from that date alone, or if a broader, more representative dataset using information from all three unmanned aircraft acquisitions would increase performance of the model. It was anticipated that multiple acquisition dates may improve classification accuracies as a result of the spectral variation of water hyacinth due to phenology [12]. To determine if a single-date-based classification performed better or worse than a combined data classification, sets of RF

models were constructed by restricting the ground reference data used in training and testing by date and then comparing those results to models trained using data from all acquisition dates. For example, models trained and tested based on only the April data (April-April) were compared to models trained on all three acquisition dates and tested only on April data (April-All). This was also done for May and June acquisitions. Once again, accuracy metrics of overall accuracy, producer's accuracy, user's accuracy and Kappa were compared using Kruskal-Wallis and pairwise Wilcoxon rank sum tests.

2.6.8 Selecting the Map-Making Model

The results of each experimental treatment were used to inform the final model selected for map-making. The experimental results supported using a model built with the 500 most important variables, with the combination of 1000 trees, and a 65/35% training/test data split, April only acquisition date, and using labelled data from both flight directions for training but only testing on labelled data collected during flights in the away from solar plane direction. After the selection was narrowed to this set of 1000 model iterations, a model with high overall accuracy as well as high producer's and user's accuracies for the species of concern, water hyacinth and water primrose, was selected to support the goal of successfully mapping those two species effectively.

2.6.9 Map Creation and Comparison

After the final model was selected, one target class, non-photosynthetic vegetation, had to be added to match with the HyMap map provided by UC Davis. Initially, NPV was omitted as a class because it was thought that the Nano multi-date acquisitions would be able to better identify species that initially presented as senescent. To add the NPV class, all plant classes that were senescent were reclassified to NPV using an NDVI threshold. NDVI is a metric used to evaluate vegetation health or greenness [72], low values correspond to targets such as water or bare soil, while higher values correspond to green vegetation. Based upon values of ground reference data, the NDVI range of 0 to 0.3 was chosen to represent non-photosynthetic vegetation. The class of 'other' was used to describe duckweed or pennywort classes present in the HyMap analysis but missing from the Nano model. For that reason, 'other' was omitted from any location specific comparative analysis. After adding these classes, the final model was applied to the Nano imagery to generate a map. To improve spatial coherency of final thematic maps, the clump classes function in ENVI (Exelis Visual Information Solutions, Boulder, Colorado) with a 3x3 pixel window was used to smooth the images and eliminate isolated pixels.

The four map comparisons outlined in Figure 1 were performed between the thematic maps derived from the Nano sensor, and the HyMap map provided by CSTARS. One comparison simply compared areal coverage, while the three remaining comparison methods incorporated a positional component to evaluate whether the same spatial areas were shared the same class label. Despite rigorous orthorectification of both the Nano and the HyMap imagery, there remained a slight georegistration error that could not be evaluated precisely due to lack of ground control points, which are extremely difficult to collect in wetland environments. Positional agreement between the maps was estimated to

be within 3.0 m based upon positions of static objects within the study region visible on both maps. This quantity corresponds with the positional accuracy of the GPS unit and IMU onboard the unmanned aircraft, meaning any point from the Nano imagery should fall within a 1.5m radius in the HyMap imagery.

The first comparison was a per-class area comparison across three maps, the Nano map, a Nano map resampled to the HyMap spatial resolution using nearest neighbor, and the HyMap map. Per-class areas for each map were calculated as the total number of pixels of each multiplied by the pixel size of the map. This comparison neglects the positional component of the maps but shows the quantity of each class detected. The second comparison was a location specific approach, which randomly sampled 100 points from each class in the HyMap imagery, then created a square buffer region equivalent in size to the HyMap pixel (1.7m) around the centroids of those pixels in the Nano imagery and extracted the values. This created a set of subpixels from the Nano image which were then compared to the sampled pixel values of the HyMap image by class. The third comparison uses a moving-window approach that reclassifies pixels in the Nano map based upon the mode of the classes occurring within a 3 x 3 m window (59 x 59 pixels), matching with the spatial uncertainty of the unmanned aircraft GPS sensor. The moving window was used to reassign each pixel in the imagery to the mode of the classes within. Then, 100 random samples of each class were selected from the reclassified Nano map, and the class values were extracted and compared with extracted values from the same locations in the HyMap map using a confusion matrix with the HyMap map as reference. The fourth comparison method resampled the Nano map to the same spatial resolution of the HyMap map using a nearest neighbor approach. After resampling, class areal comparisons were made between the two maps and 100 randomly sampled locations for each class were extracted from both maps and again evaluated for agreement using a confusion matrix with the HyMap map as reference.

For comparisons 2 through 4, the distributions of the 100 samples per class were compared using a Pearson's chi-squared test [73] to evaluate if the distribution of the points per class were the same between maps. Cramér's V statistic (Cramér 1946) and Adjusted Rand Index [75] were used to determine the degree of correlation between the two distributions. For comparison 1, the total areal distributions per class were compared using the same methods. In a chi-squared test, the value of χ^2 (Equation 4) is calculated, where O_i is the number of observations of type i, and E_i is the expected number of observations of type i, the χ^2 value can then be compared to that of a χ^2 distribution based upon the degrees of freedom. Cramér's V (V, Equation 5) uses the chi-squared statistic to measure the association of the two nominal variables, resulting in a value in the range of 0 to 1 corresponding to the degree of agreement between the two distributions. In equation 5, n represents the total number of observations, k represents the number of columns and r represents the number of rows. The Adjusted Rand Index (ARI, Equation 6), is a normalized count of the correctly classified pairs of elements, where n refers to the a value from the contingency table, i refers to the row number, j refers to the column number, a refers to the row sum, and b refers to the column sum.

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \tag{4}$$

$$V = \sqrt{\frac{\chi^2/n}{\min(k-1, r-1)}} \tag{5}$$

$$ARI = \frac{\sum_{ij}\binom{n_{ij}}{2} - \left[\sum_i \binom{a_i}{2} \sum_i \binom{b_j}{2}\right] / \binom{n}{2}}{\frac{1}{2}\left[\sum_i \binom{a_i}{2} \sum_i \binom{b_j}{2}\right] - \left[\sum_i \binom{a_i}{2} \sum_i \binom{b_j}{2}\right] / \binom{n}{2}} \tag{6}$$

3. RESULTS AND DISCUSSION

3.1 Map-Making Model

After conducting the three experimental treatments and reducing the quantity of input variables to 500, a final model set of 1000 iterations using the criteria of April collected, 65/35% training/test data split consisting of training data from both flight directions and independent test data from only the away-from-solar-plane direction was built. Overall accuracy within the 25-75th quantiles ranges from 80 to 86% and Kappa coefficient ranges from 76 to 83%. With a median overall accuracy of 82% and a median Kappa statistic of 0.79, the models performed well and are similar in performance to that of the 2019 HyMap classification map, which has an overall accuracy of 85.7 % and a Kappa statistic of 0.83. Thus, agreement of the classification scheme with test data can be considered good, falling slightly below the performance of HyMap in recent years (Table 4). Median user's accuracy of the models for identifying the species of concern, water hyacinth and water primrose were 83.3% and 66.7%, and the median producer's accuracy of the two species were lower at 62.5% and 50%. This suggests that the classification scheme may be missing some occurrences of those species in the field, but if it did identify them as present at a location, it is more likely that they are indeed present. The lower producer's accuracy may be a result of fewer training and test points for these two classes due to overall class sparsity in the study region.

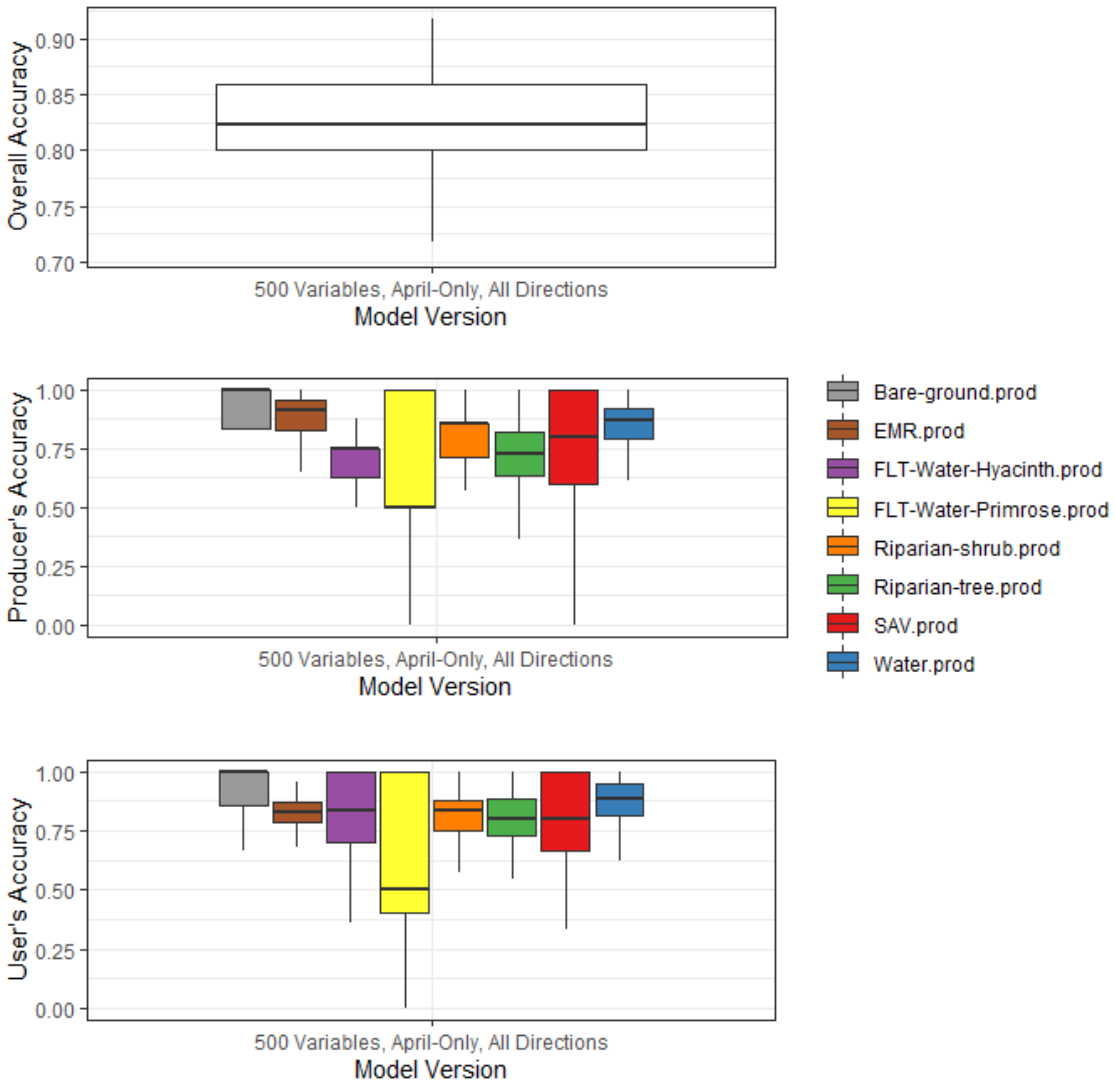


Figure 4. Overall, Producer's and User's accuracy ranges for all classes of the best performing selection of experimental treatment criteria for RF models. This set of models was constructed from the 500 most important variables with data acquired in both flight directions on April 9, 2019 only.

Of the 1000 RF model iterations, the best performing model was chosen for use in map-making. It had an overall accuracy of 91.8% and a Kappa of 0.90. For water hyacinth and water primrose, it had producer's accuracies of 87.5% and 100%, and user's accuracies of 100% and 50%. The low user's accuracy for water primrose is partially the result of low labelled data quantities. The high accuracies for water hyacinth are encouraging because at the time of acquisition what little hyacinth was present in the study area was senescent, which often makes it difficult to differentiate from other senescent vegetation. Due to the lack of healthy hyacinth to use for reference points, the HyMap classification used water hyacinth reference data collected in 2018 to classify water hyacinth in 2019 [24]. This likely lead to none of the water hyacinth in the study region being detected as such, only as NPV. From a management standpoint it is concerning that the user's accuracy of water

primrose for the selected 'best' Nano model was only 50% because that implies that half of sites identified would not contain primrose increasing operational time for treatment, but the high producer's accuracy means that if it was present in the area of concern it would be labelled properly and would not go untreated. It's reasonable to attribute the confusion of primrose and riparian shrubs to their similarity in appearance because water primrose appeared mostly inland in the study area and was not flowering at the time the imagery was acquired.

It is important to note that the selected model for map-making is in the upper quantile of the bootstrapped models, meaning it is unlikely that a repeat study would have a similar result without bootstrapping random forest iterations; but the goal in selection was to create the best map possible to fit the acquired data and the best candidate was selected. The selected model performed better than the HyMap model for 2019. One issue with the Nano model was the lack of labelled data. Having a larger set of labelled data may have further improved performance; however, this is uncertain due to lack of spectral information in the SWIR for the Nano sensor. The SWIR has been shown to be a useful predictor for classifying vegetation, specifically for discriminating between SAV, and floating or emergent vegetation [16], [76]. This is due to water's high absorption within this range of the spectrum, as well as compounds like cellulose and lignin [77].

Table 4. HyMap and Nano Classification Performance by Year

Year	Sensor	Classification Model	Overall Accuracy	Kappa Coefficient
2019	Nano	RF 153	0.918	0.901
2019	HyMap	UC Davis	0.857	0.830
2018	HyMap	UC Davis	0.908	0.900
2015	HyMap	UC Davis	0.952	0.939

3.2 Variable Importance and Variable Reduction

The 30 most important variables from the final model selected for map-making can be seen in Figure 5. These highlight the importance and usefulness of the vegetation indices calculated for classification. Reflectance at 503 nm (RF_503), and other remote sensing products calculated, including the fifth component from the minimum noise fraction transform (MNF_5) and mean texture occurrence metrics for bands 50, 52, and 158 (Mean_50, Mean_52, Mean_158) were important, though did not appear as frequently as indices.

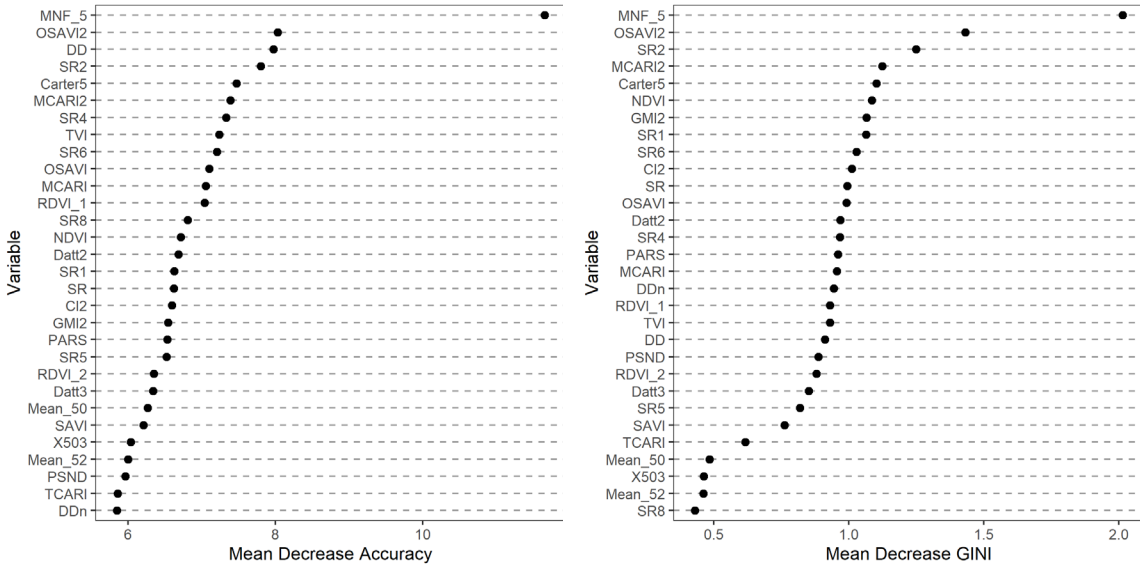


Figure 5. Variable importance plots from the RF model of the 30 most important variables used in map creation.

Variable importance of additional model sets (1000 bootstraps of the 10 models used in experiments) also showed that vegetation indices were considered most important during classification. The top subset of variables considered frequently most important to the classification can be seen in Figure 5. Other important variables across all model sets included mean kernel values from the texture analysis of bands in the NIR, and reflectance from the NIR. It is unsurprising that the NIR is important in addition to the vegetation indices. The majority of vegetation indices in this list use red and NIR bands using the large jump in reflectance that occurs near the red edge of the electromagnetic spectrum which enables separation of vegetation from other landcover classes. Wavelengths near the red edge correspond to biophysical and chemical properties of plants [78], which also enables discriminating between some classes, such as emergent vegetation and floating vegetation which typically has higher reflectance values in the NIR region [16]. Image segment also showed up as an important variable, which was expected due to its success when used for classification of high-resolution imagery of similar environments [79]. MNF 5, which showed up as the most important variable by far in final model used for map creation also appears in most important 500 variables, but does not occur as frequently (8042) which indicates that its usefulness in classification may be more dependent on the selected training and test data, rather than some of the other variables.

Table 5. Frequency of 90th Percentile Variables from the 500 most important. Green denotes a vegetation index, red a texture product, yellow a reflectance band, and white an image segment number.

Variable	Frequency (out of 10000)	Variable	Frequency (out of 10000)
Datt4	9997	SR1	9937
MSAVI	9997	GMI2	9935
MTVI	9997	Band 166 (768 nm)	9917
OSAVI	9997	DDn	9902
RDVI_1	9997	Band 156 (746 nm)	9900
SAVI	9997	Tex Mean Band 154 (741 nm)	9895
PARS	9996	Tex Mean Band 161 (757 nm)	9887
RDVI_2	9996	Green	9881
SR2	9993	Tex Mean Band 155 (743 nm)	9878
NDVI	9992	Segment	9877
SPVI	9990	Band 160 (755 nm)	9875
TVI	9988	Datt6	9871
SR	9985	Band 159 (752 nm)	9851
SR5	9985	Band 171 (779 nm)	9844
Tex Mean Band 156 (746 nm)	9974	Tex Mean Band 171 (779 nm)	9828
Datt5	9969	NDVI2	9821
PSSR	9969	Band 167 (770 nm)	9810
Band 154 (741 nm)	9969	Band 155 (743 nm)	9807
SR4	9961	Tex Mean Band 160 (755 nm)	9805
SR3	9960	mSR2	9802
PSND	9958	TCARI	9797
GMI1	9957	Tex Mean Band 210 (867 nm)	9787
NDVI3	9954	Tex Mean Band 164 (764 nm)	9772
MCARI	9949	Band 178 (795 nm)	9766
DWSI4	9945	Tex Mean Band 162 (759 nm)	9741

Reducing the overall quantity of variables for the final model had no effect median overall accuracy with each model performing at 82.4% though there was a slight increase in the variance for the reduced variable model. The minimal change in performance based on the number of variables used in the model was statistically significant based upon a Kruskal-Wallis test (χ^2= 23.609, df = 1, $p < 0.01$). For water hyacinth, the median producer's accuracy was significantly higher ($p < 0.01$) for the 500 variable model set at 75% and the median user's accuracy was the same ($p = 0.08$) at 83.3%. For water primrose, the median producer's accuracy was the same between model sets at 50%, but the data had significantly different distributions ($p < 0.01$). The user's accuracy for water primrose was higher ($p < 0.01$) for the model using all variables at 66.7%.

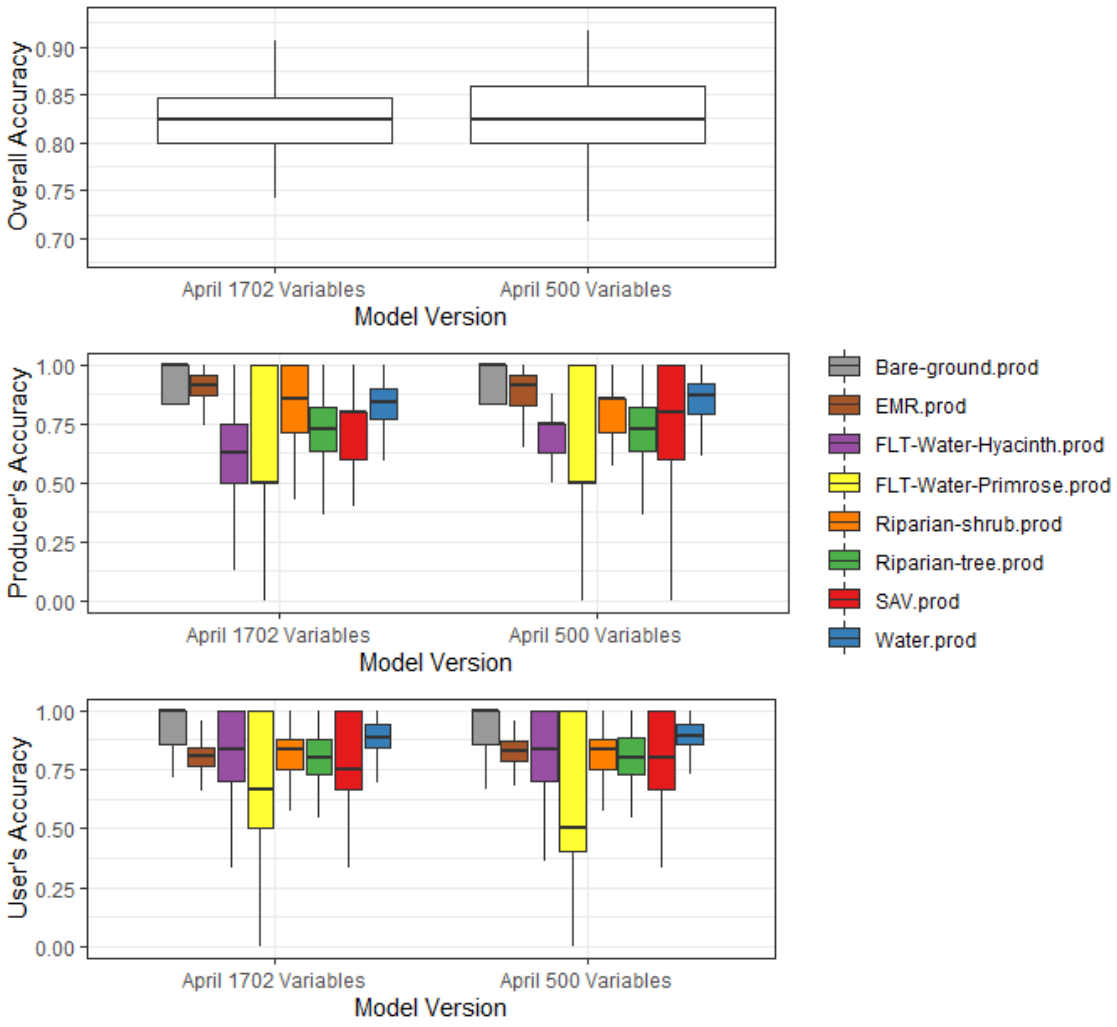


Figure 6. Overall, producer's and user's accuracies for variable reduction with model sets using all of the calculated variables (left) and the 500 most important variables only (right).

These results aligned with our expectations that variable reduction may impact accuracies in various ways but should not drastically reduce performance. The reduction in variables considered a necessity due to the impossibly high computational overhead associated with keeping all 1702 variables, a challenge common in the literature of RF classification for remote sensing. Data reduction has long been an objective of imaging spectroscopy studies to reduce computational overhead and data redundancy [80], [81] and studies have also shown that data reduction often considered necessary has improved model performance if only the most important variables are used [82]. RF is relatively robust to redundant input variables [66], though not necessarily correlated predictor variables, which may result in inflated accuracy results[67]. While RF is robust to redundant data, and computational burdens are lessened by the rapid advances and access to HPC computing across the remote sensing community, data volumes continue to grow, especially for extremely high spatial resolution, high spectral resolution unmanned aircraft-applications. Thus, eliminating redundant and unimportant variables was helpful. In this

study, the high quantity of spectral data associated with imaging spectroscopy appears useful, especially in the NIR, with several bands in the 740-770 nm region being identified as important. Conversely this could be a result of the variables being correlated which would inflate accuracy metrics. Common unmanned aircraft mountable multispectral sensors have a single NIR band located above 800nm which would miss much of this key information. Although this is the case, it is also important to consider the effect that the abundance of noise in the NIR region of the Nano spectra likely had. Noise increased upon proximity to the largest detectable wavelength (1000 nm), potentially reducing usefulness of data in this region, which would have in-turn reduced its importance in the model.

3.3 Experiment 1: Sensitivity to Training and Test Data

Varying the percentages of total data used for training and testing affected the performance of the model and served as an important predictor of model accuracy (Figure 7). This was expected as varying quantities of training data used has been shown to effect performance [83]. In this case, increasing training data quantity increased accuracies. A Kruskal-Wallis test confirmed that at least one of the overall accuracies was significantly different (χ^2= 155.72, df = 4, $p < 0.01$). Further examination using a pairwise Wilcoxon test showed that all models had statistically significant distributions of overall accuracy and kappa except for the 19% and 26% test pair (Table 6). Improvements in accuracy have been shown to occur as training data quantities are increased [83], and the same occurred here. It's likely this is because less training data means a less representative sample on which the model bases votes. The range of these results for each model set also highlights the dependency of model performance on training data selection, and how mislabeled data can affect modelling results. There are probably some mislabeled/non-representative labelled data used here as a result of due to slight geographic misalignment of the imagery or the photointerpretation used to label some reference sites, and those effects can be seen here.

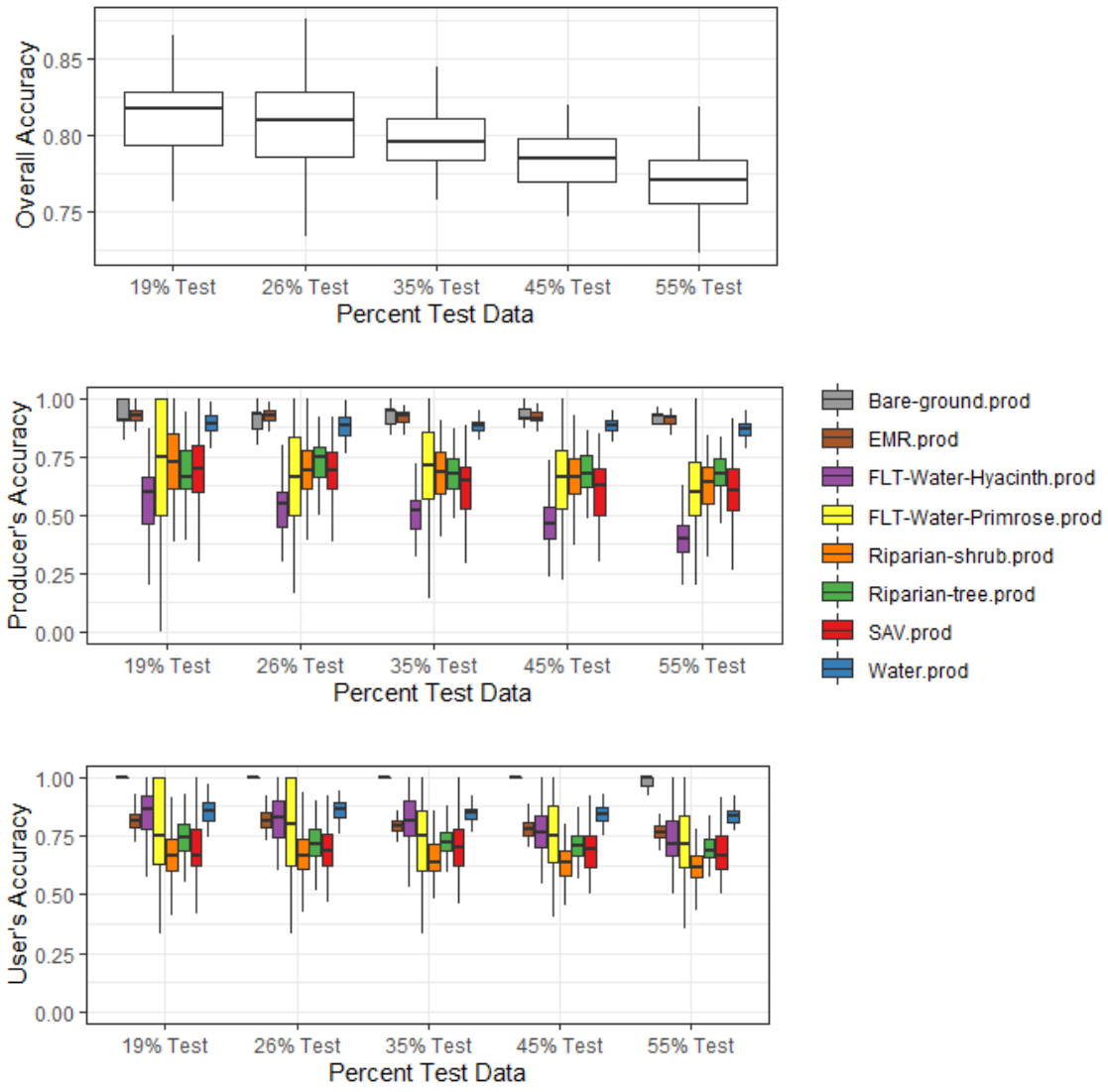


Figure 7. Overall, producer's and user's accuracies for model sets for experiment 1: examining variation in testing and training data percentages.

Table 6. Experiment 1 Wilcoxon Rank-Sum Test Statistics

Wilcoxon Rank Sum Test p-values

Overall Accuracy				
	19% Test	26% Test	35% Test	45% Test
26% Test	0.15	NA	NA	NA
35% Test	< 0.001	0.004	NA	NA
45% Test	< 0.001	< 0.001	< 0.001	NA
55% Test	< 0.001	< 0.001	< 0.001	< 0.001
Water Hyacinth Producer's Accuracy				
	19% Test	26% Test	35% Test	45% Test
26% Test	0.075	NA	NA	NA
35% Test	0.015	0.540	NA	NA
45% Test	< 0.001	< 0.001	0.001	NA
55% Test	< 0.001	< 0.001	< 0.001	0.005
Water Hyacinth User's Accuracy				
	19% Test	26% Test	35% Test	45% Test
26% Test	0.011	NA	NA	NA
35% Test	0.011	0.996	NA	NA
45% Test	< 0.001	0.004	0.004	NA
55% Test	< 0.001	< 0.001	< 0.001	0.034
Water Primrose Producer's Accuracy				
	19% Test	26% Test	35% Test	45% Test
26% Test	0.020	NA	NA	NA
35% Test	0.012	0.040	NA	NA
45% Test	< 0.001	0.078	< 0.001	NA
55% Test	< 0.001	0.072	< 0.001	0.957
Water Primrose User's Accuracy				
	19% Test	26% Test	35% Test	45% Test
26% Test	0.854	NA	NA	NA
35% Test	0.440	0.440	NA	NA
45% Test	0.497	0.464	0.854	NA
55% Test	0.280	0.280	0.584	0.440

Median producer's and user's accuracies of water hyacinth varied from 40 - 60% and 71.7- 86.6% based upon the training and test data split (Figure 7). For each class specific accuracy and training/test split, a pairwise Wilcoxon test was conducted showing that all pairs of user's and producer's accuracies were significantly different at the established p-value of 0.05, except for the 35% and 26% test splits (Table 6). For water primrose, producer's and user's median accuracies for this experiment varied from 60 – 75% and 71.4 to 80%. Producer's accuracies for primrose were significantly different except 45%

and 55% test, 26% and 55% test, and 26 and 45% (Table 6). Pairwise comparison of user's accuracies for primrose are quite different, with no significant differences in distributions between the splits (Table 6). This is likely a result of the low quantity of reference data for water primrose. When restricted to a single direction and date, there were less than 10 labelled data points for this class, much fewer than any other class. This very low quantity would greatly limit the combinations of training and test data possible and widen the variability of accuracies based upon the quality of points which were selected during training. This small quantity is representative of the ratio of primrose present in the scene, which in hindsight should have been increased using a similar procedure to water hyacinth. Although imbalances in training data relative to presence in a scene do reduce classification accuracy, it has been shown that increasing quantities of sparse classes, such as primrose here, does not adversely affect classification accuracies [44]. For mapping purposes, the 65% training and 35% test data quantities were chosen as a middle-of-the-road path because these quantities increased model performance without broadening the range of accuracy too severely. The 74% training 26% test split was also considered because of the insignificant differences between the class specific accuracies of the two species of concern, but this was decided against because of the increase in variance resulting from the decreased test data quantities.

3.4 Experiment 2: Sensitivity to Flight Direction

Flight direction of training and test data had a significant impact on model overall accuracy based upon a Kruskal-Wallis ($\chi^2 = 917.85$, df =3, $p < 0.01$), showing significant difference in overall accuracy distributions of the experimental treatments. Further examination using a pairwise Wilcoxon test shows that the overall accuracy of the model trained on both flight directions but applied to away-from-solar-plane data (Both-Away) performed best by a significant margin (Table 7). Other pairs also showed significantly different distributions except for models Away-Away and Both-Toward (Table 7). It was expected that there may be a difference in reflectance due to flight direction because of the bidirectional reflectance distribution function (BRDF). BRDF is a function resulting from the solar incidence angle, instrument viewing geometry, physical properties of the target, and shadow effects [38]. Studies have shown that these variances in reflectance will propagate through to the product maps [84]. Though some variation was expected, flying toward or away from the solar plane was expected to minimize cross-track effects [85], it was unclear exactly how performance would vary between those directions. Perhaps the use of class reflectances collected in both flight directions better represented the interclass variability due to BRDF effects, which led to a better performing model. BRDF correction has been considered extremely important for unmanned aircraft collected data because variation in reflectances due to geometry and illumination are more prevalent in unmanned aircraft imagery due to the fine spatial scale [86]. Some suspected causes of these resulting differences in model accuracy are wind direction, the slight misalignment of the sensor from the nadir angle, and platform instability. Although other environmental factors cannot be dismissed, flight design was supposed to marginalize cross-track illumination due to solar position relative to the sensor because flight lines were collected toward and away from the solar plane only. By looking at solar positions from Table 3, it is likely that the

sun moved enough during the flight to have an impact on reflectances as well. One potential reason that the model trained on both directions performed better could be that the additional variation in reflectance is mostly based upon physical properties and shadow on the ground, providing a better sample of the potential spectra of each class.

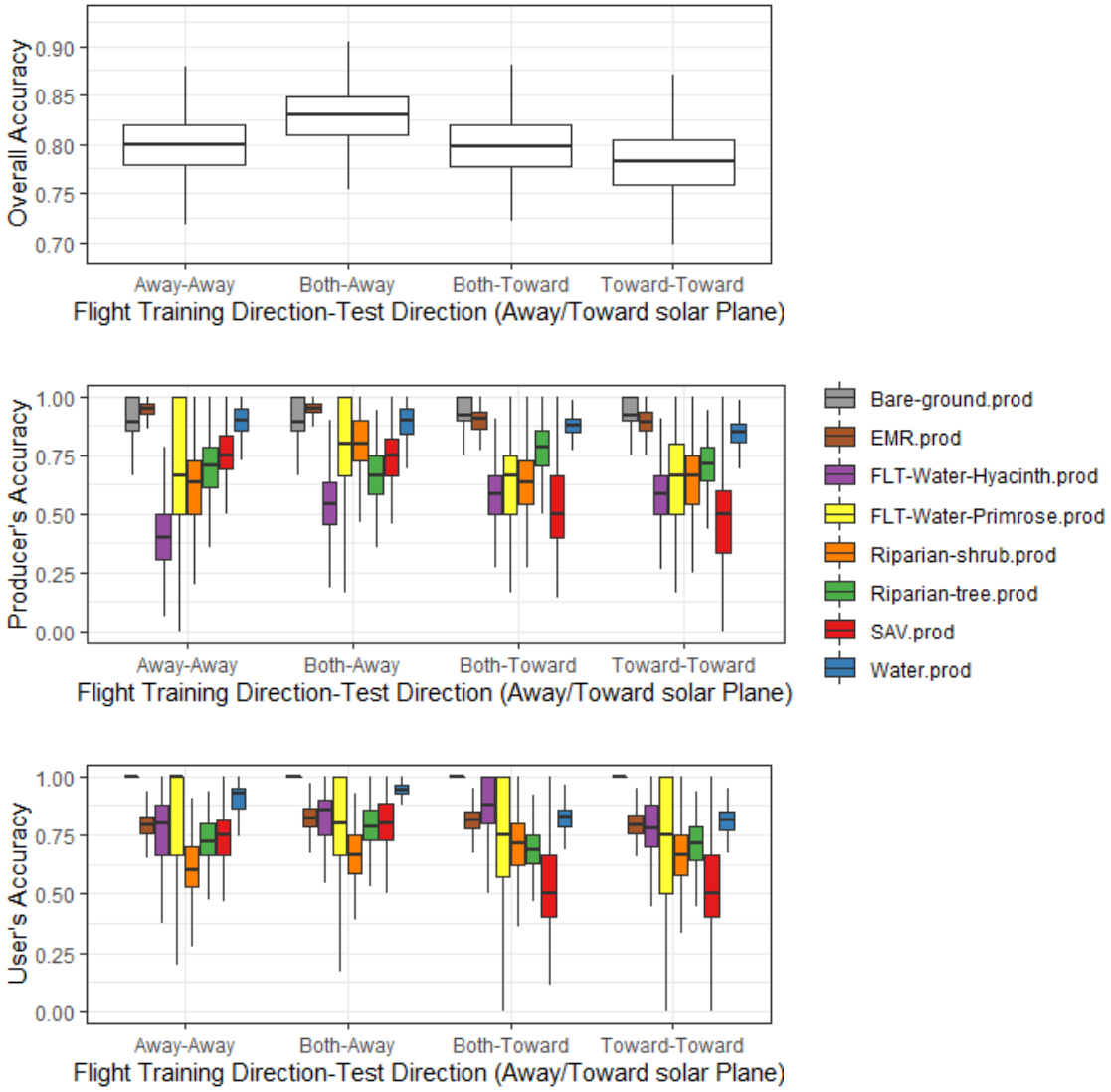


Figure 8. Overall, producer's and user's accuracies for experiment 2: varying flight direction.

Table 7. Experiment 2 Wilcoxon Rank-Sum Test Statistics

Wilcoxon Rank Sum Test p-values

Overall Accuracy			
	Away-Away	Both-Away	Both-Toward
Both-Away	< 0.001	NA	NA
Both-Toward	0.573	< 0.001	NA
Toward-Toward	< 0.001	< 0.001	< 0.001
Water Hyacinth Producer's Accuracy			
	Away-Away	Both-Away	Both-Toward
Both-Away	< 0.001	NA	NA
Both-Toward	< 0.001	< 0.001	NA
Toward-Toward	< 0.001	< 0.001	0.742
Water Hyacinth User's Accuracy			
	Away-Away	Both-Away	Both-Toward
Both-Away	< 0.001	NA	NA
Both-Toward	< 0.001	< 0.001	NA
Toward-Toward	0.198	< 0.001	< 0.001
Water Primrose Producer's Accuracy			
	Away-Away	Both-Away	Both-Toward
Both-Away	< 0.001	NA	NA
Both-Toward	< 0.001	< 0.001	NA
Toward-Toward	< 0.001	< 0.001	0.237
Water Primrose User's Accuracy			
	Away-Away	Both-Away	Both-Toward
Both-Away	0.057	NA	NA
Both-Toward	< 0.001	< 0.001	NA
Toward-Toward	< 0.001	< 0.001	0.634

Flight direction also affected class specific accuracies. Water hyacinth producer's and user's accuracies ranged from 40 – 58.3% and 77.7 – 85.7% (Figure 8). Wilcoxon rank sum tests of producer's accuracies showed significant difference between all model sets for water hyacinth except Both-Toward and Toward-Toward (Table 7). Wilcoxon tests showed significant differences between all pairs of user's accuracies for water hyacinth except Toward-Toward and Away-Away. For water primrose, producer's and user's accuracies ranged from 66.7 – 80% and 75 – 100%. Producer's accuracy for water primrose was significantly different between all model set pairs except Both-Toward and Toward-Toward (Table 7). User's accuracy for primrose were significantly different for all pairs except Both-Toward and Toward-Toward, and Both-Away and Away-Away (Table 7). Considering these metrics, the Both-Toward model set performed best for classifying water hyacinth and the Both-Away models performed best for classifying water primrose. A potential explanation for this could be the differences appearances between the two species.

There was a large difference in producer's and user's accuracies for the water and SAV classes, which had their highest accuracies and less variability between the first and fourth quantiles with the Both-Away model set (Figure 8). Median producer's and user's accuracies for water in this model set were 89.5% and 94.4 %, and producer's and user's accuracies for SAV were 75% and 80%. Pairwise tests confirmed significantly different distributions for the Both-Away model sets for these two classes, except for the producer's accuracy of SAV, which was not significantly different from the Away-Away model set (Table 7). The improved performance specific to these classes as well as overall when training data was from the away direction could be because there was much more specular reflectance of water present in the toward solar plane imagery. Based upon the overall improved performance from the Both-Away model, in addition to its higher performance for the water primrose class it was decided that the map-making model would be trained using both flight directions and applied to imagery collected in the away from solar plane direction.

3.5 Experiment 3: Sensitivity to Acquisition Date

Acquisition date had an impact on overall accuracy(Figure 9), with a Kruskal-Wallis test showing at least one model having significantly different performance ($\chi^2 = 580.78$, df =5, $p < 0.01$). Pairwise comparison revealed that all experimental treatments were significantly different from one another except May-All and May-May (Table 8), with April-April having the highest median accuracy at 83.4%. Producer's accuracies for water hyacinth ranged from 25 – 75% with May-May and June-June performing the best with non-significant differences in distributions (Table 8). User's accuracies for water hyacinth ranged from 66 - 100% with May-All performing best, but not being significantly different from April-All or April-April (Table 9).

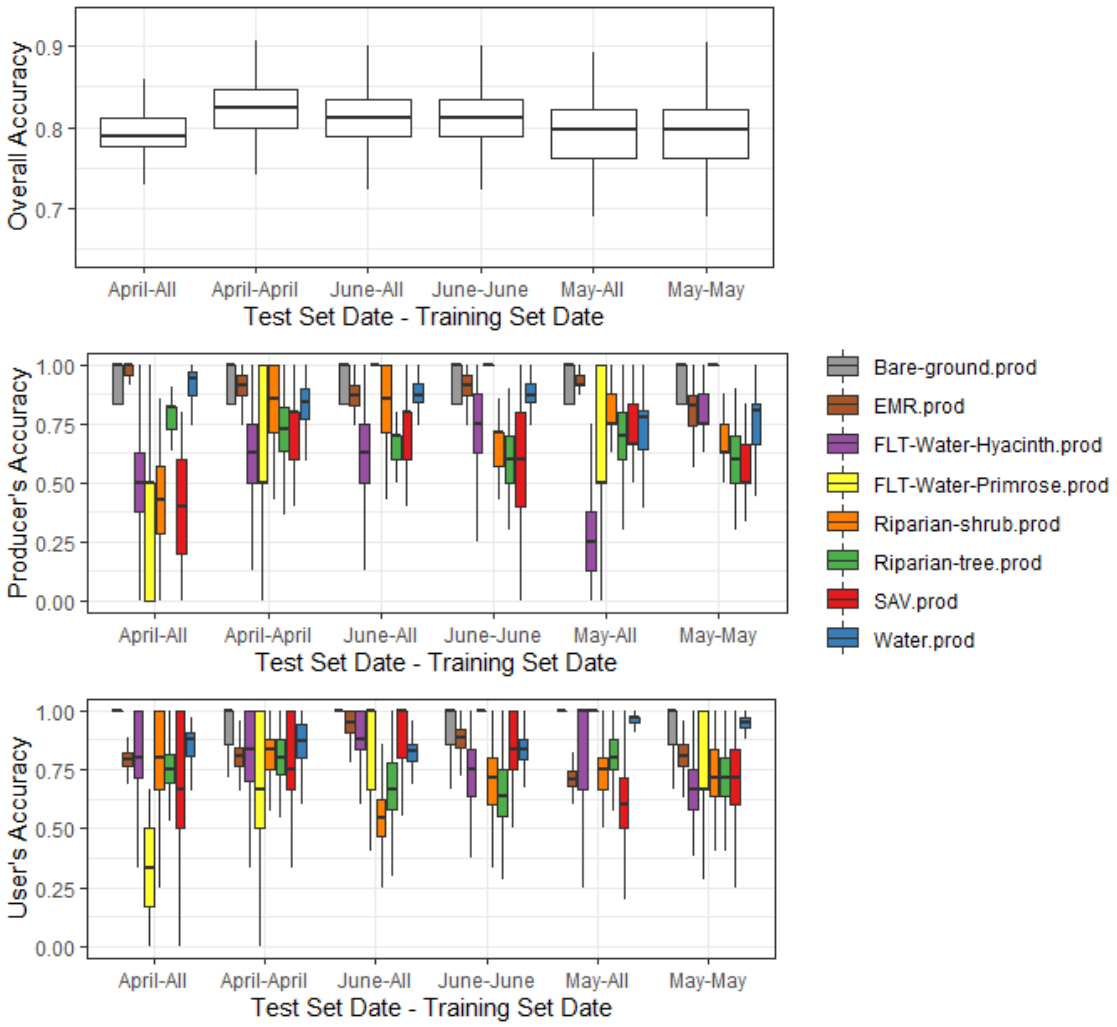


Figure 9. Overall, producer's and user's Accuracies for experiment 3: using multiple vs. single dates as training and test data.

Table 8. Experiment 3 Wilcoxon Rank-Sum Test Statistics

Wilcoxon Rank Sum Test p-values

Overall Accuracy

	April-All	April-April	June-All	June-June	May-All
April-April	< 0.001	NA	NA	NA	NA
June-All	< 0.001	< 0.001	NA	NA	NA
June-June	< 0.001	< 0.001	0.663	NA	NA
May-All	< 0.001	< 0.001	< 0.001	< 0.001	NA
May-May	0.013	< 0.001	< 0.001	< 0.001	0.190

Water Hyacinth Producer's Accuracy

	April-All	April-April	June-All	June-June	May-All
April-April	< 0.001	NA	NA	NA	NA
June-All	< 0.001	0.514	NA	NA	NA
June-June	< 0.001	< 0.001	< 0.001	NA	NA
May-All	< 0.001	< 0.001	< 0.001	< 0.001	NA
May-May	< 0.001	< 0.001	< 0.001	0.439	< 0.001

Water Hyacinth User's Accuracy

	April-All	April-April	June-All	June-June	May-All
April-April	0.650	NA	NA	NA	NA
June-All	< 0.001	< 0.001	NA	NA	NA
June-June	< 0.001	< 0.001	< 0.001	NA	NA
May-All	0.789	0.931	< 0.001	< 0.001	NA
May-May	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Water Primrose Producer's Accuracy

	April-All	April-April	June-All	June-June	May-All
April-April	< 0.001	NA	NA	NA	NA
June-All	< 0.001	< 0.001	NA	NA	NA
June-June	< 0.001	< 0.001	0.049	NA	NA
May-All	< 0.001	< 0.001	< 0.001	< 0.001	NA
May-May	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

Water Primrose User's Accuracy

	April-All	April-April	June-All	June-June	May-All
April-April	< 0.001	NA	NA	NA	NA
June-All	< 0.001	< 0.001	NA	NA	NA
June-June	< 0.001	< 0.001	< 0.001	NA	NA
May-All	< 0.001	< 0.001	< 0.001	< 0.001	NA
May-May	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

During ground data reference collection, visible changes in species were observed between April and June, increasing the variability of labelled data for those classes. Water hyacinth initially appeared in April as senescent and greened up during the April to June

period, the typical growing season for water hyacinth [87], while a large portion of riparian shrubs along the shore senesced as the spring season passed. These and other phenological changes may have led to this trend in accuracies with individual acquisition trained and tested models performing better.

For water primrose, the median producer's accuracies ranged from 50 – 100%, with June-All, June-June, and May-May performing best at 100% accuracy. Pairwise tests showed a significant difference between all of these, but it is likely an artifact of very small number of labelled data for the primrose class for testing because there was no variance in the results across the 1000 bootstraps of the selected data for these model sets for this class. User's accuracies for water primrose from 33 – 100%, with May-All and June-June performing best at 100%. These producer's and user's accuracies for water primrose can likely be attributed to the low quantities of labelled primrose data.

While previous studies have exploited differences in phenology across multi-date imagery to improve species discrimination [88]–[90], and Hestir et al. [12] and Khanna et al. [30] suggest the case could be true for the target species in the Delta, it was not effective for improving species discrimination for this study; likely because of the experimental setup and low quantities of labelled data. We considered using the geographic duplicates from the second flight direction but did not want to inflate accuracy statistics. Other studies have shown similar challenges when using small training sets to discriminate highly variable classes with spectral overlap. For example, Millard and Richardson [67] found that increasing quantity of reference points for training and testing increased independent accuracy of models, but that spatial autocorrelation should be minimized because it was found to cause an increase in error during independent accuracy assessments.

The April-trained, April-tested model was selected for use because the imagery was collected concurrently with the HyMap. In addition, the model's improved overall accuracy from the other acquisition dates coupled with the inconclusiveness of the primrose class accuracy statistics supported omission of the imagery from other dates. Although none of the ground reference locations had different species when visited in May and June, there were visible changes to the study area, which were thought to potentially impact classification. With the visible changes associated with the study site and the limited single acquisition date of April for HyMap, this choice seemed further justified because the labelled data needed to match up as best as possible between the two imagery sets.

3.6 Map Comparison

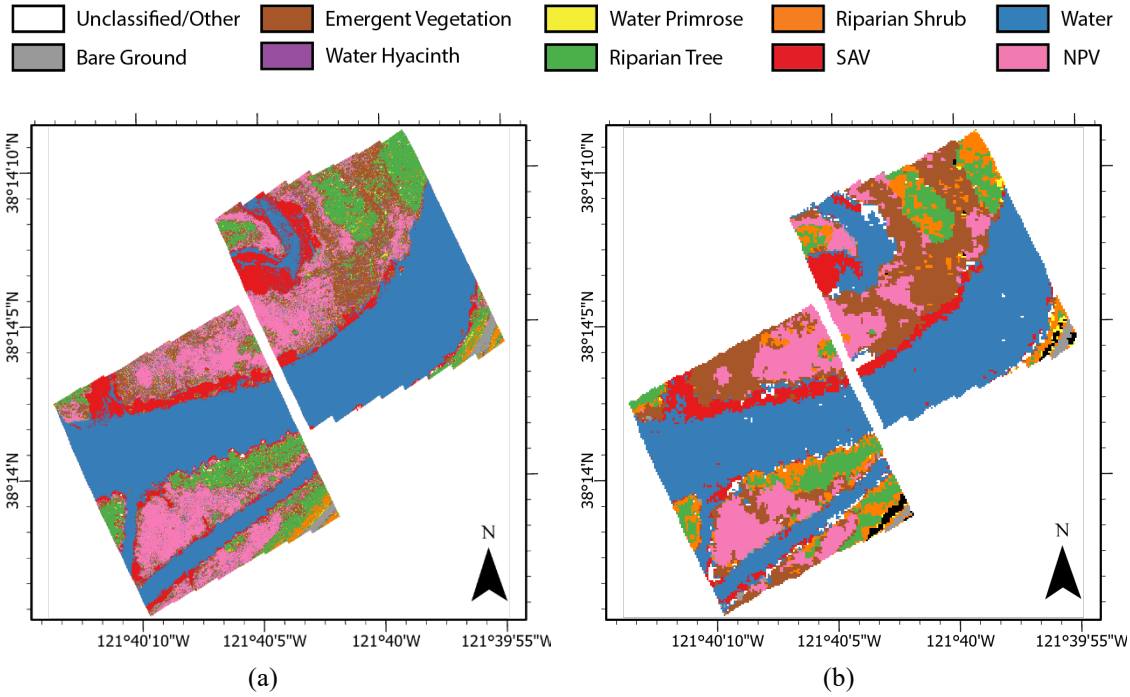


Figure 10. Classification Maps of the study region. (a) Nano classification and (b) HyMap classification

3.6.1 Comparison 1: Total Class Area

Total class area comparisons between the Nano and HyMap maps highlight the usefulness of the Nano for detecting smaller patches of classes. The Nano map showed 1589 $m^{2,}$ of water hyacinth, while the HyMap map had no occurrences (Table 8), illustrating the differences due to spatial mapping scale. Ground reference data collection confirmed water hyacinth presence in the region and the amount is significant considering its rapid growth rate [34]. The majority of water hyacinth in the Delta was senescent during April, leading to a low quantity of labelled data for the class. During the HyMap classification process, imagery from prior years was used to compensate for the lack of labelled data [24]. The earlier acquisition date may have also affected mapping performance of other classes due to different phenological stages of plant species, potentially a cause for the decrease in overall HyMap accuracy from 2018 to 2019 (Table 4).

Table 9. Map Comparison 1: Class area coverage for the Nano and nearest neighbor resampled Nano maps and percent difference for each from HyMap (reference).

Class	HyMap Area (m^2)	Nano Area (m^2)	Nano Resampled Area (m^2)	NanoDifference	Nano ResampledDifference
Unclassified	128570.32	126528.80	126899.9	-1.59%	-1.30%
Bare ground	858.33	930.35	956.59	8.39%	11.45%
EMR	19097.12	12077.12	14686.98	-36.76%	-23.09%
Water Hyacinth	0	1589.43	124.27	NA	NA
Water Primrose	349.69	1616.53	317.9	362.28%	-9.09%
Riparian Shrub	7262.57	2217.58	916.13	-69.47%	-87.39%
Riparian Tree	7690.29	10365.43	10539.83	34.79%	37.05%
SAV	6988.02	9022.95	8594.86	29.12%	22.99%
Water	41269.20	44822.89	42422.31	8.61%	2.79%
Other	679.15	0	0	NA	NA
NPV	13577.22	16989.30	20883.14	25.13%	53.81%

Typically, direct area comparisons are not helpful when assessing classification performance because errors of omission and commission can cancel each other out, in which case having good area matchups between classes may have little meaning [91]. However, in this situation the area comparison highlighted usefulness of the Nano's higher spatial resolution via the detection of water hyacinth and the differences in areal coverage of classes gives insight into potentially improved detection capabilities. Comparing these to the nearest neighbor resampled class areas shows how the total areas of more sparse classes, such as water primrose, were drastically lower after the resampling process. If the same loss of detail is occurring in the HyMap map, some invasive species may be missed, allowing them to spread. The Nano classification found significantly higher quantities of water primrose (360% more) than the HyMap classification, with an area of 1617 m^2 while the resampled Nano map showed only 318 m^2.

Additional comparison of class area coverage showed that the classes with the best matchup between the Nano and the HyMap are water, with an 8.61% difference and bare ground, with an 8.39% difference within the study region. These classes are may be in high agreement because they are spectrally unique in the region, offering little to be confused with in the image area. The unclassified areas can be ignored because they consist mostly of masked areas which exclude all values from the HyMap except those contained within the Nano flight area, the only exception being a few pixels in each image where shadowing was very heavy. NPV and SAV show the next best match between class areas with difference in total area coverage are SAV and bare ground with 25.13 % and 29.12 % differences respectively. All other classes differed by 30% or more in total area coverage. As mentioned earlier, a potential cause of these discrepancies could be the difference in spatial scales. Reducing contents of a large (1.7m pixel) to a single class, is unrealistic for heterogenous regions such as wetlands. The Nano image would contain more than 1100 pixels in that same area (0.052 m pixel), capable of classification on a much smaller spatial scale. Many of the smaller patches of vegetation would be obscured by the dominant class

present in the larger pixel. The dominance of other classes within the study region can be seen by the reduction in area of the water hyacinth when the HyMap is compared to the Nano. Additionally, when a nearest neighbor resampling was done of the Nano data there was also a large reduction in sparse classes. Though nearest neighbor is not a majority resampling technique where this is expected [92], it occurs as a consequence of the very low quantities of these classes in the region. It is less likely that one of these classes will be at the location where the nearest neighbor pixel is sampled for upscaling. This is also supported by knowledge of the study area from field observations while conducting ground reference sampling. Both water hyacinth and primrose in the study area consisted of small patches mixed in with other classes. Another explanation for mismatch between the area coverage is a potential mismatch in class definitions. For example, riparian shrubs and trees, have a high level of disagreement (30%+), even with the resampled Nano areas but they share a similar overall area when combined, 12583.01 m^2 and 14952.86 m^2 only an approximately 12% difference. Whether these differences in areas are a consequence of class definitions or simply confusion between classes of similar appearance should become apparent during the site-specific comparison.

3.6.2 Comparison 2: HyMap Pixel to Nano Subpixels by Class

The two maps have a lower agreement for more sparse classes, with water hyacinth at 0% and water primrose at 6.9% (Table 9). More homogenous classes like water, NPV and SAV had higher agreements. The Nano map captured more of the complexity of the study area, showing a larger presence of less dominant classes based on the area comparison. This adds a spatial component revealing that either the Nano is successfully mapping at a finer scale, or that there is noise present in the Nano resulting in misclassification. A surprising result was the 48.9% agreement for bare ground because it was expected to be mixed within some of the other classes. This was the only class where agreement here did not trend inversely with the difference percentages calculated in the first comparison. One explanation for this could be the large homogenous area of road that accounts for much of the bare ground within the study site.

Table 10. Map Comparison 2: Percent Agreement of Nano polygons with HyMap Pixels by class.

Class	Percent Agreement
Bare ground	0.489
EMR	0.323
Water Hyacinth	0
Water Primrose	0.069
Riparian Shrub	0.103
Riparian Tree	0.637
SAV	0.702
Water	0.924
Other	NA
NPV	0.706

3.6.3 Comparison 3: Moving window mode adjusted Nano

The moving window adjusted Nano map samples showed a showing a 44.7% overall accuracy (agreement) when the HyMap classification was used as reference. Table 8 shows the producer's and user's accuracies calculated with the HyMap map as reference.

Table 11. Map Comparison 3 and 4: User's and Producer's Accuracies (agreement) for each class using HyMap as a reference.

Map	Species	Nano-MW	Nano-Resampled
User's Accuracy (Commission)	Unclassified	87.4%	100.0%
	Bare ground	40.2%	28.2%
	EMR	38.6%	62.0%
	Water Hyacinth	NA	0.0%
	Water Primrose	NA	0.0%
	Riparian Shrub	6.0%	29.8%
	Riparian Tree	39.5%	45.7%
	SAV	58.1%	35.0%
	Water	55.2%	88.5%
	Other	NA	NA
	NPV	48.8%	52.5%
Producer's Accuracy (Omission)	Unclassified	98.0%	98.7%
	Bare ground	37.0%	36.0%
	EMR	39.0%	41.0%
	Water Hyacinth	NA	NA
	Water Primrose	0.0%	0.0%
	Riparian Shrub	5.0%	3.9%
	Riparian Tree	77.0%	73.1%
	SAV	54.0%	49.2%
	Water	90.0%	88.4%
	Other	0.0%	0.0%
	NPV	78.0%	81.4%

This table shows poor producer's accuracies (<40%) for some classes, such as bare ground, EMR, riparian shrub, and water primrose. A possibility for the low agreement for bare ground is spatial scale. Bare ground appears in several individual or small groups of pixels in the imagery due to canopy cover. This is probably also true for primrose, of which no Nano pixels sampled in the manner were. EMR seems to be getting confused with NPV frequently, which is not a surprise due to the similarity in appearance and the near senescent state of those species in April. Riparian shrubs also have a very low producer's accuracy at 5%, which is surprising, but probably due to the idea of mismatched class definitions with riparian trees, which were confused with shrubs several times, but only in one direction. Riparian trees were not confused with shrubs at a high rate when trees were the reference in the HyMap data. The classes of NPV, Water, and riparian trees all showed high producer's accuracies (>70%), while SAV showed a 51% producer's accuracy between the maps. This could easily be due to the variability in specular reflectance or surface appearance of water, or the patchiness of SAV only detectable at higher spatial resolutions. The differences in distribution between the sampled points in moving window

Nano map and the HyMap map were significant based upon a χ^2 test (χ^2 =2217.4, df = 49, $p < 0.01$) but did appear to be moderately correlated with each other based on Cramér's V (V = 0.630). This comparison omitted the water hyacinth and water primrose classes because they were absent from the sampled data from one of the maps, making the mapped locations of those classes significantly different. Having zero observed points also caused expected values in the χ^2 test denominator to be zero, prohibiting its use with those classes included. Two concerning issue with χ^2 were the extremely high value, and some expected frequencies being very low, which would support use of an exact test instead [93]. The adjusted Rand index calculation was conducted including all classes (ARI = 0.343), showing some similarity between maps.

3.6.4 Comparison 4: Nearest Neighbor Resampled Nano

A comparison of the nearest neighbor resampled Nano map to the spatial resolution of the HyMap map resulted in moderate to high agreement between maps. Using HyMap as a reference, the overall percent agreement between maps was 85%. It is not surprising that this has a higher level of agreement for than the moving window pixel to pixel comparison for two reasons. The moving window pixel to pixel comparison functions as a majority aggregation, picking the center pixel out of a region of over 1000 pixels (Nano-Hyperspec) and reassigning it based on the mode. This altered the Nano map significantly, increasing the quantity of dominant classes like water, and decreasing the quantities of sparse classes like primrose as. This was expected from a majority aggregation [92], but reduced primrose to so few occurrences that none remaining were in agreement with the HyMap map. The nearest neighbor method proved to be a less biased resampling, leaving more occurrences of primrose, and a higher level of agreement with the HyMap. Table 7 shows the similarity in producer's accuracy between the comparisons, which suggests either consistency of the Nano classification relative to the HyMap, or the correlation between the Nano's central pixel relative to each HyMap pixel and the nearest neighbor resampling method. The difference in user's accuracy can be attributed to general misidentification of classes in the model intensified by the nearest neighbor resampling process.

There was a significant difference in class distribution between the sampled points from each map (χ^2 =2140.1, df = 49, $p < 0.01$) and a moderately strong correlation between the two distributions (V= 0.618). Similar to comparison 3, the classes of water hyacinth and primrose were omitted from these statistical comparisons for the same reasons. Likewise, the very high χ^2 values cast doubt on the usefulness of these statistical metrics in this scenario. Adjusted Rand index (ARI = 0.344) shows some correlation between the distributions of sample points between classes.

The statistical tools here show that the maps are statistically different but somewhat correlated; however, based upon area comparisons and visual inspection of the maps, it's clear that classes generally occur in similar quantities and within the same spatial location in both maps, even if there is a slight geolocational mismatch. It seems unusual that they are so different statistically considering the models used for creating both had similar performance. From this one can surmise that the differences can likely be attributed to the comparative process. Two potential factors causing the statistical results could be that: 1)

The consequences of the extremely different minimum mapping units, such as spectral mixing or effects from resampling, and 2) geographic misalignment between the maps.

3.6.5 Qualitative Assessment

Additional qualitative assessment of the maps suggest that the Nano has the capability to detect smaller patches of species as expected, which is important to locating invasive plants for treatment. This is further supported from in-field experience in the study site and visual assessment of species and quantities present in the area. For example, water hyacinth patches mapped in the Nano imagery were small and mixed in with emergent and non-photosynthetic vegetation. Another example is a small tule colony visible only to the Nano can be found in Figure 9. Not only is the tule colony visible, but a small region of SAV is visible along the side, which was typical of the individual tule colonies seen in the field. This highlights the ability of the higher resolution unmanned aircraft technology to obtain this more detailed information, which may be helpful for management of invasive species in the region.

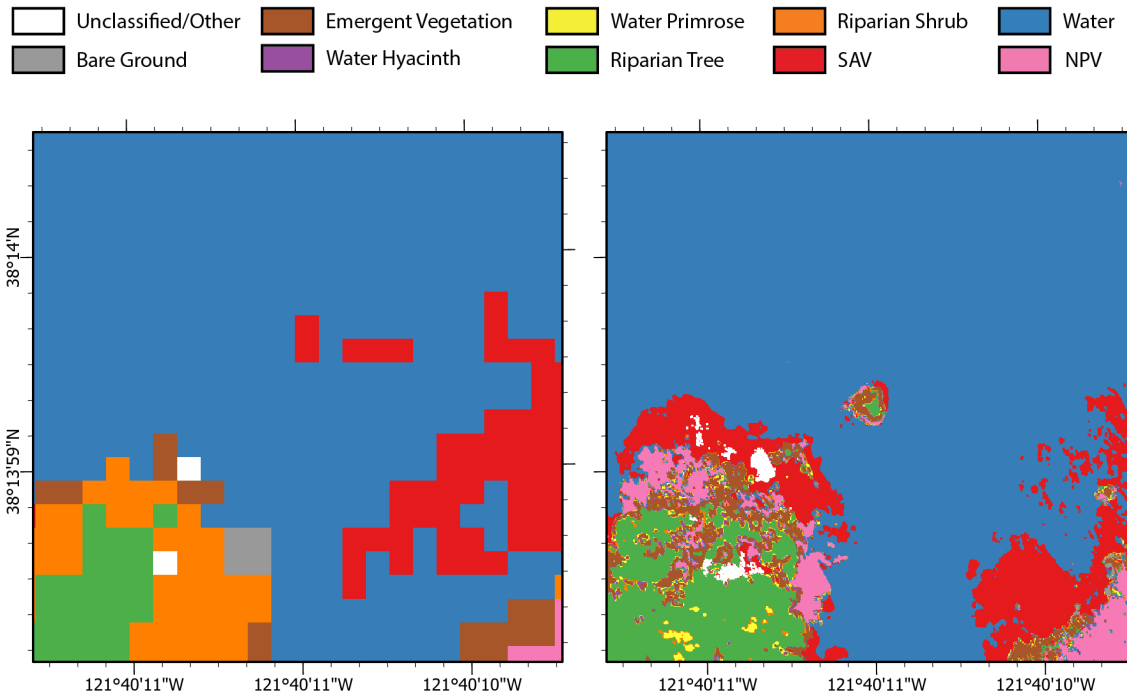


Figure 11. Close-up of a tule colony showing the improved spatial detail of the Nano classification (right) versus the HyMap classification (left).

4. MANAGEMENT RELEVANCE AND OPERATIONAL CONSIDERATIONS

The higher spatial resolution unmanned aircraft-based sensors were capable of detecting water hyacinth where the lower spatial resolution manned flight could not, which is beneficial for management. This capability to detect smaller patches allows for early detection and herbicide applications; however, the small footprint of unmanned aircraft operations would make this process difficult to implement broadly. Specific, targeted deployment at problem areas identified from the previous growing season is likely the best implementation of this process. In addition to increased detail possible with the unmanned aircraft, flexibility of deployment could be a benefit. The 2019 HyMap flight over the Delta was conducted in early April, slightly too soon to catch the greening of water hyacinth at a large scale. This is evidenced by the lack of training data reported by CSTARS and their use of data from previous years to map water hyacinth during 2019, likely resulting in their lower accuracies for that year (Table 4). Maps built using the unmanned aircraft imagery had a similar accuracy range to that of those created with HyMap during more successful years, but also provided a much higher spatial resolution. As mentioned, the primary disadvantage of using unmanned aircraft would be the cost and logistics of putting large scale unmanned aircraft mapping into practice.

Table 12. Nano and HyMap Operational Information and Cost Estimates

Sensor	Area Covered (hectares)	Time Estimate (hrs)	Approximate Data Volume (GB)	Approximate Deployment Costs
HyMap	74123.42	5	700	$150,000*
Nano - This Study	10.53	1	600	$62,780
Nano - Entire Delta	74123.42	7040	2111880	$865,014

*Estimate, pers. Comm, Shruti Khanna (2020)

The small spatial footprint relative to the HyMap would require thousands of flights, even with multiple units to cover the entire Delta region covered by HyMap. Table 9 shows data associated with the conducted HyMap and Nano flights, and estimates of time, data volume, and deployment costs to fly the entire Delta with the Nano. For this estimate, the high overlap used in this study to acquire imagery from two directions was abandoned, effectively halving the time estimates, data volume, and associated deployment costs. Even with this reduction, using the Nano to cover the entire Delta would likely be cost prohibitive. Time estimates include 4 hours of collection time per day corresponding to solar windows, plus launch preparation time, downloading data, and travel amounting to roughly the other 4 working hours per day. Approximate data volume estimates for HyMap only consist of orthorectified reflectance data, where Nano data volume includes raw image cubes, orthorectified reflectance data, and the intermediate steps. Approximate deployment cost includes initial purchase cost for the Headwall Nano-Hyperspec turnkey package, assumes a team of 3 members, pilot, designated observer, and calibration technician to operate a portable field spectrometer working 8 hour days at $20 per hour, a rental vehicle

for $60/day, a 60% overhead to cover additional travel expenses such as hotel rooms, and data storage priced at $0.02 per GB. With this design, a survey of the entire Delta would take more than a year to complete with one team, negating much of the advantages of using unmanned aircraft. Multiple teams with multiple units may be able to complete a survey of the entire Delta in a shorter time frame but would have additional costs associated with personnel operating on multiple teams and capital expenditures on multiple pieces of equipment. An additional concern for large-scale usage of the unmanned aircraft for mapping would be data volume. To cover the entire data following the protocols outlined here, roughly 2100 TB of storage space at an estimated cost of $42,238 would be necessary to collect and convert the data to a orthorectified reflectance product, not including the storage necessary to create remote sensing products or construct a map. One potential way to reduce data volume could be by using a multispectral camera, but a different methodology that did not utilize narrow band indices would be required. The small geographic footprint associated with the high spatial and spectral resolution capabilities of unmanned aircraft imaging spectrometers make them difficult to implement as a detection method over such a large area. For this reason, if unmanned aircraft mounted imaging spectrometers were to be utilized in invasive species management applications, targeted flights in a low quantity of locations of high concern would be recommended.

5. CONCLUSIONS

This study outlines a procedure for determining an optimal model and mapping invasive plants in a wetland ecosystem using a Headwall Nano-Hyperspec, and shows the higher spatial resolution provided by the Nano can improve detection of small patches of invasives relative to an operational government funded program. Additionally, we show that a similar methodology using a RF classification with data acquired from an unmanned aircraft mounted Nano-Hyperspec can result in similar accuracies to that of the same operational procedure that is conducted annually. This improved fine scale detection does come at a cost, the operational footprint. The small footprint makes coverage of the entire Delta with unmanned aircraft highly impractical if not impossible, restricting unmanned aircraft applications to specific areas of concern with the specified procedure.

There were several challenges to overcome, including the inability to resolve the spatial misalignment between Nano and HyMap imagery, which made location specific comparisons difficult. Though great considerations were taken in attempts to eliminate these issues up front, such as use of high precision GPS equipment and several attempts at orthorectification of the unmanned aircraft data, they seem inevitable working at a scale of 5cm pixels. Another challenge was data volume. The project require 8.6 TB of data storage due to the several remote sensing products used in the project, amounting to the 1702 variables originally used in the model, which proved to be too large for efficient processing, which was resolved by reducing the quantity of variables to only those of high importance.

In the future, this study may be used as a template for invasive species mapping using a Nano-Hyperspec. The capability to detect small patches as shown here is highly useful for weed management applications, as well as mapping that requires fine resolution. Improvement of the orthorectification procedure for the Nano through automation, would

greatly simplify this mapping process and make utilizing the Nano for mapping more practical. Though there are still hurdles to overcome, technological advances will make this application more practical in the near future.

REFERENCES

[1] E. A. Bolch *et al.*, "Remote Detection of Invasive Alien Species," in *Remote Sensing of Plant Biodiversity*, J. Cavender-Bares, J. Gamon, and P. Townsend, Eds. Cham, Switzerland: Springer Nature, 2020, pp. 267–308.

[2] D. Pimentel, R. Zuniga, and D. Morrison, "Update on the environmental and economic costs associated with alien-invasive species in the United States," *Ecol. Econ.*, vol. 52, no. 3, pp. 273–288, Feb. 2005.

[3] J. E. Byers and E. G. Noonburg, "Scale dependent effects of biotic resistance to biological invasion," *Ecology*, vol. 84, no. 6, pp. 1428–1433, Jun. 2003.

[4] S. Khanna, M. J. Santos, J. D. Boyer, K. D. Shapiro, J. Bellvert, and S. L. Ustin, "Water primrose invasion changes successional pathways in an estuarine ecosystem," *Ecosphere*, vol. 9, no. 9, p. e02418, Sep. 2018.

[5] J. M. Allen and B. A. Bradley, "Out of the weeds? Reduced plant invasion risk with climate change in the continental United States," *Biol. Conserv.*, vol. 203, pp. 306–312, Nov. 2016.

[6] A. Ricciardi, "Are modern biological invasions an unprecedented form of global change?," *Conservation Biology*. 2007.

[7] H. Seebens *et al.*, "No saturation in the accumulation of alien species worldwide," *Nat. Commun.*, vol. 8, p. 14435, Feb. 2017.

[8] D. A. Mortensen, E. S. J. Rauschert, A. N. Nord, and B. P. Jones, "Forest Roads Facilitate the Spread of Invasive Plants," *Invasive Plant Sci. Manag.*, vol. 2, no. 03, pp. 191–199, Jul. 2009.

[9] G. Masters and L. Norgrove, "KNOWLEDGE FOR LIFE Climate Change and Invasive Alien Species," 2010.

[10] P. E. Hulme, "Invasion pathways at a crossroad: Policy and research challenges for managing alien species introductions," *J. Appl. Ecol.*, vol. 52, no. 6, pp. 1418–1424, Dec. 2015.

[11] UN General Assembly, "No Title," in *Transforming our world : the 2030 Agenda for Sustainable Development*, 2015, p. A/RES/70/1.

[12] E. L. Hestir *et al.*, "Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem," *Remote Sens. Environ.*, vol. 112, no. 11, pp. 4034–4047, 2008.

[13] M. Y. Jollineau and P. J. Howarth, "Mapping an inland wetland complex using hyperspectral imagery," *Int. J. Remote Sens.*, vol. 29, no. 12, pp. 3609–3631, Jun. 2008.

[14] P. D. Hunter, D. J. Gilvear, A. N. Tyler, N. J. Willby, and A. Kelly, "Mapping macrophytic vegetation in shallow lakes using the Compact Airborne Spectrographic Imager (CASI)," *Aquat. Conserv. Mar. Freshw. Ecosyst.*, vol. 20, no. 7, pp. 717–727, Nov. 2010.

[15] S. Khanna, M. J. Santos, S. L. Ustin, and P. J. Haverkamp, "International Journal of Remote Sensing An integrated approach to a biophysiologically based classification of floating aquatic macrophytes An integrated approach to a

biophysiologically based classification of floating aquatic macrophytes," *Int. J. Remote Sens.*, vol. 32, no. 4, pp. 1067–1094, 2011.
[16] E. L. Hestir, J. A. Greenberg, and S. L. Ustin, "Classification trees for aquatic vegetation community prediction using imaging spectroscopy," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 5, no. 5, pp. 1572–1584, Oct. 2012.
[17] D. Zhao, H. Jiang, T. Yang, Y. Cai, D. Xu, and S. An, "Remote sensing of aquatic vegetation distribution in Taihu Lake using an improved classification tree with modified thresholds," *J. Environ. Manage.*, vol. 95, no. 1, pp. 98–107, Mar. 2012.
[18] M. J. Santos *et al.*, "Use of Hyperspectral Remote Sensing to Evaluate Efficacy of Aquatic Plant Management," *Invasive Plant Sci. Manag.*, vol. 2, no. 03, pp. 216–229, 2009.
[19] C. L. Doughty and K. C. Cavanaugh, "Mapping coastal wetland biomass from high resolution unmanned aerial vehicle (UAV) imagery," *Remote Sens.*, vol. 11, no. 5, 2019.
[20] Y. Zhong *et al.*, "Mini-UAV-Borne Hyperspectral Remote Sensing: From Observation and Processing to Applications," *IEEE Geosci. Remote Sens. Mag.*, vol. 6, no. 4, pp. 46–62, 2018.
[21] D. J. Turner, Z. Malenovsky, A. Lucieer, J. D. Turnbull, and S. A. Robinson, "Optimizing Spectral and Spatial Resolutions of Unmanned Aerial System Imaging Sensors for Monitoring Antarctic Vegetation," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 10, pp. 3813–3825, Oct. 2019.
[22] B. P. Banerjee, S. Raval, and P. J. Cullen, "UAV-hyperspectral imaging of spectrally complex environments," *Int. J. Remote Sens.*, vol. 41, no. 11, pp. 4136–4159, Jun. 2020.
[23] B. Melville, A. Lucieer, and J. Aryal, "Classification of Lowland Native Grassland Communities Using Hyperspectral Unmanned Aircraft System (UAS) Imagery in the Tasmanian Midlands," *Drones*, vol. 3, no. 1, p. 5, Jan. 2019.
[24] S. L. Ustin, S. Khanna, M. Lay, K. Shapiro, and N. Ghajarnia, "Remote sensing of the Sacramento-San Joaquin Delta to enhance mapping for invasive and native aquatic vegetation plant species," 2020.
[25] E. Tabacchi, D. L. Correll, R. Hauer, G. Pinay, A. M. Planty-Tabacchi, and R. C. Wissmar, "Development, maintenance and role of riparian vegetation in the river landscape," *Freshw. Biol.*, vol. 40, no. 3, pp. 497–516, 1998.
[26] E. B. Barbier, S. D. Hacker, C. Kennedy, E. W. Koch, A. C. Stier, and B. R. Silliman, "The value of estuarine and coastal ecosystem services," *Ecol. Monogr.*, vol. 81, no. 2, pp. 169–193, May 2011.
[27] California Department of Food and Agriculture, "California Agricultural Statistics Review," Sacramento, 2019.
[28] E. C. Underwood, M. J. Mulitsch, J. A. Greenberg, M. L. Whiting, S. L. Ustin, and S. C. Kefauver, "Mapping invasive aquatic vegetation in the sacramento-san Joaquin Delta using hyperspectral imagery," *Environ. Monit. Assess.*, vol. 121, no. 1–3, pp. 47–64, Oct. 2006.
[29] W. Kimmerer *et al.*, "Effects of drought and the emergency drought barrier on the ecosystem of the California Delta," *San Fr. Estuary Watershed Sci.*, vol. 17, no. 3, 2019.

[30] S. Khanna, M. J. Santos, E. L. Hestir, and S. L. Ustin, "Plant community dynamics relative to the changing distribution of a highly invasive species, Eichhornia crassipes: A remote sensing perspective," *Biol. Invasions*, vol. 14, no. 3, pp. 717–733, 2012.
[31] M. J. Santos, E. L. Hestir, S. Khanna, and S. L. Ustin, "Image spectroscopy and stable isotopes elucidate functional dissimilarity between native and nonnative plant species in the aquatic environment," *New Phytol.*, vol. 193, no. 3, pp. 683–695, Nov. 2012.
[32] E. L. Hestir, D. H. Schoellhamer, J. Greenberg, T. Morgan-King, and S. L. Ustin, "The Effect of Submerged Aquatic Vegetation Expansion on a Declining Turbidity Trend in the Sacramento-San Joaquin River Delta," *Estuaries and Coasts*, vol. 39, no. 4, pp. 1100–1112, 2016.
[33] A. N. Cohen and J. T. Carlton, "Nonindigenous Aquatic Species in a United States Estuary: A Case Study of the Biological Invasions of the San Francisco Bay and Delta," 1995.
[34] G. Venugopal, "Monitoring the Effects of Biological Control of Water Hyacinths Using Remotely Sensed Data: A Case Study of Bangalore, India," *Singap. J. Trop. Geogr.*, vol. 19, no. 1, pp. 91–105, Dec. 2002.
[35] K. M. Jetter and K. Nes, "The Cost to Manage Invasive Aquatic Weeds in the California Bay-Delta The Division of Boating and Waterways (DBW)," 2015.
[36] J. D. Toft, C. A. Simenstad, J. R. Cordell, and L. F. Grimaldo, "The Effects of Introduced Water Hyacinth on Habitat Structure, Invertebrate Assemblages, and Fish Diets," 2003.
[37] S. L. Ustin, S. Khanna, M. Lay, and K. D. Shapiro, "Enhancement of Delta Smelt (Hypomesus transpacificus) habitat through adaptive management of invasive aquatic weeds in the Sacramento-San Joaquin Delta & Suisun," 2019.
[38] D. J. Jensen, M. Simard, K. C. Cavanaugh, and D. R. Thompson, "Imaging Spectroscopy BRDF Correction for Mapping Louisiana's Coastal Ecosystems," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1739–1748, Mar. 2018.
[39] SPH Engineering, "UgCS." 2018.
[40] Headwall Photonics, "SpectralView." Fitchburg, MA, 2017.
[41] Exelis Visual Information Solutions, "ENVI." Boulder, CO.
[42] S. L. Ustin, S. Khanna, M. Lay, and K. D. Shapiro, "Enhancement of Delta Smelt (Hypomesus transpacificus) habitat through adaptive management of invasive aquatic weeds in the Sacramento-San Joaquin Delta," 2018.
[43] Trimble, "GPS Pathfinder Office." Westminster, CO, 2018.
[44] A. Mellor, S. Boukir, A. Haywood, and S. Jones, "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 155–168, Jul. 2015.
[45] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
[46] J. W. Coulston, C. E. Blinn, V. A. Thomas, and R. H. Wynne, "Approximating prediction uncertainty for random forest regression models," *Photogramm. Eng. Remote Sensing*, vol. 82, no. 3, pp. 189–197, 2016.
[47] L. Loosvelt *et al.*, "Random Forests as a tool for estimating uncertainty at pixel-

level in SAR image classification," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 19, no. 1, pp. 173–184, 2012.
[48] P. Sinha, A. E. Gaughan, F. R. Stevens, J. J. Nieves, A. Sorichetta, and A. J. Tatem, "Assessing the spatial sensitivity of a random forest model: Application in gridded population modeling," *Comput. Environ. Urban Syst.*, vol. 75, pp. 132–145, May 2019.
[49] M. Story and R. G. Congalton, "Accuracy Assessment: A User's Perspective," *Photogramm. Eng. Remote Sensing*, vol. 52, pp. 397–399, 1986.
[50] G. H. Rosenfield and K. Fitzpatrick-Lins, "A coefficient of agreement as a measure of thematic classifcation accuracy.," *Photogramm. Eng. Remote Sensing*, vol. 52, no. 2, pp. 223–227, 1986.
[51] T. M. Lillesand, R. M. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*, 5th ed. New York, NY: Wiley, 2004.
[52] G. M. Foody, "Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification," *Remote Sens. Environ.*, vol. 239, p. 111630, Mar. 2020.
[53] R. G. Pontius and M. Millones, "Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment," *International Journal of Remote Sensing*, vol. 32, no. 15. Taylor and Francis Ltd., pp. 4407–4429, 2011.
[54] M. Kuhn, "Building Predictive Models in R Using the caret Package," *J. Stat. Software, Artic.*, vol. 28, no. 5, pp. 1–26, 2008.
[55] A. Liaw and M. Wiener, "Classification and Regression by randomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
[56] R Core Team, "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing, Vienna, Austria, 2019.
[57] A. Green, "The use of multivariate statistical techniques for the analysis and display of AEM data," *Explor. Geophys.*, vol. 29, no. 2, pp. 77–82, 1998.
[58] G. Luo, G. Chen, L. Tian, K. Qin, and S.-E. Qian, "Minimum Noise Fraction versus Principal Component Analysis as a Preprocessing Step for Hyperspectral Imagery Denoising," *Can. J. Remote Sens.*, vol. 42, no. 2, pp. 106–116, Mar. 2016.
[59] L. W. Lenhert *et al.*, "Hyperspectral Data Analysis in R: The hsdar Package," *J. Stat. Softw.*, vol. 89, no. 12, pp. 1–23, 2019.
[60] T. G. Whiteside, G. S. Boggs, and S. W. Maier, "Comparing object-based and pixel-based classifications for mapping savannas," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 13, no. 6, pp. 884–893, 2011.
[61] Y. Gao and J. F. Mas, "A Comparison of the Performance of Pixel Based and Object Based Classifications over Images with Various Spatial Resolutions," *Online J. Earth Sci.*, vol. 2, no. 1, pp. 27–35, 2008.
[62] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sens. Environ.*, vol. 115, no. 5, pp. 1145–1161, May 2011.
[63] J. Gonçalves, I. Pôças, B. Marcos, C. A. Mücher, and J. P. Honrado, "SegOptim-A new R package for optimizing object-based image analyses of high-spatial resolution remotely-sensed data," 2018.

[64] M. Grizonnet, J. Michel, V. Poughon, J. Inglada, M. Savinaud, and R. Cresson, "Orfeo ToolBox: open source processing of remote sensing images," *Open Geospatial Data, Softw. Stand.*, vol. 2, no. 1, p. 15, Dec. 2017.
[65] J. Radoux and P. Defourny, "A quantitative assessment of boundaries in automated forest stand delineation using very high resolution imagery," *Remote Sens. Environ.*, vol. 110, no. 4, pp. 468–475, Oct. 2007.
[66] E. W. Fox, R. A. Hill, S. G. Leibowitz, A. R. Olsen, D. J. Thornbrugh, and M. H. Weber, "Assessing the accuracy and stability of variable selection methods for random forest modeling in ecology," *Environ. Monit. Assess.*, vol. 189, no. 7, p. 316, Jul. 2017.
[67] K. Millard and M. Richardson, "On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping," *Remote Sens.*, vol. 7, no. 7, pp. 8489–8515, Jul. 2015.
[68] J. Breidenbach, E. Næsset, V. Lien, T. Gobakken, and S. Solberg, "Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data," *Remote Sens. Environ.*, vol. 114, no. 4, pp. 911–924, Apr. 2010.
[69] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts, "Understanding variable importances in Forests of randomized trees," in *Advances in Neural Information Processing Systems*, 2013.
[70] H. Han, X. Guo, and H. Yu, "Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest," in *Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS*, 2016, vol. 0, pp. 219–224.
[71] J. Hastie, T., Tibshirani, R., & Friedman, "Random forests," in *The elements of statistical learning*, New York, NY: Springer, 2009, pp. 587–604.
[72] D. W. Rouse Jr, J. W., Haas, R. H., Schell, J. A., & Deering, "Monitoring Vegetation Systems in the Great Plains with ERTS," in *Third Earth Resources Technology Satellite-1 Symposium: The Proceedings of a Symposium Held by Goddard Space Flight Center at Washington, DC on December 10-14, 1973*, 1974, p. 309.
[73] K. Pearson, " X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling ," *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 50, no. 302, pp. 157–175, Jul. 1900.
[74] H. Cramer, *Mathematical methods of statistics*. Princeton: Princeton University Press, 1946.
[75] L. Hubert and P. Arabie, "Comparing partitions," *J. Classif.*, vol. 2, no. 1, pp. 193–218, Dec. 1985.
[76] E. L. Hestir *et al.*, "Measuring freshwater aquatic ecosystems: The need for a hyperspectral global mapping satellite mission," *Remote Sens. Environ.*, vol. 167, pp. 181–195, 2015.
[77] S. L. Ustin, D. A. Roberts, J. A. Gamon, G. P. Asner, and R. O. Green, "Using imaging spectroscopy to study ecosystem processes and properties," *BioScience*, vol. 54, no. 6. American Institute of Biological Sciences, pp. 523–534, 01-Jun-

2004.
[78] J. R. Jensen, “Biophysical Remote Sensing,” *Ann. Assoc. Am. Geogr.*, vol. 73, no. 1, pp. 111–132, 1983.
[79] E. Husson, F. Ecke, and H. Reese, “Comparison of Manual Mapping and Automated Object-Based Image Analysis of Non-Submerged Aquatic Vegetation from Very-High-Resolution UAS Images,” *Remote Sens.*, vol. 8, no. 9, p. 724, Sep. 2016.
[80] S. Kaewpijit, J. Le Moigne, and T. El-Ghazawi, “Automatic reduction of hyperspectral imagery using wavelet spectral analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 4 PART I, pp. 863–871, Apr. 2003.
[81] A. Agarwal, T. El-Ghazawi, H. El-Askary, and J. Le-Moigne, “Efficient hierarchical-PCA dimension reduction for hyperspectral imagery,” in *ISSPIT 2007 - 2007 IEEE International Symposium on Signal Processing and Information Technology*, 2007, pp. 353–356.
[82] K. Millard and M. Richardson, “Wetland mapping with LiDAR derivatives, SAR polarimetric decompositions, and LiDAR–SAR fusion using a random forest classifier,” *Can. J. Remote Sens.*, vol. 39, no. 4, pp. 290–307, Oct. 2013.
[83] A. Sabat-Tomala, E. Raczko, and B. Zagajewski, “Comparison of Support Vector Machine and Random Forest Algorithms for Invasive and Expansive Species Classification Using Airborne Hyperspectral Data,” *Remote Sens.*, vol. 12, no. 3, p. 516, Feb. 2020.
[84] M. Colgan, C. Baldeck, J.-B. Féret, and G. Asner, “Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data,” *Remote Sens.*, vol. 4, no. 11, pp. 3462–3480, Nov. 2012.
[85] J. S. Myers and R. L. Miller, “Optical Airborne Remote Sensing,” Springer, Dordrecht, 2007, pp. 51–67.
[86] Y.-H. Tu, S. Phinn, K. Johansen, and A. Robson, “Assessing Radiometric Correction Approaches for Multi-Spectral UAS Imagery for Horticultural Applications,” *Remote Sens.*, vol. 10, no. 11, p. 1684, Oct. 2018.
[87] D. F. Spencer and G. G. Ksander, “Seasonal growth of waterhyacinth in the Sacramento/San Joaquin Delta, California,” *J. Aquat. Plant Manag.*, vol. 43, no. JUL., pp. 91–94, 2005.
[88] P. Evangelista, T. Stohlgren, J. Morisette, and S. Kumar, “Mapping Invasive Tamarisk (Tamarix): A Comparison of Single-Scene and Time-Series Analyses of Remotely Sensed Data,” *Remote Sens.*, vol. 1, no. 3, pp. 519–533, Aug. 2009.
[89] M. E. Andrew and S. L. Ustin, “The role of environmental context in mapping invasive plants with hyperspectral image data,” *Remote Sens. Environ.*, vol. 112, no. 12, pp. 4301–4317, 2008.
[90] B. Somers and G. P. Asner, “Invasive species mapping in hawaiian rainforests using multi-temporal hyperion spaceborne imaging spectroscopy,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 6, no. 2, pp. 351–359, 2013.
[91] R. G. Congalton and K. Green, *Assessing the Accuracy of Remotely Sensed Data*. Boca Raton, FL, USA: CRC Press, 2019.
[92] H. S. He, S. J. Ventura, and D. J. Mladenoff, “Effects of spatial aggregation

approaches on classified satellite imagery," *Int. J. Geogr. Inf. Sci.*, vol. 16, no. 1, pp. 93–109, 2002.
[93] J. H. McDonald, *Handbook of Biological Statistics*, 3rd ed. Baltimore, MD: Sparky House Publishing, 2014.

APPENDIX

A. Table of Vegetation Indices

Abbreviation	Name	Relevance
Boochs	Boochs/SB703	Chlorophyll/Red Edge
Boochs2	Boochs2/SB720	Chlorophyll/Red Edge
CARI	Chlorophyll Absorption Ratio Index	Chlorophyll/Red Edge
Carter	Carter	Chlorophyll/Red Edge
Carter2	Carter2	Chlorophyll/Red Edge
Carter3	Carter3	Chlorophyll/Red Edge
Carter4	Carter4	Chlorophyll/Red Edge
Carter5	Carter5	Chlorophyll/Red Edge
Carter6	Carter6	Chlorophyll/Red Edge
CI	Coloration Index	Other
CI2	Coloration Index 2	Other
CRI1	Carotenoid Reflectance Index 1	Other
CRI2	Carotenoid Reflectance Index 2	Other
CRI3	Carotenoid Reflectance Index 3	Other
CRI4	Carotenoid Reflectance Index 4	Other
D1	Derivative Index 1	Chlorophyll/Red Edge
D2	Derivative Index 2	Chlorophyll/Red Edge
Datt	Datt	Chlorophyll/Red Edge
Datt2	Datt2	Chlorophyll/Red Edge
Datt3	Datt3	Chlorophyll/Red Edge
Datt4	Datt4	Chlorophyll/Red Edge
Datt5	Datt5	Chlorophyll/Red Edge
Datt6	Datt6	Chlorophyll/Red Edge
DD	Double Difference Index	Chlorophyll/Red Edge
DDn	New Double Difference Index	Chlorophyll/Red Edge

DPI	Double Peak Index	Chlorophyll/Red Edge
DWSI4	Disease Water Stress Index 4	Other
EVI	Enhanved Vegetation Index	Chlorophyll/Red Edge
GI	Greenness Index	Chlorophyll/Red Edge
Gitelson	Gitelson	Chlorophyll/Red Edge
Gitelson2	Gitelson2	Chlorophyll/Red Edge
GMI1	Gitelson and Merzlyak Index 1	Chlorophyll/Red Edge
GMI2	Gitelson and Merzlyak Index 2	Chlorophyll/Red Edge
Green NDVI	Green Normalized Difference Vegetation Index	Chlorophyll/Red Edge
Maccioni	Maccioni	Chlorophyll/Red Edge
MCARI	Modified Chlorophyll Absorption in Reflectance Index	Chlorophyll/Red Edge
MCARI2	Modified Chlorophyll Absorption in Reflectance Index 2	Chlorophyll/Red Edge
mND705	Modified Normalized Difference Vegetation Index 705	Chlorophyll/Red Edge
mNDVI	Modified Normalized Difference Vegetation Index	Chlorophyll/Red Edge
MPRI	Modified Photochemical Reflectance Index	Other
MSAVI	Modified Soil Adjusted Vegetation Index	Chlorophyll/Red Edge
mSR	Modified Simple Ratio	Chlorophyll/Red Edge
mSR2	Modified Simple Ratio 2	Chlorophyll/Red Edge
mSR705	Modified Simple Ratio 705	Chlorophyll/Red Edge
MTCI	MERIS Terrestrial Chlorophyll Index	Chlorophyll/Red Edge
MTVI	Modified Triangular Vegetation Index	Structural
NDVI	Normalized Difference Vegetation Index	Structural
NDVI2	Normalized Difference Vegetation Index 2	Chlorophyll/Red Edge
NDVI3	Normalized Difference Vegetation Index 3	Chlorophyll/Red Edge
NPCI	Normalized Pigment Chlorophyll Index	Chlorophyll/Red Edge
OSAVI	Optimized Soil Adjusted Vegetation Index	Structural
OSAVI2	Optimized Soil Adjusted Vegetation Index 2	Chlorophyll/Red Edge
PARS	Published Analysis of Reflectance Spectra	Other
PRI	Photochemical Reflective Index	Other
PSND	Pigment Specific Normalized Difference	Other

PSRI	Plant Sensecence Reflectance Index	Other
PSSR	Pigment Specific Simple Ratio	Other
PWI	Plant Water Index	Water
RDVI	Renormalized Difference Vegetation Index	Structural
RDVI	Renormalized Difference Vegetation Index	Chlorophyll/Red Edge
SAVI	Soil Adjusted Vegetation Index	Structural
SPVI	Spectral Polygon Vegetation Index	Other
SR	Simple Ratio	Chlorophyll/Red Edge
SR1	Simple Ratio 1	Chlorophyll/Red Edge
SR2	Simple Ratio 2	Chlorophyll/Red Edge
SR3	Simple Ratio 3	Chlorophyll/Red Edge
SR4	Simple Ratio 4	Chlorophyll/Red Edge
SR5	Simple Ratio 5	Chlorophyll/Red Edge
SR6	Simple Ratio 6	Chlorophyll/Red Edge
SR7	Simple Ratio 7	Chlorophyll/Red Edge
SR8	Simple Ratio 8	Chlorophyll/Red Edge
SRPI	Simple Ratio Pigment Index	Chlorophyll/Red Edge
TCARI	Transformed Cholorophyll Absorbtion Ration	Chlorophyll/Red Edge
TCARI2	Transformed Cholorophyll Absorbtion Ration 2	Chlorophyll/Red Edge
TGI	Triangular Greenness Index	Other
TVI	Transformed Vegetation Index	Chlorophyll/Red Edge
Vogelmann	Vogelmann	Chlorophyll/Red Edge
Vogelmann2	Vogelmann2	Chlorophyll/Red Edge
Vogelmann3	Vogelmann3	Chlorophyll/Red Edge
Vogelmann4	Vogelmann4	Chlorophyll/Red Edge
WBI	Water Band Index	Water

Note: All indices formulas can be found in the hsdar packages for R.

B. Map Comparison Contingency Tables

Comparison 3: MW Resampled Nano v. HyMap

	Unclassified	Bare Ground	EMR	Water Hyacinth	Water Primrose	Riparian Shrub	Riparian Tree	SAV	Water	Other	NPV	Sum
Unclassified	97	2	1	0	4	0	0	1	0	5	1	111
Bare ground	0	37	1	0	23	3	0	0	1	26	1	92
EMR	0	2	39	0	8	21	15	1	1	2	12	101
Water Hyacinth	0	0	0	0	0	0	0	0	0	0	0	0
Water Primrose	0	0	0	0	0	0	0	0	0	0	0	0
Riparian Shrub	0	14	0	0	7	5	1	0	0	57	0	84
Riparian Tree	0	7	8	0	40	51	77	0	2	8	2	195
SAV	1	16	6	0	3	4	0	54	5	0	4	93
Water	0	7	3	0	12	5	0	44	90	0	2	163
Other	0	0	0	0	0	0	0	0	0	0	0	0
NPV	1	15	42	0	3	11	7	0	1	2	78	160
Sum	99	100	100	0	100	100	100	100	100	100	100	999

Comparison 4: Nearest Neighbor Resampled Nano v. HyMap

	Unclassified	Bare Ground	EMR	Water Hyacinth	Water Primrose	Riparian Shrub	Riparian Tree	SAV	Water	Other	NPV	Sum
Unclassified	98	0	0	0	0	0	0	0	0	0	0	98
Bare ground	0	35	1	0	25	3	0	0	1	27	1	93
EMR	0	3	43	0	10	23	15	1	2	5	11	113
Water Hyacinth	0	0	0	0	0	0	0	0	0	0	0	0
Water Primrose	0	0	0	0	0	0	0	0	0	1	0	1
Riparian Shrub	0	17	0	0	10	4	1	0	0	56	0	88
Riparian Tree	0	9	8	0	38	51	78	0	2	10	3	199
SAV	1	19	7	0	4	2	0	54	8	0	4	99
Water	0	5	2	0	11	7	0	44	86	0	2	157
Other	0	0	0	0	0	0	0	0	0	0	0	0
NPV	1	12	39	0	2	10	6	1	1	1	79	152
Sum	100	100	100	0	100	100	100	100	100	100	100	1000