A LOOK AT LAND COVER CLASSIFICATION METHODS IN NORTHERN CALIFORNIA WITH THE USE OF HIGH SPATIAL RESOLUTION GEOSPATIAL DATA

By

Lucila Corro

A Thesis Presented to

The Faculty of Humboldt State University

In Partial Fulfillment of the Requirements for the Degree

Master of Science in Natural Resources: Environmental Science and Management

Committee Membership

Dr. James Graham, Committee Chair

Dr. Frank Shaughnessy, Committee Member

Dr. David Gwenzi, Committee Member

Dr. Erin Kelly, Graduate Coordinator

May 2021

ABSTRACT

A LOOK AT LAND COVER CLASSIFICATION METHODS IN NORTHERN CALIFORNIA WITH THE USE OF HIGH SPATIAL RESOLUTION GEOSPATIAL DATA

Lucila Corro

Land use and land cover (LULC) mapping plays a vital role in understanding the state of the world, showing us a visual representation of the natural and anthropogenic features covering our planet. Northern California in the United States is home to many critical habitats that provide for a variety of endemic and some threatened and engendered species, making it an area of particular concern to better understand and monitor. There is a greater need to identify specific methods for vegetation modeling in Northern California due to its unique species; to do this we examined two case studies with the following objectives: 1) Determine whether unmanned aerial system (UAS) image analysis can provide similar estimates of eelgrass biometrics, such as percent coverage, to those obtained in situ using traditional field survey methods; 2) To develop a GIS data fusion workflow for high-resolution habitat classification in the Napa Watershed of central California with a focus on oak savanna habitat. UAS Imagery for two eelgrass sites were collected during June, 2019 using a DJI Matrice 100 equipped with MicaSense RedEdge Multispectral sensor (5-band). Following UAS image collection, ground survey data were collected at three tidal elevation transects per site,

with 20 quadrats stationed randomly along each transect. Eelgrass percent coverage was measured for each quadrat and then compared to eelgrass classification models derived from UAS derived imagery. In the Napa watershed, we examined methods necessary to accurately incorporate ancillary geospatial spatial datasets into a remote sensing land cover classification. By doing so, I developed a habitat distribution dataset that may better analyze interactions of wildlife, humans, and the endemic habitat types of the Napa watershed in California. UAVs provided a means to obtain high resolution remote sensing imagery of eelgrass at a resolution of 3.46 – 3.70 cm per pixel or greater at specific tidal periods, providing a useful methodology that allowed for percent coverage estimates with an R^2 value of 0.6496 compared to *in situ* measurements. While developing a land cover classification workflow for the Napa watershed, I found that by incorporating ancillary geospatial data, remotely sensed data, and threshold classification, I could obtain a LULC model that more accurately depicts the endemic land use and land cover features of the Napa watershed. With an overall accuracy of 70.20% and a kappa statistic of 0.6140, this modeling method proved more accurate than traditional image classification methods. With ground sampled reference data and remotely sensed data gathered at the same temporal and spatial scales these classification methods would be robust and replicable for future analyses.

ACKNOWLEDGEMENTS

I would like to thank all my family and friends that helped me realize my goal of obtaining a graduate degree in environmental science and management. I want to thank my major advisor Dr. Jim Graham for providing guidance, assistance, and understanding while working at Humboldt State University. I would also like to thank my thesis committee for all their time and commitment in aiding me to complete my thesis. I would like to thank James Lamping, Shayne Magstadt, Marina Gagarina, Kevin Landaw and Dr. David Gwenzi for assisting in UAV and field data collection. I would also like to thank Humboldt State University's Institute for Spatial Analysis, Monitoring and Modeling for providing UAV equipment and software for image processing and analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

INTRODUCTION

Land use and land cover (LULC) classification is a method of remote sensing that separates land cover and land use based on the surface reflectance of the earth or other geospatial data. Land cover describes what physical material covers the earth or what is seen on land from above. Land use differs from land cover in that it focuses on how the land is being utilized. LULC plays a vital role in understanding the state of the world, showing us a visual representation of the natural and anthropogenic features covering our planet (Tesfaw et al., 2018). Land cover models have been developed for many years to classify landscapes into descriptive informational groups based on the land cover characteristics seen on land from above (Lillesand et al., 2008). Land cover maps can also be used to describe the different land uses, land cover patterns and habitat types. These types of maps are also used to assess land cover changes (Green et al., 1994). Land cover maps that are classified based on habitat needs have been used by biologist to better understand patterns of wildlife distributions (Shi et al., 2006), reproduction (Brown et al., 2017), and behavior (Hargrove et al., 2005).

Remote sensing is a scientific method that aims to gather information about a subject without encountering that object and has been a vital resource to land managers for many years (Lillesand, Kiefer, & Chipman, 2008). With the use of remote sensing data, researchers can observe trends across the landscape more efficiently (Hardin, 1999). Spatial modeling is often used in remote sensing to classify features or objects of interest from imagery (Lillesand et al., 2008). The use of spatial modeling as a tool in natural

resource management has proliferated in recent years as computing power has become more affordable and therefore more accessible to researchers and land managers (Digruttolo & Mohamed, 2010). Now, remote sensing data collection is becoming even more accessible with the advent of Unmanned Aerial Vehicles (UAVs).

To better understand the natural ecosystems around us, it is important to have representative data of natural phenomenon (John R Jensen, 2007; Plummer, 2000). Gathering these data is often a costly and time-consuming process that requires large on the ground campaigns to gather accurate and standardized data (Buiten & Clevers, 1993). However, with the advancement of aero-nautical science and imaging sensor technology, aerial images of the Earth are readily available and have emerged as a more cost-effective mapping tool (Lillesand et al., 2008). Remote sensing data, such as aerial imagery, is collected without encountering the object of interest. This means that remote sensing data is often less invasive, less costly and more accurate (Congalton & Green, 1999). Using remotely sensed data, such as aerial imagery, scientists have gained a unique perspective to landscape patterns and characteristics allowing for a greater understanding of the natural world around us (Suribabu et al., 2012).

While land cover maps are some of the most widely available data, they are limited in information to discrete nominal classes (Hollister et al., 2004). These data are also often coarse in resolution (30 m or greater) and consist of generalized classifications not specific to regional habitat types (J. Robinson et al., 1994). It is also true that with coarser resolutions of 30 meters per pixel or greater, most of these land cover data are also contain more inherit uncertainty (Cracknell, 1998). These land cover data are often

classified into broad land cover classes that may not address the finer spatial scale of research needs (Hollister et al., 2004).

Though it is clear there is a need to map land use and land cover, it can often be difficult to identify the best methods for the desired final product (Moody & Woodcock, 1995). Most LULC mapping is done using remotely sensed imagery, by separating land cover classes based on surface reflectance. Aerial mapping methods began as early as the 1860s with maps drawn using some of the first aerial images (Graham & Read, 1986). While this emerged as a useful tool, it wasn't until the 1990s when satellite imagery became a more available and accurate resource for LULC mapping (Wulder et al., 2012). With higher temporal, spectral, and spatial resolution in remotely sensed imagery, remote sensing experts developed more complex, detailed and often more accurate land cover maps.

Higher spatial resolution imagery has proven useful in land cover classifications as it allows for more detailed mapping and more fine-scale delineation of features (West, 2007). Higher resolution imagery with relatively low spectral resolution has been used to classify urban, water, and vegetation land cover features (Perumal & Bhaskaran, 2010). Imagery with high spatial resolution (0.1-1.0m) can be very useful in delineating features that exhibit a fine-scale pattern on the landscape. For instance, urban land cover can be difficult to identify with moderate resolution (10.0m-30.0m) imagery (Hu et al., 2016). Using images with high spatial resolution can improve the final resolution of land cover mapping and can improve accuracy. While high spatial resolution imagery, such as NAIP (USDA National Agriculture Imagery Program), can be used to delineate three feature

classes (urban, water, and vegetation), with limited spectral information, it is difficult to further classify different types of vegetation (Srivastava et al., 2012).

For instance, it is often necessary to include supplementary data either collected on the ground with geographic positioning systems or using remote sensing (Vogelmann et al., 1998). For example, lidar height data can be used to supplement a spectrally-limited remote sensing image classification (Pham et al., 2016). The addition of lidar height data can help to further distinguish different types of vegetation. In addition to lidar data, the Napa County GIS Department developed a ground-validated spatial dataset that accurately mapped agricultural land cover at a high resolution (Matt Lamborn, 2010). Using both lidar height data and vector based agricultural coverage and high-resolution hydrography data, it is possible that a more detailed and comprehensive model can be produced with higher accuracy for the Napa Valley.

Northern California is characterized by a Mediterranean climate, with more mild temperatures and relatively more rainfall than Southern California (Elford, 1963). With this productive climate and diverse landscape it's no wonder this area of California is also considered a hotspot for threatened and endangered species (Flather et al., 1998). Like many natural phenomena, there is a large degree of heterogeneity in spatial patterns of habitat characteristics and these characteristics are often overlooked in coarse satellite-based classification methods (Morgan et al., 2010). In addition to this natural complexity, Northern California is known for difficult atmospheric conditions, including, haze, fog, and clouds (Augyte & Simona, 2011; Schlosser & Eicher, 2012). All of these factors make it difficult to accurately map habitats in Northern California from above. Here we

will discuss two indicative habitats of Northern California that present particularly complex mapping challenges, these are; intertidal submerged aquatic vegetation (SAV) in Humboldt Bay and oak dominant savanna habitat in the Napa Valley Watershed.

Eelgrass (*Zostera marina*) is a seagrass found in Northeast Pacific and North Atlantic estuaries where it occurs in low intertidal and shallow subtidal zones (Carter-Griffin et al., 2010; Waycott et al., 2009). Humboldt Bay in Northern California is home to a species of eelgrass, *Zostera marina*. Seagrasses like eelgrass are some of the most productive primary producers on the planet (Duarte & Chiscano, 1999) and play a vital role in providing habitat to diverse species (Carter-Griffin et al., 2010; Gilkerson, 2008). Eelgrass communities also contribute to complex detrital and grazer food webs (Phillips & Watson, 1984). All of the ecosystem services provided by eelgrass are susceptible to a wide range of disturbances that are responsible for direct habitat loss, due both to anthropogenic and natural processes (Waycott et al., 2009). Furthermore, as climate change and anthropogenic disturbances begin to change the dynamics of coastal estuaries, it is essential to examine how those changes may impact the services provided by eelgrass ecosystems (Shaughnessy et al., 2012; Waycott et al., 2009).

Quantitative monitoring of eelgrass and other seagrass systems is necessary for knowing how these ecosystems are changing through time. A series of *in situ* sampling methods have been employed; including scuba diving (Whippo et al., 2018), active sensing with side-scan sonar (Xu et al., 2020), and transect and quadrat sampling. While ground sampling of intertidal eelgrass habitat is detailed and reliable, it is also a potentially destructive, invasive, expensive, exhausting, and the spatial extent of a habitat

that actually get sampled is often low (Merkel & Consultant, 2017). This situation has motivated scientists to identify the most cost-effective methods for seagrass surveys, eventually leading to the development of standard operating protocols for ground sampled coverage estimates (Dean & Bodkin, 2016; Radloff et al., 2013), benthic habitat mapping, and even aerial image analysis (Clinton et al., 2007). More recent efforts have been made to track eelgrass coverage using remotely sensed satellite imagery and image classification analysis techniques rather than intensive on the ground efforts (O'Neill & Costa, 2013).

While there have been many efforts to identify eelgrass with remote sensing methods, researchers still face several obstacles. One limitation is the temporal frequency of image capture which is either dependent on satellite configuration or pre-planned flight schedules (Al-Wassai & Kalyankar, 2013). To survey intertidal phenomenon, it is important to capture data during low tide cycles that also align with good lighting conditions (Nesbit, 2018). With an uncontrollable data capture time in most remote sensing platforms, it is often difficult to obtain data captured at low tides that also possess optimal light conditions to yield sufficient image quality (Digruttolo & Mohamed, 2010). For this reason, remote sensing data, although , is quite limited for intertidal areas (Klemas, 2016). Remote sensing data of eelgrass is also known to underestimate areas of subtidal eelgrass as well as sparse eelgrass (Meehan et al., 2005). Furthermore, with the coastal influence and climatic conditions of Humboldt Bay, the frequently foggy or hazy days reduce the ability to capture intertidal remote sensing imagery (Judd et al., 2007). Water conditions such as turbidity, the extent of eutrophication, and wind wave action

can make it difficult to obtain clear remotely sensed images of eelgrass in temperate estuaries because of light attenuation and spectrally confusing surface reflectance (Reshitnyk et al., 2014).

The second study area for my thesis is the Napa watershed of Northern California, which is located north of San Francisco Bay and marks the northern portion of the Central Valley. There are unique habitat types that possess particular characteristics of ecological interest within the Napa watershed area. One example is the expansive oak dominant savanna, characterized by a combination of grassland and sparse or clustered oak tree canopy cover (~10-70%) (Grossinger et al., 2008; Law et al., 1994). A variety of endemic wildlife depend on the unique habitat provided by oak woodland and savanna, including hundreds of vertebrate species, thousands of invertebrate species, and almost 1500 flowering plants (Grossinger et al., 2008). In addition to hosting several important native wildlife species, the Napa Valley falls within the Pacific Flyway, attracting large populations of migratory birds (Grossinger et al., 2008). These habitats are an important ecological indicator of ecosystem function and biodiversity (Mahall et al., 2005). This area of California is also well known for its expansive vineyard agriculture. With an estimated 45,000 acres of vineyards in Napa Valley, it is expected many wildlife depend on these agriculturally modified areas for critical habitat (Wendt, 2016). These distinct habitat types, like oak dominant savanna, are not distinguished from other forested land cover types in most available LULC models for the Napa Valley, such as the National Land Cover Dataset. In addition to oak dominant savanna and vineyard agriculture, other

habitat types of ecological interest are often overlooked in most available LULC models, including riparian habitat and other specific agricultural types.

Previous efforts have been made to monitor oak savanna characteristics such as canopy cover, tree species composition, distribution, mortality and basal area (Colgan et al., 2012). Most of these efforts have involved on the ground sampling campaigns, though some of these ground data have also been incorporated in spatial models to identify more detailed ecosystem function such as soil dynamics (D. A. Robinson et al., 2010). While most of these efforts focus on oak savanna habitat, they often rely on ground sampled data (Karlik & Chojnacky, 2014) or spectrally calibrated *in situ* data (Colgan et al., 2012). That is why efforts have focused on satellite based modeling methods for oak savanna classification (Wolter et al., 2014), allowing researchers to map larger areal extents at different temporal intervals. Satellite mapping methods have been successfully developed to map oak savanna habitat (Wolter et al., 2014); however, these efforts focused on pristine oak savanna habitat, where model confusion with other similar vegetation and vegetation structure was not a problem. In addition to this oversight, these models are often coarse in resolution (30m or greater). Research has begun to focus on savanna classification in the face of new high spatial resolution data, employing new classification modeling methods such as machine learning and object based image classification (Whiteside et al., 2011). Object-based image analysis is an image classification method that works to incorporate contextual spectral and spatial information in the classification process (Blaschke, 2010). While object-based

classification, methods have shown the most promise in oak savanna mapping, these methods often require costly software and specialized software knowledge.

My thesis objectives are to present two case studies from Northern California that examine the use of land cover modeling in two challenging ecological scenarios. For the remote monitoring of intertidal eelgrass abundance in Humboldt Bay, CA, we compared a measure of eelgrass abundance derived from UAS multispectral imagery and image classification to measures of abundance from ground sampling at two eelgrass study sites. In order to address the second objective – the production of a more targeted land cover model in the Napa Valley watershed of California - we developed a GIS data fusion workflow that combined remotely sensed imagery with ancillary geospatial data.

MATERIALS AND METHODS

Humboldt Bay Modeling Methods

Study Area

Humboldt Bay is located in Northern California (40.75471°N; -124.21509°E) and is one of the largest marine estuary systems in the state, with a total area of 28.0 km^2 at mean lower low water (MLLW). It is comprised of three main sections, the shallow and relatively large South Bay, the deeper and larger Arcata Bay, and finally, near where the two Bays meet is a narrow and deep entrance channel, leading to the Pacific Ocean. Unmanned Aerial Vehicle (UAV) imagery was collected for two ~ 6 ha study areas of South Humboldt Bay (Figure 1). South Humboldt Bay is characterized by fairly uniform depth throughout and shallow waters. Eelgrass is known to occur in large distributions within South Bay, exhibiting some of the most pristine eelgrass populations along the West Coast (Schlosser & Eicher, 2012) . These two study areas are characterized by two unique management regulations, where one site is located within a Marine Protected Area (MPA), while the other does not fall within an MPA. The two sites selected are also known to exhibit diverse growth characteristics representative of different eelgrass conditions in Humboldt Bay, where the MPA site overall exhibits shallower and less dense eelgrass growth and the SBW site overall exhibits deeper and denser eelgrass growth. This provides representative data along the natural ecological gradient of eelgrass growth.

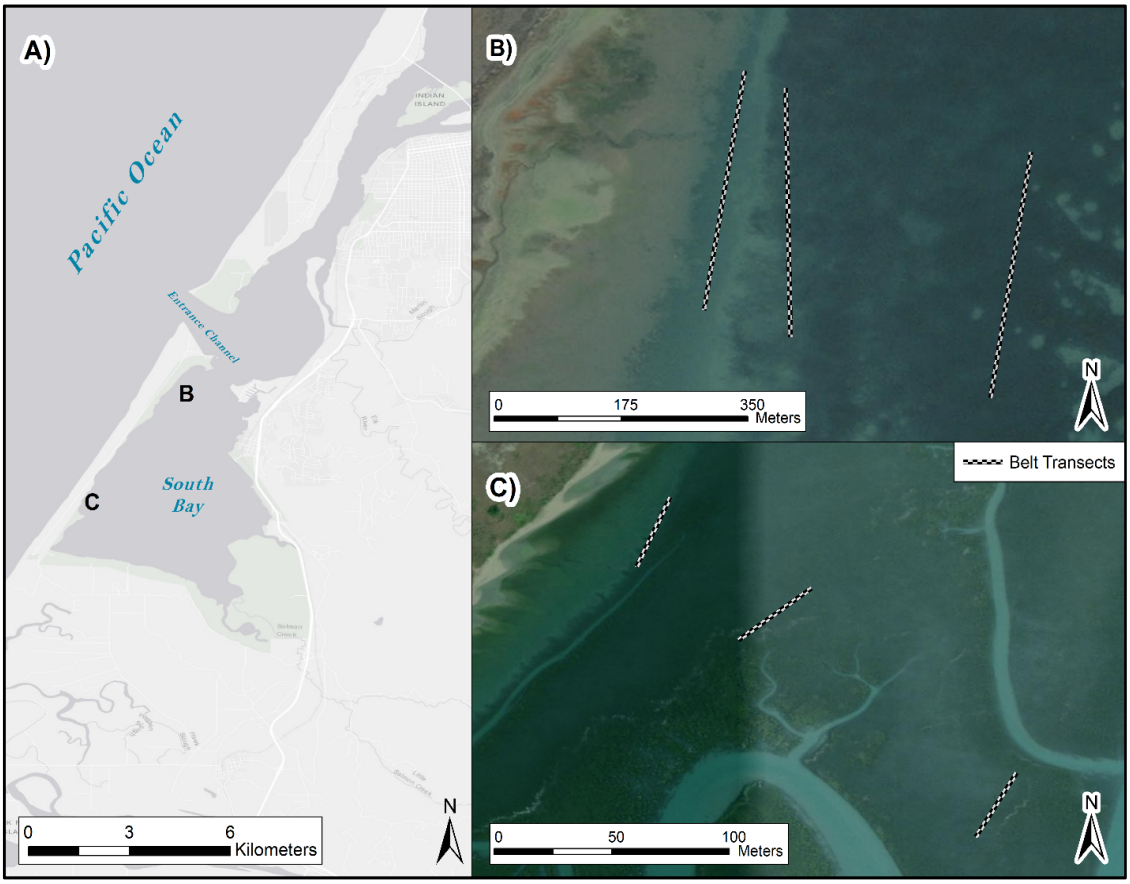


Figure 1. A) Maps showing the two study area locations within South Humboldt Bay, California with insets. B) Marine Protected Area study site showing the three tidal ground sampling transects. C) South Bay West study site showing the three tidal ground sampling transects.

UAV Data Collection and Processing

UAV data was collected using a DJI Matrice 100 equipped with a MicaSense RedEdge sensor. The DJI Matrice quad-copter drone with a customizable platform allows for the integration of multispectral image sensors and precise GPS data. MicaSense RedEdge sensor images were captured in raw, TIF format, with five spectral bands ranging from 475 nm to 840 nm and a bandwidth of 20 nm. Images of the MicaSense

Reflectance Calibration Panel were collected prior and post each UAS flight. UAS image data was collected at a flying altitude of 50 m, resulting in an image resolution of approximately 3.88 cm per pixel. UAS flights began at the MPA site at 08:22 am on June 19, 2019 and ended at 10:06 am PST, while the UAS flights began at the SBW site at 08:21 am and ended at 10:03 am PST. UAS flight conditions at the MPA site on June 19, 2019, consisted of little to no wind, light overcast skies, and stable conditions with a low tide of -1.21 ft MLLW, estimated at 08:05 am PST (NOAA, 2019). UAS flight conditions at the SBW site on June 20, 2019 differed from those at the MPA site, where hazy skies turned to clear and sunny skies with greater wind gusts, creating more inconsistent conditions with a low tide of -0.95 ft MLLW, estimated at 08:43 am PST (NOAA, 2019). The images obtained from the UAS flight were processed using AgiSoft Photoscan Professional (version 12.5) to create and georectify an orthomosaic raster image. UAS images were calibrated for reflectance using calibration reflectance images, sunshine sensor data, and down-welling light senor data. Images were georeferenced, a method used to tie an image to a geodetic network, using ground tie points obtained with a Reach RS+ Emlid Global Navigation Satellite System (GNSS). Reach RS+ Emlid systems collect GPS data using real-time kinematic positioning to obtain highly precise data. Five, 18-inch ground control point (GCP) targets were distributed along the shore near the eelgrass meadow study sites. The GPS location of the center point of each target was recorded using the GNSS system.

Field Data Collection

In situ data were collected for both sites during the summer of 2019. Three permanent 100 m sampling transects were established at the low, middle and high portions of the eelgrass beds in each site (Figure 1). PVC poles were inserted at the ends of the transects (0.0 m and 100 m) with poles at 20 m increments, where all PVC pole points were recorded using a Garmin Montana Series GPS, with accuracies within 3.65 meters. Along each 100 m transect, 20 sampling quadrats were distributed along each tidal transect at random. Each quadrat was photographed using a waterproof Fujifilm camera (mo. FinePix XP70) on a constructed square white PVC photograph frame of either 0.5 m^2, 0.25 m^2, 0.0625 m^2, depending on the density of eelgrass shoots. To sample eelgrass metrics, 30 random points were placed within the photo quadrat. Percent cover of eelgrass and green algae (*Rhizoclonium*) were then determined using a point intercept method, by finding the proportion of the 30 sampled points that were visually directly above eelgrass shoots and green algae. These data were collected under the Humboldt Ocean Carbon Observatory & Eelgrass Monitoring Baseline project and funded by the California Ocean Protection Council in 2017. Project investigators included Humboldt State University Professors Dr. Jeffery Abell, Dr. Frank Shaughnessy, Dr. Joe Tyburczy, University of California Davis Professor Dr. Tessa Hill, Wiyot Tribe personnel, and California Department of Fish and Wildlife personnel.

Feature Selection

To classify eelgrass presence and percent cover using high resolution multispectral UAS imagery, I tested the utility of two different classification methods;

supervised classification and unsupervised classification. However, prior to classification, a series of vegetation indices, band ratios, and textural component raster mosaics were developed to improve model predictions. Studies have shown simple vegetation indices such as the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973) could be applied to aquatic vegetation classification when vegetation was emergent from the water surface (Darvishzadeh et al., 2006). However, submerged aquatic vegetation such as eelgrass may be better evaluated using a water adjusted vegetation index such as the Normalized Difference Aquatic Vegetation Index (NDAVI) (Casado, 2019). With the implementation of vegetation index raster data, we were able to more accurately model eelgrass in intertidal areas.

In addition to vegetation index information, other efforts have shown textural components to be valuable in the discernment of eelgrass from other intertidal phenomenon (Duffy et al., 2018). Textural statistic matrices are derived using grey scale images and moving window calculations to create co-occurrence matrices (Gebejes & Huertas, 2013). Using this matrix, a series of statistics such as mean, variance, entropy, etc., can be calculated to create a raster with data regarding the cell neighborhood characteristics (Zvoleff, 2019). These types of statistics have been used to identify unique features in a landscape and help to better discern patterns of pixels rather than pixel values (Marceau et al., 1990). With the use of additional textural information variables, we were able to predict eelgrass coverage and distribution.

Feature selection is a method of statistical analysis that helps to reduce collinearity in model covariates (Dash & Liu, 1997). Collinearity exists when two

covariates overlap highly in a feature space, providing redundant information for model training (Dormann et al., 2013). Collinearity can introduce unnecessary complexity and reduce the model's predictive power (Dormann et al., 2013). One method of feature selection uses learning vector quantization.

Supervised Classification with Support Vector Machines

Supervised classification is a type of classification that requires the input of user defined training data (Lillesand et al., 2008). Training data is a sample of pixel values representing each desired feature class, creating a spectral signature for each feature class (Srivastava et al., 2012). Using this sample, supervised classification models can predict the distribution of feature classes throughout an image based on pixel values (Lillesand et al., 2008). Supervised classification in this study was implemented with Support Vector Machine (SVM) classifiers.

Support vector machines are a popular classification algorithm in image analysis and mining, where a set of non-linear decision boundaries are drawn in n-dimensional variable space (Hsu et al., 2003). These linear decisions are made using a series of training support vectors that together define a hyperplane separating the feature classes (Hsu et al., 2003). SVM algorithms were first developed in 1963 by Vladimir Vapnik (Vapnik & Lerner, 1963), and were later adapted to define a non-linear classification method (Cortes & Vapnik, 1995). SVMs are now one of the most successful modeling methods and have been applied to classification and regression problems (Ebrahimi et al., 2017). These models are best suited for binary classification (Harrington, 2015). SVMs have also become a preferred remote sensing classification method (Zhang et al., 2013)

and are often used to classify high resolution imagery (Anthony et al., 2009). This type of classification is robust to small sample sizes and does not make any assumptions of the data, allowing for flexibility in its application (Clemmensen et al., 2011). These classifiers are often criticized for overfitting (Han & Jiang, 2014). To combat this concern, model parameters can be adjusted to avoid overfitting (Cawley et al., 2007). This type of learning algorithm is often controlled with two parameters, sigma and cost (Jed Wing et al., 2019). To classify eelgrass coverage and density, SVM models were developed using the classifying features derived from feature selection.

Unsupervised Classification

Unlike supervised classification, unsupervised classification is a method of classification that does not require training data (J R Jensen, 1996). Unsupervised classification creates spectral classes by grouping pixels based on their spectral similarities; whereas, supervised classification uses reference data to group pixels into informational classes (Long & Srihann, 2004). Unsupervised classification can provide a more objective classification of eelgrass, because user bias from training data is removed (J R Jensen, 1996). This pixel-based method of classification uses a series of algorithms to differentiate spectral classes based on the natural clustering of image reflectance values and sometimes image textural components (Villanueva-Rivera et al., 2011). Using this objective method of classification, eelgrass percent cover classifications may be more robust. The unsupervised classification algorithm K-means was used within the RSToolbox package in R (Leutner et al., 2017). This algorithm was similarly used to classify eelgrass presence and coverage the United Kingdom (Duffy et al., 2018).

Accuracy and Uncertainty

Classified images were assessed in the statistical software R, using the 'caret' package (Jed Wing et al., 2019). To assess the eelgrass classification, accuracy, specificity, sensitivity, and the kappa coefficient of agreement can be derived from an error matrix. An error matrix is an array of columns and rows used to compare the number of samples of each feature class in the reference data relative to the classification model (Congalton & Green, 1999). To create an error matrix, samples of the image classification are taken repeatedly over 100 iterations to incorporate chance. Overall accuracy is an estimate based on the percent of correctly classified data (Remesan & Mathew, 2014). Sensitivity is defined as the ability to classify presence correctly, while specificity is defined as the ability to classify absence correctly (Parikh et al., 2008). Cohen's Kappa coefficient of agreement is considered to be a more conservative or more pessimistic accuracy assessment (Kottner, 2009), where the models overall agreement is assessed with the addition of this agreement occurring by chance (Verostek, 2014).

Virtual quadrats were created in a GIS for prediction assessment. Each transect start and end point locations were defined, then a transect was drawn between these end points. Finally, a set of quadrats were delineated along each transect to represent ground sampled quadrats, this yielded $n = 56$ virtual quadrats at SBW, and $n = 54$ virtual quadrats at MPA. Predicted eelgrass percent coverage was evaluated against ground reference data using linear regression (Verostek, 2014). Using linear regression, the degree of similarity between the predicted percent coverage and the reference percent coverage data can be quantified with statistics like R^2 and the root mean squared

deviation (RMSD) (Piñeiro et al., 2008). R^2 measures goodness of fit in linear regression, demonstrating the proportion of variance in the observed values explained by the variance in the predicted values (Verostek, 2014). The RMSD can be used to quantify the mean deviation between the predicted values and the observed values (Piñeiro et al., 2008). If the linear regression yields a significant relationship, the model prediction of eelgrass percent coverage will be considered a good fit.

Napa Valley Modeling Methods

Study Site

The study area encompasses the greater Napa watershed of Northern California. The Napa watershed is located just northeast of San Francisco Bay. The Napa watershed covers an estimated 1,451 km^2. This area of California is characterized by a dry Mediterranean climate, influenced by the nearby ocean and mountain ranges. The study area is comprised of a diverse geographic landscape, with dominant land cover uses comprising forests, vineyard agriculture and urban development. In addition to the man-made land cover types in this area, there is also a set of unique natural and endemic land cover types such as oak dominant savanna, lush grasslands, and riparian forests. The study area was defined by the areal extent of available spatial data, see Figure 2.

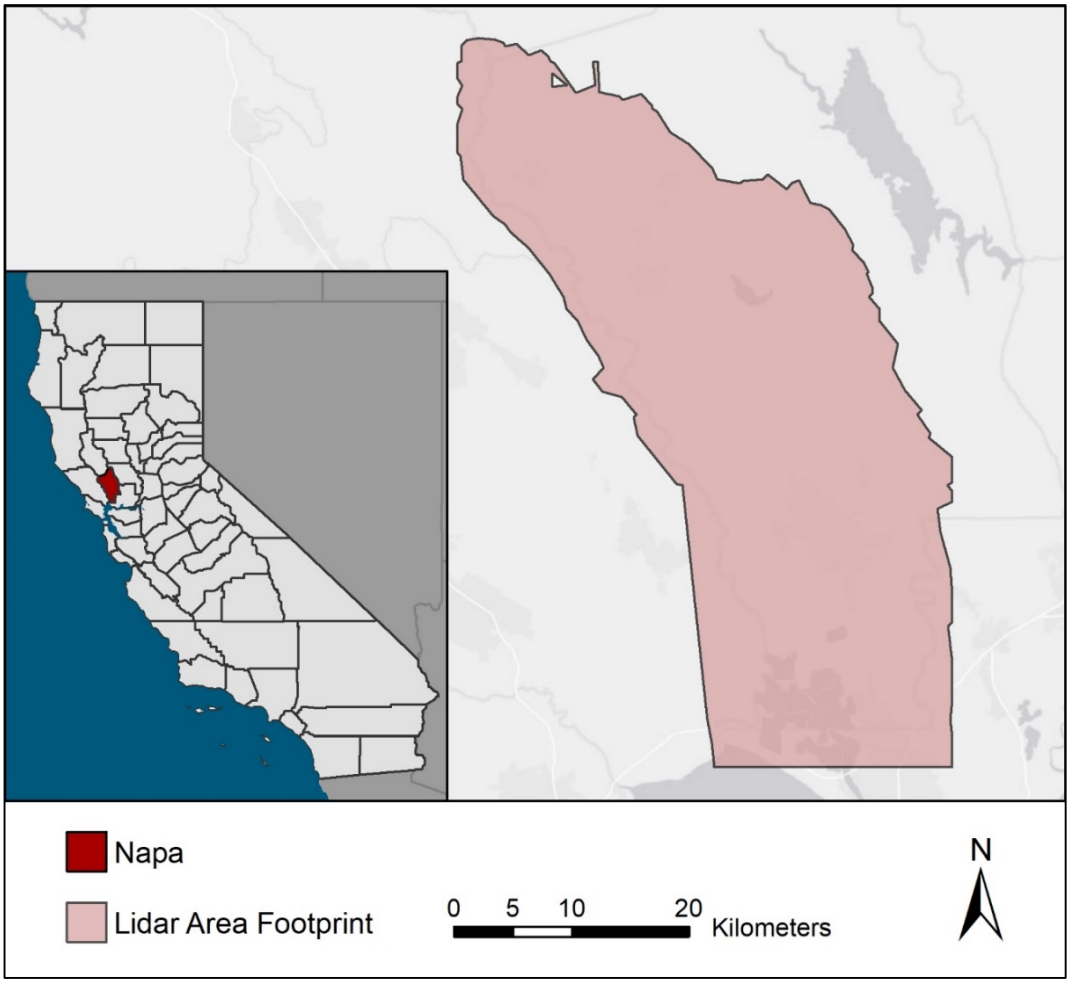


Figure 2. Map shows the collection area footprint for the lidar data collected by NCALM in 2003. The area is shown over the Napa Valley watershed, with an inset visualizing Napa County relative to the state of California.

LULC and Habitat Cover Types

The set of feature classifications for this study (water, vineyard agriculture, forest, grassland, oak dominant savanna, riparian and other) are unique to the Napa Valley area of California and require specific definitions for successful mapping efforts. Water land cover is defined as an area where water is the primary and persistent land cover type. The

first vegetation feature class to be defined is oak dominant savanna. According to the San Francisco Estuary Institute, oak savanna woodlands are known to exhibit a tree canopy cover of about 10 to 60% (Grossinger et al., 2008). While other studies suggest oak savanna can exhibit tree canopy cover of 10-70%, and that it is more typically between 25% and 75% cover (Minnesota, 2010). For this study we classified oak savanna using the widest range of tree canopy cover reported (10-75% tree canopy cover). This oak savanna definition also suggests that oak dominant savanna exists where both grassland and sparse tree canopy cover are present. Using this definition, we can then define forested cover as land covered by vegetation with heights of 1 meter or greater and a total canopy cover of greater than 75%. This leaves grassland to be defined as vegetative land cover with canopy heights of less than 1 meter. In addition, vineyard agriculture is defined as agricultural land dedicated to viticulture. Riparian areas are defined as areas that are within 30 meters of a river or stream, and where forest or oak dominant savanna is present. Finally land cover that does not exhibit any of the above characteristics was classified as other and is considered to include land cover that is urban, developed, or bare ground.

Data Collection and Preprocessing

To classify seven unique land cover types across the Napa Watershed, a series of remotely sensed and ground derived geospatial data were collected and analyzed. Firstly, multispectral imagery was downloaded for the study area. Multispectral imagery was collected from the USDA National Agriculture Imagery Program in 2010; acquisition and initial preprocessing is managed by the USDA. This year was selected based on the

limited cloud cover throughout the scene, this image collection year also matches collection year for other data sources. NAIP image data are available online in digital ortho quad quarters (DOQQ), each covering just under 50 km^2. To cover the entire scene of interest, 100 tiles were downloaded, mosaiced together and clipped to the region of interest. The resulting image had a spatial resolution of 1 meter and a four-band spectral resolution with nominal spectral bands (Blue, Green, Red, and Near-Infrared).

In addition to multispectral imagery, lidar data was obtained for the study area to classify vegetation heights. Lidar data was collected by the National Center for Airborne Laser Mapping (NCALM) in the summer of 2003 and was funded by the National Science Foundation. The lidar survey area covered the entire Napa Watershed, containing approximately 1,230 km^2, and required 14 aerial missions at a height of 700 meters above ground level. All survey efforts took place from May 15, 2003 to June 1, 2003. The aerial missions resulted in 129 flight lines and required an additional 20 lines to densify point data.

In addition to the remotely sensed data, two vector-based spatial datasets were utilized to aid in the classification of vineyard agriculture and riparian areas. Firstly, to clearly identify vineyard agriculture, a high-resolution and highly accurate vector-based agricultural coverage data were obtained from the Napa County GIS Department. These data were created using high-resolution digitizing over NAIP imagery from 2010. Secondly, to clearly identify riparian areas, a high resolution hydrography dataset delineating all streams and rivers was obtained from the USGS (U.S., 2009). Both vector-based datasets were visually inspected in a GIS to ensure data was properly aligned for

further use and integration with remotely sensed data. While all data were visually inspected, there are still several sources of error and uncertainty when multiple sources of data are utilized at different temporal intervals.

Habitat Classification Overview

The primary steps involved in this classification can be summarized by the flowchart in Figure 3. The NAIP imagery was classified into three feature classes (water, vegetation, and other) using supervised classification. Then, to further classify vegetation types, lidar data was processed to create a canopy height model with a spatial resolution of 1 meter. To extract height data for vegetation only, the previously classified NAIP image was used to extract lidar point data only for pixels classified as vegetation. Using the extracted lidar point data, grass or herbaceous vegetation was differentiated from tree vegetation with a canopy height threshold at an elevation of 1 meter (grass defined at < 1 meter and trees defined at ≥ 1 meter). Dense forested land cover was separated from individual trees and tree clusters to identify patchy oak savanna habitat, using a patch size threshold. Using the individual trees and tree clusters, the ESRI aggregate polygons function was used to aggregate these individual trees and tree clusters into oak savanna. Riparian areas were identified by buffering the acquired hydrography dataset to 30m and only selecting pixels with forest or oak dominant savanna land cover. Finally, all layers were combined in a GIS and rasterized to develop a single band classified raster surface.

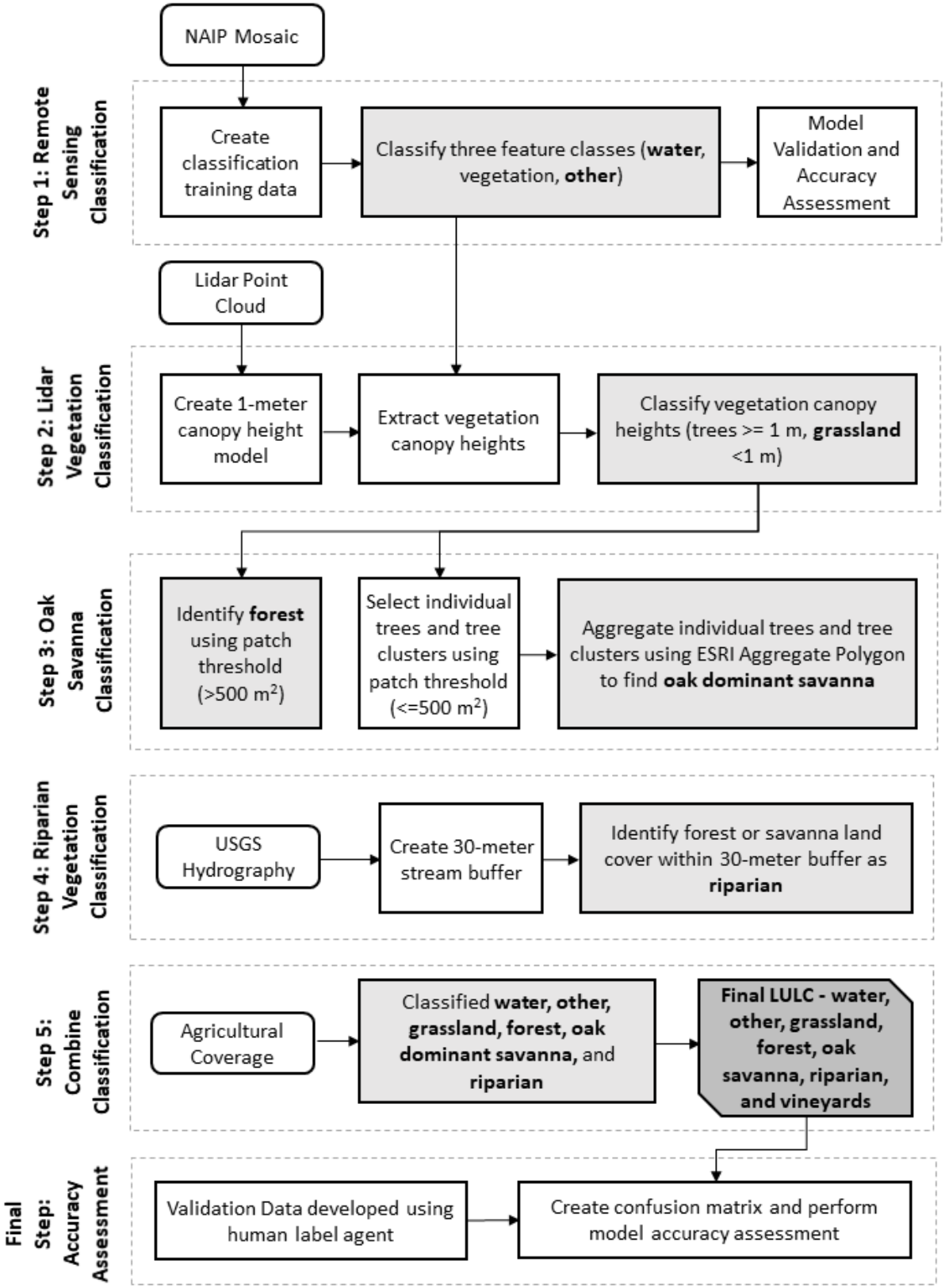

NAIP Mosaic
Step 1: Remote Sensing Classification
Create classification training data
Classify three feature classes (**water**, vegetation, **other**)
Model Validation and Accuracy Assessment
Lidar Point Cloud
Step 2: Lidar Vegetation Classification
Create 1-meter canopy height model
Extract vegetation canopy heights
Classify vegetation canopy heights (trees >= 1 m, **grassland** <1 m)
Step 3: Oak Savanna Classification
Identify **forest** using patch threshold (>500 m²)
Select individual trees and tree clusters using patch threshold (<=500 m²)
Aggregate individual trees and tree clusters using ESRI Aggregate Polygon to find **oak dominant savanna**
Step 4: Riparian Vegetation Classification
USGS Hydrography
Create 30-meter stream buffer
Identify forest or savanna land cover within 30-meter buffer as **riparian**
Step 5: Combine Classification
Agricultural Coverage
Classified **water, other, grassland, forest, oak dominant savanna,** and **riparian**
Final LULC - water, other, grassland, forest, oak savanna, riparian, and vineyards
Final Step: Accuracy Assessment
Validation Data developed using human label agent
Create confusion matrix and perform model accuracy assessment

Figure 3. A flowchart outlines the necessary steps to complete the land use/ land cover classification.

Remote Sensing Image Classification

The first and most necessary distinction in the modeled LULC classification is the identification of the three feature classes, vegetation, water, and other land cover. To classify these three feature classes, supervised classification was performed on a color infrared NAIP image with a 1-meter spatial resolution. Training data were collected in a GIS using heads up digitizing and high-resolution image interpretation. The standard is to use field collected training data points and accuracy assessment points. However, in this study these methods could not be done because of the large areal extent of the study region of interest. Training data were manual interpreted using expert knowledge. Training points were classified multiple times ($n = 5$), then averaged using a mode function to reduce interpretation bias. Using the collected training data, a radial-kernel support vector machine model was developed using the 'caret' package (Jed Wing et al., 2019), in the statistical software R (R Core Team, 2018). The classified image of the Napa watershed comprising the three desired feature classes is later used to further classify vegetation types. Because this classified image would be used in later analysis it was important to assess model accuracy prior to more detailed classification. To do so, a confusion matrix was developed during the model training process using the training data as ground truth data. The classified image was only used in further classification when accuracy standards were comparable to scientific literature, i.e. when a Kappa coefficient of 0.7 or greater was obtained, (Congalton & Green, 1999).

Vegetation Height Classification and Savanna Mapping

To differentiate the different types of vegetation, lidar point data collected in 2003, with a density of 1.43 pts/m^2 was used to create a 1-meter spatial resolution canopy height model. Canopy height pixels were extracted only for areas classified as vegetation in the three-class multispectral image classification. Using lidar with relatively low density, it is difficult to achieve accurate classification of different vegetation structures (Bujan et al., 2013). For this reason, lidar canopy heights derived from this dataset were only used to differentiate trees and grass vegetation. The extracted canopy heights of vegetation were classified as forest being 1 meter or taller and grass was defined as less than 1 meter. Because lidar densities were relatively low, further classification of vegetation types was too difficult.

To overcome these difficulties with further vegetation classification, a series of rule-based classifications were applied. Oak dominant savanna characteristically exhibits patchy canopy cover with a mixture of sparse individual trees, small clusters of trees and lush grasslands. First, individual trees and small tree clusters needed to be identified and separated from continuous dense forest. To do this we applied a patch size threshold for the classified tree pixels, with an area of 500 square meters. In other words, any tree patches equal to or greater than 500 m^2 were classified as forest land cover. The remaining tree pixels in patches less than 500 m^2 were considered individual trees and small tree clusters. These individual trees and tree clusters were then aggregated using the Esri aggregate polygons function, to capture a variation in canopy cover that was representative of oak dominant savanna in the Napa Valley. The aggregation distance for

oak savanna was set to 1000 meters. The areas captured in the aggregation were classified as oak dominant savanna.

Ancillary Geospatial Data

Classifying riparian vegetation and vineyard agriculture with remotely sensed data proved challenging, where streams were visually obstructed by tree canopy cover, and vineyard vegetation was difficult to distinguish both structurally and spectrally. For these reasons, it was necessary to incorporate ancillary geospatial data to properly map vineyard agriculture and riparian areas. Vineyard LULC was extracted from the Napa County GIS Department's ground-validated spatial dataset of all Williamson Act agricultural lands. Riparian areas could not be identified using remotely sensed imagery or lidar. This phenomenon has been seen in other classification and is due to the complex definition of riparian areas (Cunningham, 2006). In this classification riparian areas are a combination of forest or oak dominant savanna land cover that is within 30 meters of a stream or river. In order to identify riparian areas, it was necessary to incorporate contextual data with the addition of GIS data. First streams, rivers and other flow lines needed to be identified. Because these flow lines and streams were unable to be identified using remote sensing imagery, we incorporated a high resolution hydrography dataset from the USGS (U.S., 2009). These data are mapped at a 1:24,000 scale and identify the spatial geometry of predominantly rivers and streams.

Accuracy Assessment

Confusion matrices are a commonly used method for assessing the correctness of image classification models (Congalton & Green, 1999). Ground reference data were

created for the study area using a simple random sample of the classified image (Nowak et al., 2006). Prior to classification, 500 points were manually classified using high-resolution imagery from 2003, using three human label agents similar to previously used methods in land cover classification accuracy assessment (Kennedy et al., 2015; Li et al., 2017). The three sets of human-labeled points were then combined using a mode function. Using a GIS, the predicted model classifications were extracted for each of the 500 points for cross tabulation. The 500 points with manually classified labels and the predicted land cover model labels were then cross tabulated in a confusion matrix. A confusion matrix offers a series of measures for model accuracy, making it a useful tool in land cover model assessment (Verostek, 2014). These matrices provided estimates such as specificity and sensitivity as well as balanced accuracies. The kappa coefficient of agreement was also derived from the confusion matrix. Using the kappa coefficient of agreement, we hope to account for the possibility of chance agreement.

RESULTS

Humboldt Bay Modeling Results

UAV Data Collected

The unmanned aerial vehicle flights conducted at the MPA site, on June 19, 2019, flown at a height of 52.7 meters, yielded 3,025 useable images with approximately 75% image overlap. The orthomosaic derived from the AgiSoft Photoscan processing had a ground resolution of 3.50 centimeters per pixel with a reprojection error of 0.686 pixels. With an overall size of 10,124 by 12,288 pixels the orthomosaic covered approximately 0.07 km^2, Figure 4.

The unmanned aerial vehicle flights conducted at the SBW site, on June 21, 2019, flown at a height of 55.7 meters, yielded 9,390 useable images with approximately 75% image overlap. The orthomosaic derived from the AgiSoft Photoscan processing had a ground resolution of 3.76 centimeters per pixel with a reprojection error of 0.700 pixels. With an overall size of 21,503 by 20,480 pixels the orthomosaic covered approximately 0.17 km^2, Figure 5.

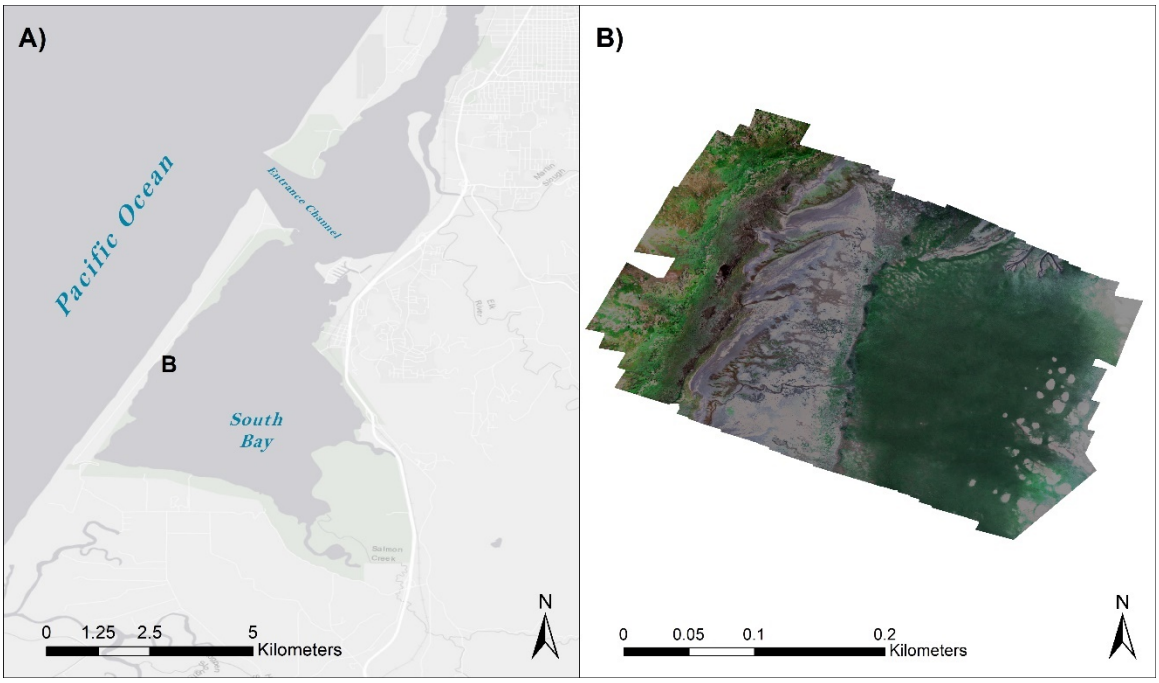


Figure 4. A map shows A) the location of the Marine Protected Area (MPA) sample location and B) the true color orthomosaic obtained from UAV imagery at the MPA site.

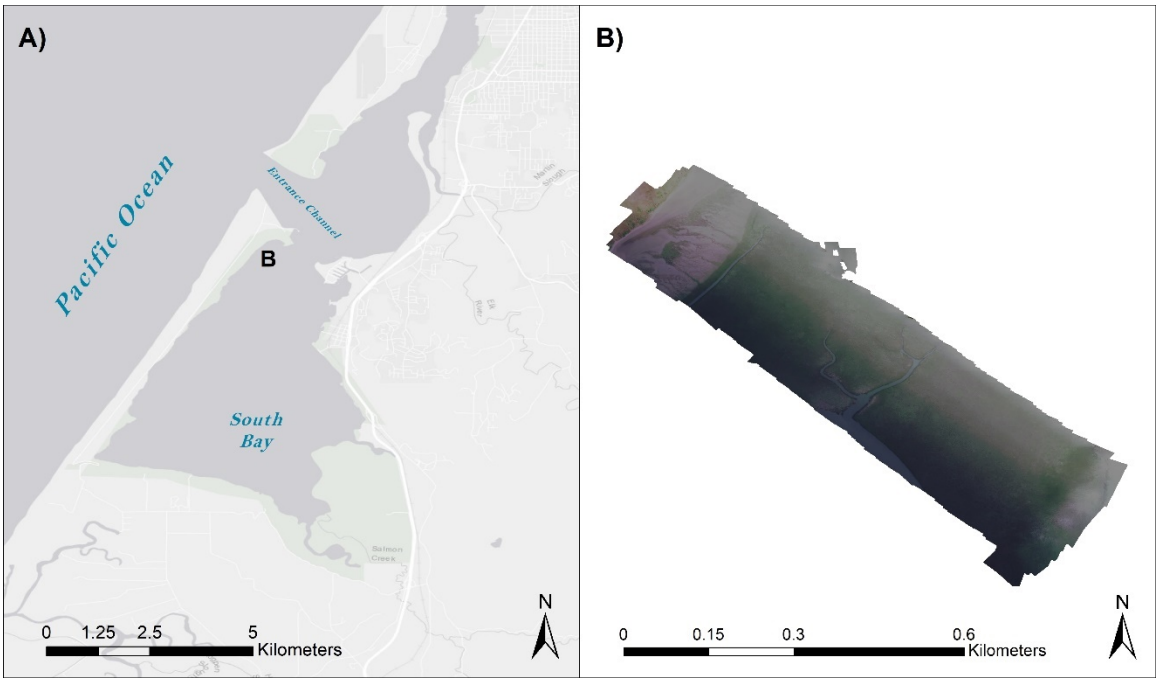


Figure 5. A map shows A) the location of the South Bay West (SBW) sample location and B) the true color orthomosaic obtained from UAV imagery at the SBW site.

Percent Cover Estimates

Percent cover proportions were extracted to virtual quadrats using the modeled presence/absence binary raster image. The proportion of pixels classified as presence or absence was extracted as a percentage for each virtual quadrat to compare with on the ground quadrat percent cover. The best predicted percent coverage and ground sampled percent coverage for the MPA site resulted in a coefficient of correlation or R^2 value of 0.6496 and a root mean squared error of 27.44, Figure 6, while the best model for the SBW site resulted in an R^2 value of 0.1036 and a root mean squared error of 50.23, Figure 7.

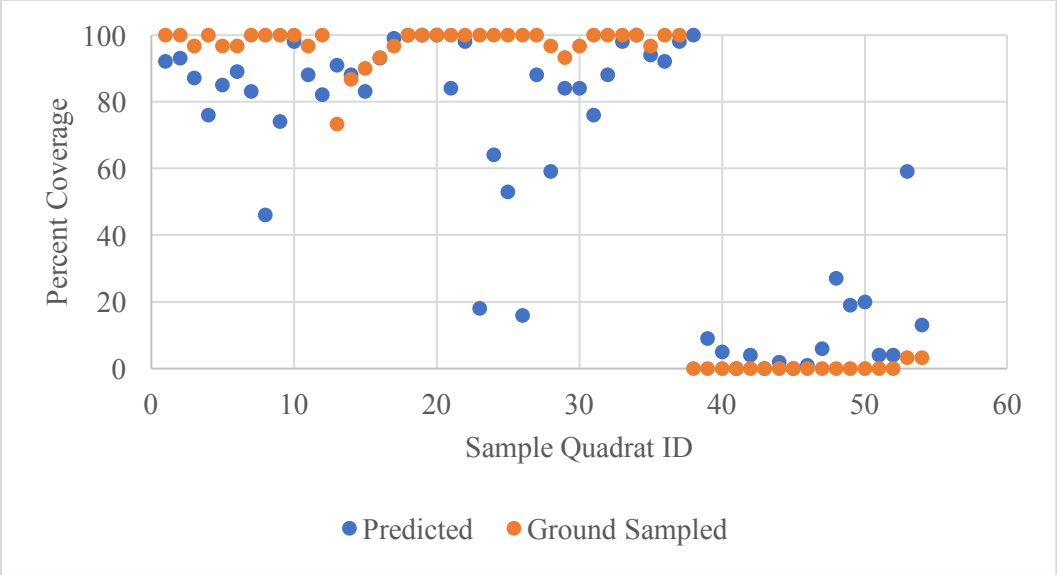


Figure 6. The predicted (blue) and ground sampled (orange) percent cover of eelgrass are shown for the MPA site with quadrat sample ID shown across the x-axis.

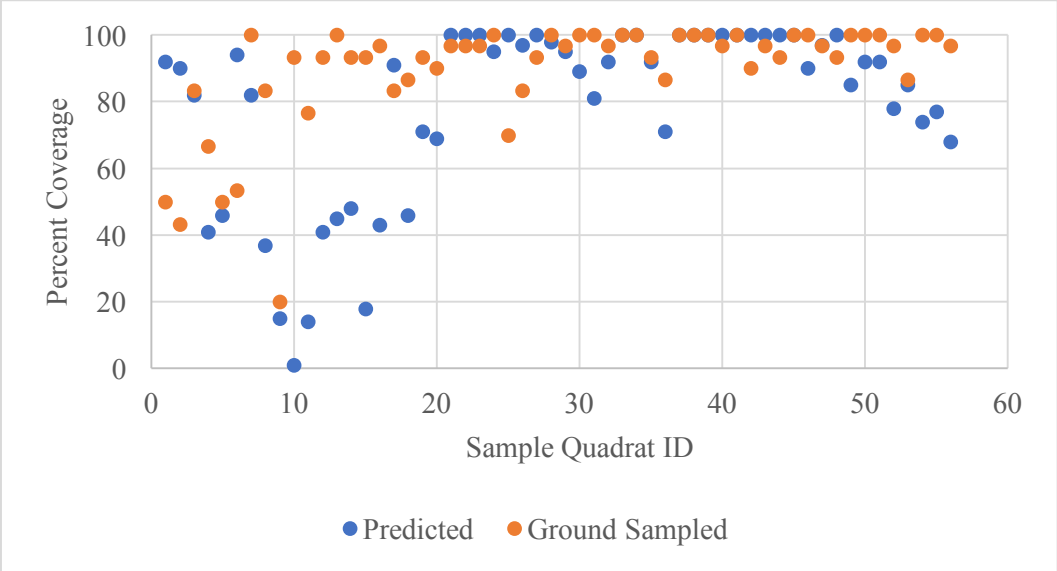


Figure 7. The predicted (blue) and ground sampled (orange) percent cover eelgrass are shown for the SBW site with quadrat sample ID shown across the x-axis.

The results of all models at the MPA site are shown in Table 1 showing the two best model to be the supervised SVM classification with all variables and the unsupervised classification with all variables and $n = 4$ classes. The results for all models at the SBW site are shown in Table 2, indicating again the best models are the SVM supervised classification with all variables and the unsupervised classification with all variable and $n = 4$ classes.

Table 1. Summarized model information is shown with a correlation coefficient for the results of a linear regression between each model prediction and ground sampled percent coverage data for the MPA site.

Model	Variables	No. Classes	R-Squared
SVM	Multispectral	2	0.1816
SVM	Multispectral + Indices	2	0.6496
K-means	Multispectral + Indices	2	0.0001
K-means	Multispectral + Indices	3	0.3699
K-means	Multispectral + Indices	4	0.6038
K-means	Multispectral + Indices	5	0.0368

Table 2. Summarized model information is shown with a correlation coefficient for the results of a linear regression between each model prediction and ground sampled percent coverage data for the SBW site.

Model	Variables	No. Classes	R-Squared
SVM	Multispectral	2	0.0808
SVM	Multispectral + Indices	2	0.1036
K-means	Multispectral + Indices	2	0.0032
K-means	Multispectral + Indices	3	0.0553
K-means	Multispectral + Indices	4	0.2023
K-means	Multispectral + Indices	5	0.0014

Predicted eelgrass coverage and ground sampled eelgrass coverage are broken down by transect, this isolates each transect and highlights the sources of error in the predictions. This breakdown is shown for the SBW site in Appendix A and the MPA site in Appendix B.

Napa Valley Modeling Results

The results of the habitat classification are shown in Figure 4. Most of the cultivated land cover types are seen in the central portion of the Napa Valley, with other vegetative types covering the less populated areas of the valley. The majority of classified riparian habitat is located within forested land cover. Water bodies were primarily found in the southern most portion of the scene where the San Pablo Bay is met by the Napa River. In addition, a minority of the water bodies are shown in the north central area of

the study site. Most of the classified oak dominant savanna is shown in the eastern side of the valley and a minority of the pixels are also shown in the northern portion of the scene.

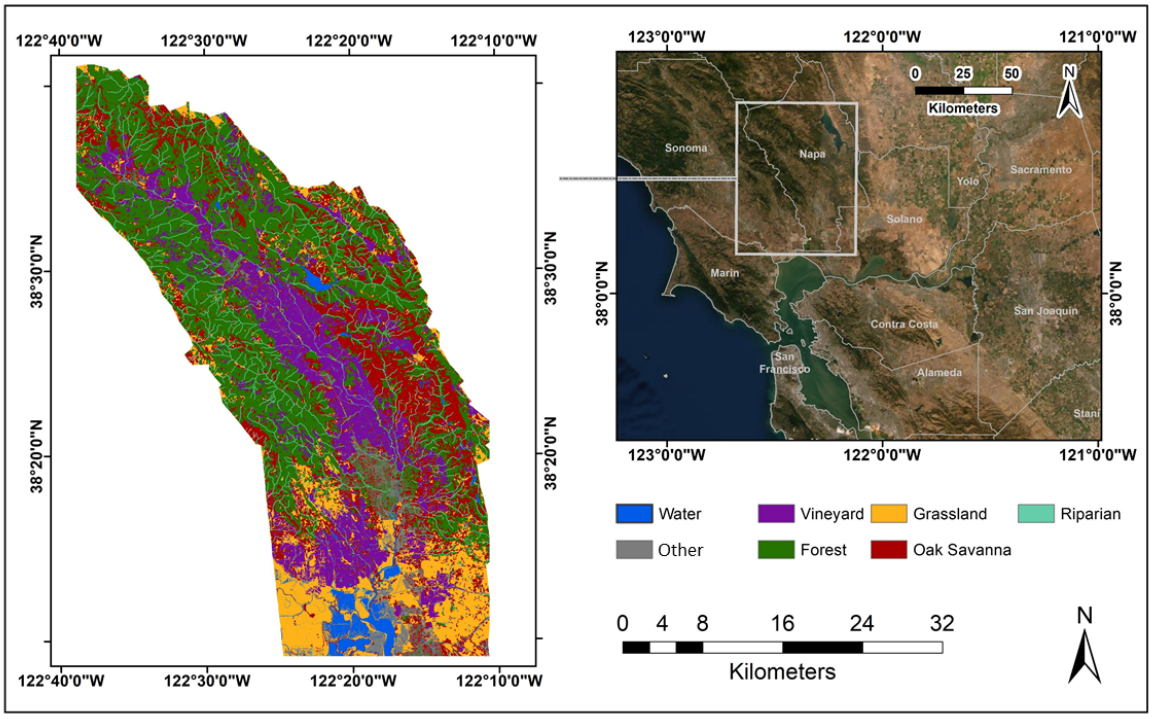


Figure 8. The final habitat land cover classification is shown in respect to California and neighboring areas.

Classification accuracy statistics were calculated for each feature class (Table 1) using a confusion matrix (Table 2). Forested land cover was the dominant land cover class comprising 36.76% of all classified pixels. The least dominant land cover classes were water, bare ground, and riparian with 3.06%, 4.96%, and 5.79% of all classified pixels, respectively. Therefore, these three classes exhibited the lowest detection rates of all the feature classes, Table 1. The remaining land cover classes (vineyards, grassland, and savanna) were more frequently classified with 16.24%, 12.89%, and 20.30% of all classified pixels respectively. Classified vineyard land use obtained the highest balanced

accuracy, 95.66%. The high accuracy of vineyard classification is due to the accuracy of the vector-based agriculture data. In addition to vineyard land use, the feature classes water and other obtained the next highest balanced accuracies. The remaining feature classes obtained balanced accuracies between 71.99% and 81.51%. These remaining feature classes, forest, grassland, oak savanna, and riparian have the lowest balanced accuracies, and this is largely due to their spectral similarities as these are all different vegetation types.

Table 3. Accuracy assessment statistics for individual classified features.

	Water	Other	Vineyard	Forest	Grassland	Savanna	Riparian
Sensitivity	0.7895	0.7500	0.9326	0.7409	0.4904	0.5938	0.5455
Specificity	0.9958	0.9854	0.9805	0.8893	0.9495	0.8601	0.9652
Pos Pred Value	0.8824	0.6818	0.9121	0.8079	0.7183	0.3838	0.2609
Neg Pred Value	0.9917	0.9895	0.9853	0.8452	0.8765	0.9352	0.9895
Prevalence	0.0380	0.0400	0.1780	0.3860	0.2080	0.1280	0.0220
Detection Rate	0.0300	0.0300	0.1660	0.2860	0.1020	0.0760	0.0120
Detection Prevalence	0.0340	0.0440	0.1820	0.3540	0.1420	0.1980	0.0460
Balanced Accuracy	0.8927	0.8677	0.9566	0.8151	0.7199	0.7269	0.7553

Using the manually classified ground reference data, created using multiple human label agents, and the model prediction a confusion matrix was cross tabulated, Figure 9. By summing the diagonal of the confusion matrix, we derived an overall model accuracy of 70.20%. In addition to the overall accuracy, we also calculated the Kappa coefficient of agreement, obtaining a value of 0.6140. We can see the most frequently

occurring feature class was forest land cover and therefore had the greatest number of ground reference points. Using the confusion matrix, we can also see the land cover classes with the most confusion were the forest, grassland, and oak savanna. The classification methods described in this study yielded an overall accuracy of 70.2% and was 18.6% more accurate than the next most accurate readily available LULC datasets (NLCD, 2004).

		Reference						
		Water	Other	Vineyard	Forest	Grassland	Savanna	Riparian
Prediction	Water	**15**	0	0	1	1	0	0
	Other	3	**15**	0	0	4	0	0
	Vineyard	0	2	**83**	2	3	0	1
	Forest	0	1	1	**143**	10	20	2
	Grassland	0	0	4	9	**51**	6	1
	Savanna	1	2	1	26	30	**38**	1
	Riparian	0	0	0	12	5	0	**6**

Figure 9. Cross-tabulated prediction and reference data make up a confusion matrix for the final model, used to derive accuracy statistics.

DISCUSSION

Humboldt Bay Vegetation Modeling Discussion

UAS Survey Data for Eelgrass Monitoring

While UAS imagery is relatively easy to obtain, there are several factors that can impact the quality and consistency of these data. Some of these factors include light conditions, atmospheric conditions, wind speed, flight speed, and sensor capture orientation (Whitehead & Hugenholtz, 2014). Often times image artifacts are exemplified when all flight images are mosaicked, resulting in geometric artifacts making the image blurry, choppy, striped and holes may even appear in the final mosaic (Whitehead & Hugenholtz, 2014). Therefore, it is necessary to perform careful flight planning with a goal of obtaining consistent conditions across all study sites, including tidal height of the water at the time of data capture. It may be necessary to obtain more detailed tidal height predictions to optimize data capture, this is because tidal predictions are provided at a course spatial resolution. This course spatial resolution may mean tidal predictions in one area of the Bay do not represent tidal heights throughout the entire bay.

It is likely that conditions at the SBW site were less favorable for eelgrass monitoring and with higher wind speeds and hazy conditions image capture was more inconsistent than was at the MPA site. More specifically, the image data for the two sites were visually disparate, as a result of hazy and intermittent sunny conditions present at the SBW site on June 21, 2019. This resulted in image artifacts such as haze, vignetting,

and hot spots (Whitehead & Hugenholtz, 2014). While image calibration was completed for all sensor data, conditions were too highly variable for all flights within the same day, making image calibration difficult to perform across all flight data.

There is a considerable difference in eelgrass growth characteristics at each site, where eelgrass was found at deeper tidal elevations at the SBW site than the MPA site. More specifically, the lowest transect at the MPA site is at a higher tidal elevation (0.14 MLLW) than the highest of the transects at the SBW site (-0.01 MLWW). These characteristics along with the difference in tidal height during the two UAS flights could explain the difference in model accuracies across study sites. More notably, eelgrass at the deep tidal elevations within the SBW site were more likely covered with water which likely obscured the spectral signature of eelgrass.

The discrepancies captured in this study data exemplify the difficulty experienced when collecting image data of intertidal phenomenon at multiple sample locations and especially in the Northern California area. This makes the case that UAS deployment is still the best option for intertidal remote sensing, with the ability to easily perform repeat UAS flights to capture more consistent conditions. UAS can allow for more targeted data capture with the ability to adapt to tidal cycles in order to obtain data at low tides. UAS also allows for adjustable flight height for the target subject or land cover type. All of which can be costly or impossible with other remote sensing platforms.

Mapping Eelgrass with UAS Imagery

Both sets of imagery obtained for both eelgrass sites yielded a resolution of 3.50 to 3.76 centimeters; however, it is possible that this resolution is not sufficient to capture

the level of detail exhibited in eelgrass meadows (O'Neill et al., 2011) . Individual eelgrass leaves were not visible at 3.50 – 3.76-centimeter resolution. It may beneficial to decrease the flight height of data capture to increase the spatial resolution of image data (Whitehead & Hugenholtz, 2014). However, this may require increased flight times and larger data storage. With finer spatial resolution imagery, we may be able to distinguish individual eelgrass leaves which could be useful for more detailed health and vigor analysis (Duffy et al., 2018). With finer spatial resolution imagery, it may not be necessary to include textural features and vegetation indices (Duffy et al., 2018).

While eelgrass percent coverage estimates achieved an r-squared value of 0.6496, it is difficult to compare this metric to other studies, most of which map other eelgrass metrics (such as distribution and total area) or utilize different accuracy statistics (Duffy et al., 2018). It appears that eelgrass percent coverage may be over estimated at the MPA site which is likely due to specificity error in the classification, where eelgrass absence is classified as presence instead (Parikh et al., 2008). On the other hand, eelgrass percent cover predictions at the SBW site appear to underestimate eelgrass coverage, which may be due to more water obstructing the spectral signature of eelgrass. In order to overcome issues with classification accuracy others have developed methods for water column corrections. However, these methods often require detailed spectral signatures and specialized algorithms, both of which can be costly and time consuming (Rowan & Kalacska, 2021).

In addition to environmental sources of error, our virtual quadrat to ground sampled quadrat comparison is a likely source of spatial error. It is necessary to improve

the spatial accuracy of the ground sampled data to reduce the error in virtual quadrat creation. It is recommended that a GNSS GPS system should be used to collect the precise locations of ground sample quadrat data. These methods have similarly been applied in the United Kingdom with success (Duffy et al., 2018). Green algae coverage can be compared with eelgrass percent cover estimates at both sites to understand if spectral confusion in vegetation types influenced model success. More specifically, we can see at the SBW site there is no reported green algae coverage, from the ground sampled data, Appendix C. This indicates, there is minimal vegetative confusion from green algae reflectance at the SBW site. On the other hand, at the MPA site we see considerable green algae coverage at the high tidal transect, ranging from 6.66 -100 percent coverage, Appendix D. This could indicate predicted eelgrass coverages could be overestimated at the MPA high transect, due to the presence of green algae, creating a confusing reflectance signature similar to that of eelgrass.

Napa Valley Land Cover Modeling Discussion

With an overall accuracy of 70.20%, the model proved to be more accurate than other available land cover datasets. However, a more informative model assessment may be individual feature class accuracies. The feature class with the highest accuracy was vineyard land use, with a balanced classification accuracy of 95.66%. This high classification accuracy of vineyard is due to the high accuracy of the Napa County Agricultural coverage data. These data were developed using high resolution GPS mapped coverage as well as ground-based surveys. I found the classes with the next

highest individual accuracies were the water and other feature classes. The other feature class is comprised primarily of bare ground, developed, and urban areas. This high individual accuracy in water and other features (mostly bare ground and developed areas) is likely due to their unique spectral signature, demonstrating these classes were easily classified using the methods described in this study (Srivastava et al., 2012).

I also found the feature classes with the lowest individual accuracies were those of different vegetation types. This confusion is expected due to the spectral similarity between different vegetation types. Much of the confusion between vegetation types involved the oak savanna feature class. It is possible that the definition of oak savanna is causing confusion in the model classification as oak dominant savanna is known to exhibit a combination of both grassland and forest canopy cover.

The most readily available LULC model for California is the National Land Cover Dataset provided by the USDA. These models are produced for the entire United States and are derived using satellite imagery from Landsat, resulting in a spatial resolution of 30 meters (Hollister et al., 2004). These models can provide general information with a standardized set of feature classes for the entire country. While these data are easy to access and does not require additional processing, these data do not always best represent the true LULC patterns on the earth's surface or the locally endemic land cover types (Homer et al., 2015). With the use of higher spatial resolution remote sensing data, I was better able to map the fine scale patterns of land cover and identify more informative feature classes. These more informative feature classes, such as

oak dominant savanna and riparian are more representative of the local vegetation characteristics.

Model Limitations

One of the limitations in this model was that the data used in this study were collected on different temporal scales. By using data from different years, it is possible that variability in the actual land cover between data acquisition times may cause uncertainty in the model predictions (Lin et al., 2002). For instance, a wildfire that burned through part of the Napa Valley Watershed in 2004 could impact the validity of lidar data collected in 2003 (Liu, 2008). In addition to temporal disparity in the data, there may also be classification error due to the acquisition of remotely sensed imagery. Specifically, by using NAIP imagery, it is possible to inherit some uncertainty from image preprocessing and georectification and that this uncertainty may be exacerbated by classifying as many as 100 individual image scenes. Finally, while lidar data was available for a large area, it was not acquired with a high frequency of laser pulses and therefore possessed a relatively low point density. While point density is a major limitation when classifying vegetation types based on canopy height, it is clear with higher density lidar data it may easier to distinguish vegetation types in the future (Bujan et al., 2013).

Oak Dominant Savanna Classification

Classifying oak dominant savanna proved most difficult. With a combination of variable canopy cover and grass understory, oak dominant savanna exhibited confusing spectral and canopy height characteristics. For this reason, we saw the most overlap in the feature classes oak savanna, forest, and grassland.

Management Implications

This model has the potential to benefit a variety of local stakeholders in the Napa Valley. In particular this LULC model could be used to model change after a disturbance (Kennedy et al., 2015). LULC mapping has been a vital resource in change detection and urban planning as well (Palamuleni et al., 2007; Suribabu et al., 2012). In addition to land cover change mapping and land use planning, this model can be used for other natural resource applications. For instance, this model could be used to model wildlife behavior and habitat selection similarly to other studies in this area (Huysman & Johnson, 2021; Wendt, 2016).

Future Iterations

The methods described here demonstrate the ability to combine remotely sensed multispectral imagery, lidar data and other ancillary geospatial data to map LULC types. The classified feature classes have been defined based on local habitat characteristics and biologist knowledge. This means future applications of the LULC model can be adapted to other habitat definitions based on local wildlife needs and vegetation characteristics. With the need for land cover data increasing and the need for landscape level data, it is important that modeling applications work with large datasets (Ray, Ibironke, Kommalapati, & Fares, 2019). While this method was applied programmatically and was time efficient, it did require costly geospatial software for savanna classification. In order to improve accessibility and cost efficiency, future iterations of this method should be created using fully open source software.

CONCLUSIONS AND RECOMMENDATIONS

Humboldt Bay Modeling Conclusions

In this study I demonstrate the value of using a commercial grade image sensor and a small UAV to model eelgrass distribution within Southern Humboldt Bay. Traditional methods used to survey eelgrass are often labor intensive, and usually involve destructive sampling. Here I demonstrate a novel approach to intertidal surveying using unmanned aerial systems. The described method allows for speedy and simple implementation, while also allowing for adaptation to difficult environmental and tidal conditions.

With the threat of human development (Gao et al., 2017), eelgrass wasting disease (Short, 2014) and other disturbances (Unsworth et al., 2015), I believe that this methodology will allow for more consistent monitoring of these critical environments. These methods can provide high resolution time series remote sensing imagery that will allow for more detailed fine scale change analysis of eelgrass and other intertidal ecosystems, which are, demonstrated here, as useful to resource managers. The value of these methods is also made clear by the ability to obtain imagery at low tide cycles to better evaluate eelgrass meadows and reduce confusion in eelgrass models from water column reflectance (Nesbit, 2018). Eelgrass meadows are known to be dynamic, making these systems difficult to monitor (Cunha et al., 2005). UAVs have given researchers the ability to utilize remotely sensed data at a finer scale that may allow for more analysis

within eelgrass beds (Duffy et al., 2018). For this reason, the methods described here may be well suited to long term monitoring where eelgrass surveys can be performed more frequently and consistently (Digruttolo & Mohamed, 2010). While some analyses in this study show there is good chance UAV monitoring is a practical option for eelgrass bed sampling, it is also clear there are several obstacles that need to be considered when conducting UAV monitoring in the intertidal zone of Humboldt Bay, California. In conclusion, I believe it is necessary to further evaluate these methods and define a clearer definition of protocols for eelgrass monitoring with UAVs. These protocols should be focused on defining optimal tidal heights for intertidal and subtidal eelgrass remote sensing, determining the optimal flight heights to obtain the proper resolution for the species of concern, and evaluating the frequency at which remote sensing data should be collected. All of the above factors proved important in modeling eelgrass with UAV remotely sensed data.

Napa Valley Modeling Conclusions

The goal of this study was to create a land cover classification model of the Napa Valley Watershed that included land cover types such as vineyard land use, forests, grasslands, oak dominant savanna and riparian. The classification methods in this study used seven different feature classes based on unique wildlife habitat needs. The definitions for classification were defined above and utilize a collection of previously defined classification schemes. Of particular interest, the different vegetative classifications were difficult to distinguish. The savanna definition utilized in this

analysis created a confused feature class with unique forest structure of sparse to mixed canopy cover and grassland understory. With higher resolution datasets, such as NAIP 1-meter imagery and 1-meter lidar data, classification of unique vegetation classes was much clearer.

Due to the available geospatial and remote sensing data, I focused on the fusion of multiple data sources to classify these habitat land cover types. To this end, I used low density lidar data, airborne multispectral imagery, and high-resolution vector data to create a high-resolution habitat classification. One of the primary conclusions from the analysis was that by using this data fusion method, we increased overall accuracies by 18.74% compared to the National Land Cover Dataset. This study demonstrably classified certain habitat land cover types of which were otherwise not mapped in other LULC maps, such as the National Land Cover Dataset. Aside from overall accuracy of the land cover model, I was also able to obtain a resolution that is 30 times finer than any other available land cover model. This study defined a novel data fusion workflow to create a high-resolution habitat map for the Napa Watershed. Going forward, these methods can be utilized to develop and update existing land cover maps using newer data and other expert knowledge. While I have developed a specific set of rule-based classifications for oak dominant savanna, these parameters have the potential to be adapted to site specific habitat characteristics. These data have provided useful information for modeling natural phenomenon in the Napa Valley and will continue to be a useful resource in wildlife and natural resource management.

REFERENCES OR LITERATURE CITED

Al-Wassai, F. A., & Kalyankar, N. V. (2013). Major limitations of satellite images. *ArXiv Preprint ArXiv:1307.2434*.

Anthony, G., Greg, H., & Tshilidzi, M. (2009). *Classification of Images Using Support Vector Machines*. https://arxiv.org/ftp/arxiv/papers/0709/0709.3967.pdf

Augyte, S., & Simona. (2011). *A floristic analysis of the marine algae and seagrasses between Cape Mendocino, California and Cape Blanco, Oregon*. http://humboldt-dspace.calstate.edu/handle/2148/851

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(1), 2–16. https://doi.org/https://doi.org/10.1016/j.isprsjprs.2009.06.004

Brown, L. M., Fuda, R. K., Schtickzelle, N., Coffman, H., Jost, A., Kazberouk, A., Kemper, E., Sass, E., & Crone, E. E. (2017). Using animal movement behavior to categorize land cover and predict consequences for connectivity and patch residence times. *Landscape Ecology*, *32*(8), 1657–1670.

Buiten, H. J., & Clevers, J. G. P. W. (1993). *Land observation by remote sensing: theory and applications*. Yverdon, Switzerland.

Bujan, S., Gonzalez-Ferreiro, E., Barreiro-Fernandez, L., Sante, I., Corbelle, D., & Miranda, D. (2013). International journal of remote sensing.: Classification of rural landscapes from low-density lidar data: is it theoretically possible? *International Journal of Remote Sensing*, *34*(16), 5666–5689.

Carter-Griffin, C., Hubauer, C., Minks, A., & Robinson, E. (2010). *A pre-feasibility study examining oyster mariculture expansion in Humboldt Bay, California*. http://gsp.humboldt.edu/Websites/SLR/reports/Practicum_Mariculture_Final_2011 (2013).pdf

Casado, M. R. (2019). *Novel Advances in Aquatic Vegetation Monitoring in Ocean, Lakes and Rivers*. MDPI.

Cawley, G. C., Talbot, N. L. C., Guyon, I., & Saffari, A. (2007). Preventing Over-Fitting during Model Selection via Bayesian Regularisation of the Hyper-Parameters. *Journal of Machine Learning Research*, *8*, 841–861.

Clemmensen, L., Hastie, T., Witten, D., & Ersbøll, B. (2011). *Sparse Discriminant Analysis*. https://web.stanford.edu/~hastie/Papers/sda_resubm_daniela-final.pdf

Clinton, P. J., Young, D. R., Specht, D. T., & Lee II, H. (2007). *A Guide to Mapping Intertidal Eelgrass and Nonvegetated Habitats in Estuaries of the Pacific Northwest USA*.

Colgan, M., Baldeck, C., Féret, J.-B., Asner, G., Colgan, M. S., Baldeck, C. A., Féret, J.-B., & Asner, G. P. (2012). Mapping Savanna Tree Species at Ecosystem Scales Using Support Vector Machine Classification and BRDF Correction on Airborne Hyperspectral and LiDAR Data. *Remote Sensing*, *4*(11), 3462–3480. https://doi.org/10.3390/rs4113462

Congalton, R. G., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: principles and practices* (K. Green (ed.)). Boca Raton : Lewis Publications.

Cortes, C., & Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Cracknell, A. P. (1998). Review article Synergy in remote sensing-what's in a pixel? *International Journal of Remote Sensing*, *19*(11), 2025–2047.

Cunha, A. H., Santos, R. P., Gaspar, A. P., & Bairros, M. F. (2005). Seagrass landscape-scale changes in response to disturbance created by the dynamics of barrier-islands: A case study from Ria Formosa (Southern Portugal). *Estuarine, Coastal and Shelf Science*, *64*(4), 636–644. https://doi.org/10.1016/j.ecss.2005.03.018

Cunningham, M. A. (2006). Accuracy assessment of digitized and classified land cover data for wildlife habitat. *Landscape and Urban Planning*, *78*(3), 217–228. https://doi.org/10.1016/j.landurbplan.2005.08.002

Darvishzadeh, R., Atzberger, C., & Skidmore, A. K. (2006). Hyperspectral vegetation indices for estimation of leaf area index. *Proceedings of the ISPRS Commission VII Symposium 'Remote Sensing: From Pixels to Processes*, 6. http://www.isprs.org/proceedings/XXXVI/part7/PDF/233.pdf

Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, *1*(1–4), 131–156.

Dean, T. A., & Bodkin, J. L. (2016). *Eelgrass (Zostera marina) Bed Sampling*.

Digruttolo, N., & Mohamed, A. (2010). *Emerging Unmanned Aerial Remote Sensing System for Intertidal Zone Modeling : A Low-Cost Method of Collecting Remote Sensing Data for Modeling Short-Term Effects of Sea Level Rise , Part II : Close-Range Airborne Remote Sensing Nicholas DiGruttolo and Ahm*. *3*, 119–129.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., & others. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46.

Duarte, C. M., & Chiscano, C. L. (1999). Seagrass biomass and production: a reassessment. In *Aquatic Botany* (Vol. 65). https://ac-els-cdn-com.ezproxy.humboldt.edu/S0304377099000388/1-s2.0-S0304377099000388-main.pdf?_tid=e0ee7a5e-ecdf-4f75-ae64-c4baa8cdf9e5&acdnat=1540431943_e7c65feef88e5ce2e9653f350d07a68d

Duffy, J. P., Pratt, L., Anderson, K., Land, P. E., & Shutler, J. D. (2018). Spatial assessment of intertidal seagrass meadows using optical imaging systems and a lightweight drone. *Estuarine, Coastal and Shelf Science*, *200*, 169–180. https://doi.org/10.1016/j.ecss.2017.11.001

Ebrahimi, M. A., Khoshtaghaza, M. H., Minaei, S., & Jamshidi, B. (2017). Vision-based pest detection based on SVM classification method. *Computers and Electronics in Agriculture*, *137*(October), 52–58. https://doi.org/10.1016/j.compag.2017.03.016

Elford, R. C. (1963). *[Climate of California counties]*. Place of publication not identified : publisher not identified.

Flather, C. H., Knowles, M. S., & Kendall, I. A. (1998). Threatened and Endangered

Species Geography. *BioScience*, *48*(5), 365–376. https://doi.org/10.1525/bio.2010.60.10.17

Gao, Y., Fang, J., Du, M., Fang, J., Jiang, W., & Jiang, Z. (2017). Response of the eelgrass (Zostera marina L.) to the combined effects of high temperatures and the herbicide, atrazine. *Aquatic Botany*, *142*, 41–47. https://doi.org/10.1016/j.aquabot.2017.06.005

Gebejes, A., & Huertas, R. (2013). Texture characterization based on grey-level co-occurrence matrix. *Databases*, *9*, 10.

Gilkerson, W. (2008). *A spatial model of eelgrass (Zostera marina) habitat in Humboldt Bay, California*. http://humboldt-dspace.calstate.edu/handle/2148/378

Graham, R., & Read, R. E. (1986). *Manual of Aerial Photography*. Focal Press. https://books.google.com/books?id=4359ZQpqOjwC

Green, K., Kempka, D., & Lackey, L. (1994). Using remote sensing to detect and monitor land-cover and land-use change. *Photogrammetric Engineering and Remote Sensing*, *60*(3), 331–337.

Grossinger, R., Beller, E., Salomon, M., Whipple, A., Askevold, R., Striplen, C., Brewster, E., & Leidy, R. (2008). *Final Report: South Santa Clara Valley Historical Ecology Study*.

Han, H., & Jiang, X. (2014). Overcome support vector machine diagnosis overfitting. *Cancer Informatics*, *13*(Suppl 1), 145–158. https://doi.org/10.4137/CIN.S13875

Hargrove, W. W., Hoffman, F. M., & Efroymson, R. A. (2005). A practical map-analysis tool for detecting potential dispersal corridors. *Landscape Ecology*, *20*(4), 361–373.

Harrington, P. de B. (2015). Support Vector Machine Classification Trees. *Analytical Chemistry*, *87*(21), 11065–11071. https://doi.org/10.1021/acs.analchem.5b03113

Hollister, J. W., Gonzalez, M. L., Paul, J. F., August, P. V., & Copeland, J. L. (2004). Assessing the accuracy of national land cover dataset area estimates at multiple spatial extents. *Photogrammetric Engineering and Remote Sensing*, *70*(4), 405–414. https://doi.org/10.14358/PERS.70.4.405

Homer, C., Dewitz, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., & Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States--representing a decade of land cover change information. *Photogrammetric Engineering \& Remote Sensing*, *81*(5), 345–354.

Hsu, C.-W., Chang, C.-C., & Lin, C.-J. (2003). *A Practical Guide to Support Vector Classification*. http://www.csie.ntu.edu.tw/~cjlin

Hu, T., Yang, J., Li, X., & Gong, P. (2016). Mapping urban land use by using landsat images and open social data. *Remote Sensing*, *8*(2). https://doi.org/10.3390/rs8020151

Huysman, A., & Johnson, M. (2021). Multi- year nest box occupancy and short- term resilience to wildfire disturbance by barn owls in a vineyard agroecosystem. *Ecosphere*, *12*. https://doi.org/10.1002/ecs2.3438

Jed Wing, M. K. C., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., the R Core Team, Benesty, M., Lescarbeau, R., Ziem, A.,

Scrucca, L., Tang, Y., Candan, C., & Hunt., T. (2019). *caret: Classification and Regression Training*. https://cran.r-project.org/package=caret

Jensen, J R. (1996). Introduction to digital image processing of remotely sensed data. *Introductory Digital Image Processing: A Remote Sensing Perspective. Prentice-Hall, Inc., New Jersey*, 1–2.

Jensen, John R. (2007). *Remote sensing of the environment: an earth resource perspective* (2nd ed.). Upper Saddle River, NJ : Pearson Prentice Hall.

Judd, C., Steinberg, S., Shaughnessy, F., & Crawford, G. (2007). Mapping salt marsh vegetation using aerial hyperspectral imagery and linear unmixing in Humboldt Bay, California. *Wetlands*, *27*, 1144–1152. https://doi.org/10.1672/0277-5212(2007)27[1144:MSMVUA]2.0.CO;2

Karlik, J. F., & Chojnacky, D. C. (2014). Biomass and carbon data from blue oaks in a California oak savanna. *Biomass and Bioenergy*, *62*, 228–232.

Kennedy, R. E., Yang, Z., Braaten, J., Copass, C., Antonova, N., Jordan, C., & Nelson, P. (2015). Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA. *Remote Sensing of Environment*, *166*, 271–285. https://doi.org/10.1016/j.rse.2015.05.005

Klemas, V. (2016). Airborne Remote Sensing of Coastal Features and Processes : An Overview. *Journal of Coastal Research*, *29*(2), 239–255.

Kottner, J. (2009). Interrater reliability and the kappa statistic: A comment on Morris et al. (2008). *International Journal of Nursing Studies*, *46*(1), 141–142. https://doi.org/10.1016/j.ijnurstu.2008.04.001

Law, J. R., Johnson, P. S., & Houf, G. (1994). *A crown cover chart for oak savannas. Tech. Bull*(1), 4 p.

Leutner, B., Horning, N., Schwalb-Willmann, J., & Hijmans, R. J. (2017). RStoolbox: tools for remote sensing data analysis. *R Package Version 0.1*, *7*.

Li, W., Fu, H., Yu, L., & Cracknell, A. (2017). Deep Learning Based Oil Palm Tree Detection and Counting for High-Resolution Remote Sensing Images. *Remote Sensing*, *9*(1). https://doi.org/10.3390/rs9010022

Lillesand, T. M., Kiefer, R. W., & Chipman, J. W. (2008). *Remote Sensing and Image Interpretation* (Sixth Edit). John Wiley & Sons, Inc.

Lin, W., Orgun, M. A., Williams, G. J., & others. (2002). An Overview Of Temporal Data Mining. *AusDM*, 83–90.

Liu, X. (2008). Airborne LiDAR for DEM generation: some critical issues. *Progress in Physical Geography*, *32*(1), 31–49.

Long, W., & Srihann, S. (2004). Land cover classification of SSC image: unsupervised and supervised classification using ERDAS Imagine. *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium*, *4*, 2707–2712.

Mahall, B. E., Davis, F. W., & Tyler, C. M. (2005). Santa Barbara County Oak Restoration Program, August 1994-August 2005: Final Report. *Prepared for: County of Santa Barbara Department of Planning and Development Energy Division*.

Marceau, D. J., Howarth, P. J., Dubois, J.-M. M., Gratton, D. J., & others. (1990).

Evaluation of the grey-level co-occurrence matrix method for land-cover classification using SPOT imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(4), 513–519.

Matt Lamborn. (2010). *Napa County Agriculture*. Napa County Planning, Building, Environmental Health Services. http://gis.napa.ca.gov/giscatalog/catalog_xml.asp?srch_opt=all&db_name=x&theme=x&sort_order=layer&meta_style=fgdc&submit=Submit

Meehan, A. J., Williams, R. J., & Watford, F. A. (2005). Detecting trends in seagrass abundance using aerial photograph interpretation: problems arising with the evolution of mapping methods. *Estuaries*, *28*(3), 462–472.

Merkel, K. W., & Consultant, P. (2017). *Humboldt Bay Eelgrass Comprehensive Management Plan*.

Minnesota, C. (2010). *Appendix L : Oak Savanna Definition Working Definition of " Savanna " for Restoration Efforts at Crane Meadows General Definition of Southern Dry Savanna :* 197–200.

Moody, A., & Woodcock, C. E. (1995). The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing. *Landscape Ecology*, *10*(6), 363–379.

Morgan, J. L., Gergel, S. E., & Coops, N. C. (2010). Aerial Photography: A Rapidly Evolving Tool for Ecological Management. *BioScience*, *60*(1), 47–59. https://doi.org/10.1525/bio.2010.60.1.9

Nesbit, S. (2018). Mapping Subtidal Eelgrass Beds using Unmanned Aerial Systems (UAS). *American Association of Geographers Annual Meeting*. https://aag.secure-abstracts.com/AAG Annual Meeting 2018/abstracts-gallery/13509

NOAA. (2019). *NOAA Tide Predictions*. 1–5.

Nowak, E., Jurie, F., & Triggs, B. (2006). Sampling Strategies for Bag-of-Features Image Classification. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision -- ECCV 2006* (pp. 490–503). Springer Berlin Heidelberg.

O'Neill, J. D., & Costa, M. (2013). Mapping eelgrass (Zostera marina) in the Gulf Islands National Park Reserve of Canada using high spatial resolution satellite and airborne imagery. *Remote Sensing of Environment*, *133*, 152–167. https://doi.org/10.1016/j.rse.2013.02.010

O'Neill, J. D., Costa, M., Sharma, T., O'Neill, J. D., Costa, M., & Sharma, T. (2011). Remote Sensing of Shallow Coastal Benthic Substrates: In situ Spectra and Mapping of Eelgrass (Zostera marina) in the Gulf Islands National Park Reserve of Canada. *Remote Sensing*, *3*(5), 975–1005. https://doi.org/10.3390/rs3050975

Palamuleni, L., Annegarn, H., Kneen, M., & Landmann, T. (2007). Mapping rural savanna woodlands in Malawi: A comparison of maximum likelihood and fuzzy classifiers. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 1260–1264. https://doi.org/10.1109/IGARSS.2007.4423035

Parikh, R., Mathai, A., Parikh, S., Sekhar, G. C., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology*, *56*(1), 45.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2636062/

Perumal, K., & Bhaskaran, R. (2010). Supervised classification performance of multispectral images. *ArXiv Preprint ArXiv:1002.4046*.

Pham, L. T. H., Brabyn, L., & Ashraf, S. (2016). Combining QuickBird, LiDAR, and GIS topography indices to identify a single native tree species in a complex landscape using an object-based classification approach. *International Journal of Applied Earth Observation and Geoinformation*, *50*, 187–197. https://doi.org/10.1016/j.jag.2016.03.015

Phillips, R. C., & Watson, J. F. (1984). *The ecology of eelgrass meadows in the Pacific northwest: A community profile. U.S. Fish and Wildlife Service. FWS/OBS-84/24*. 85.

Piñeiro, G., Perelman, S., Guerschman, J., & Paruelo, J. (2008). How to Evaluate Models: Observed vs. Predicted or Predicted vs. Observed? *Ecological Modelling*, *216*, 316–322. https://doi.org/10.1016/j.ecolmodel.2008.05.006

Plummer, S. E. (2000). Perspectives on combining ecological process models and remotely sensed data. *Ecological Modelling*. https://doi.org/10.1016/S0304-3800(00)00233-7

R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. https://www.r-project.org/

Radloff, P. L., Hobson, C., Whisenant, A., & Bronson Warren, J. M. (2013). *Statewide Seagrass Monitoring Protocol Development - Phase 2 Final Report*.

Remesan, R., & Mathew, J. (2014). *Hydrological data driven modelling : a case study approach* (Illustrate). Springer. https://books.google.com/books?id=lessBQAAQBAJ&dq=Monte-Carlo+Cross+Validation+vs+Repeated+random+sub-sampling+validation&source=gbs_navlinks_s

Reshitnyk, L., Costa, M., Robinson, C., & Dearden, P. (2014). Evaluation of WorldView-2 and acoustic remote sensing for mapping benthic habitats in temperate coastal Pacific waters. *Remote Sensing of Environment*, *153*, 7–23. https://doi.org/10.1016/j.rse.2014.07.016

Robinson, D. A., Lebron, I., & Querejeta, J. I. (2010). Determining Soil--Tree--Grass Relationships in a California Oak Savanna Using Eco-Geophysics. *Vadose Zone Journal*, *9*(3), 528–536.

Robinson, J., Brush, S., Douglas, I., Graedel, T. E., Graetz, D., Hodge, W., Liverman, D., Melillo, J., Moss, R., Naumov, A., & others. (1994). Land-use and land-cover projections: Report of working Group C. *Changes in Land Use and Land Cover: A Global Perspective.(Eds.) W. Meyer and BL Turner II. Cambridge University Press*, 73–92.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., & Freden, S. C. (1973). Monitoring vegetation systems in the Great Plains with ERTS.(pp. 309--317). *Proceedings of 3rd Earth Resources Technology Satellite-1 Symposium*, *1*.

Rowan, G. S. L., & Kalacska, M. (2021). A Review of Remote Sensing of Submerged Aquatic Vegetation for Non-Specialists. *Remote Sensing*, *13*(4), 623.

Schlosser, S., & Eicher, A. (2012). *Humboldt Bay and Eel River Estuary Benthic Habitat Project*.

Shaughnessy, F. J., Gilkerson, W., Black, J. M., Ward, D. H., & Petrie, M. (2012). Predicted eelgrass response to sea level rise and its availability to foraging Black Brant in Pacific coast estuaries. In *Source: Ecological Applications* (Vol. 22, Issue 6). https://www-jstor-org.ezproxy.humboldt.edu/stable/pdf/41722890.pdf?refreqid=excelsior%3A419b2ab37ff0524d3db76586d1d85559

Shi, H., Laurent, E. J., LeBouton, J., Racevskis, L., Hall, K. R., Donovan, M., Doepker, R. V, Walters, M. B., Lupi, F., & Liu, J. (2006). Local spatial modeling of white-tailed deer distribution. *Ecological Modelling*, *190*(1–2), 171–189.

Short, F. (2014). *Western CEDAR Eelgrass Wasting Disease: an Overview*. https://cedar.wwu.edu/ssec

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012). Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research*, *50*(9), 1250–1265. https://doi.org/10.1016/j.asr.2012.06.032

Suribabu, C. R., Bhaskar, J., & Neelakantan, T. R. (2012). Land Use/Cover Change Detection of Tiruchirapalli City, India, Using Integrated Remote Sensing and GIS Tools. *Journal of the Indian Society of Remote Sensing*, *40*(4), 699–708. https://doi.org/10.1007/s12524-011-0196-x

Tesfaw, A. T., Pfaff, A., Golden Kroner, R. E., Qin, S., Medeiros, R., & Mascia, M. B. (2018). Land-use and land-cover change shape the sustainability and impacts of protected areas. *Proceedings of the National Academy of Sciences*, *115*(9), 2084–2089. https://doi.org/10.1073/pnas.1716462115

U.S., G. S. (2009). *National Hydrography Dataset*. [Reston, Va.] : U.S. Dept. of the Interior, U.S. Geological Survey, 2004-. https://search.library.wisc.edu/catalog/9910061259502121

Unsworth, R. K. F., Collier, C. J., Waycott, M., Mckenzie, L. J., & Cullen-Unsworth, L. C. (2015). *A framework for the resilience of seagrass ecosystems*. https://doi.org/10.1016/j.marpolbul.2015.08.016

Vapnik, V. N., & Lerner, A. Y. (1963). Recognition of Patterns with help of Generalized Portraits. In *Avtomat. i Telemekh* (Vol. 24, Issue 6). http://www.mathnet.ru/eng/agreement

Verostek, J. (2014). *Evaluating Model Performance*. http://www.johnverostek.com/wp-content/uploads/2014/06/Chapter-10.pdf

Villanueva-Rivera, L. J., Pijanowski, B. C., Doucette, J., & Pekin, B. (2011). A primer of acoustic analysis for landscape ecologists. *Landscape Ecology*, *26*(9), 1233. https://doi.org/10.1007/s10980-011-9636-9

Vogelmann, J. E., Sohl, T. L., Campbell, P. V, & Shaw, D. M. (1998). Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. *Environmental Monitoring and Assessment*, *51*(1–2), 415–428.

Waycott, M., Duarte, C. M., Carruthers, T. J. B., Orth, R. J., Dennison, W. C., Olyarnik, S., Calladine, A., Fourqurean, J. W., Heck, K. L., Hughes, A. R., Kendrick, G. A.,

Kenworthy, W. J., Short, F. T., & Williams, S. L. (2009). Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(30), 12377–12381. https://doi.org/10.1073/pnas.0905620106

Wendt, C. (2016). *EXAMINING BARN OWL NEST BOX SELECTION AT THREE SPATIAL SCALES ON NAPA VALLEY VINEYARDS* [Humboldt State University]. http://humboldt-dspace.calstate.edu/bitstream/handle/10211.3/177093/Wendt_Carrie_Su2016.pdf?sequence=1

West, C. (2007). *A Comparison of High Spatial Resolution Images for Fine Scale Vegetation Mapping*. Humboldt State University.

Whippo, R., Knight, N. S., Prentice, C., Cristiani, J., Siegle, M. R., & O'Connor, M. I. (2018). Epifaunal diversity patterns within and among seagrass meadows suggest landscape-scale biodiversity processes. *Ecosphere*, *9*(11), e02490.

Whitehead, K., & Hugenholtz, C. (2014). Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges. *Journal of Unmanned Vehicle Systems*, *02*, 69–85. https://doi.org/10.1139/juvs-2014-0006

Whiteside, T. G., Boggs, G. S., & Maier, S. W. (2011). Comparing object-based and pixel-based classifications for mapping savannas. *International Journal of Applied Earth Observation and Geoinformation*. https://doi.org/10.1016/j.jag.2011.06.008

Wolter, P. T., Berkley, E. A., Peckham, S. D., & Singh, A. (2014). Satellite-based management tool for oak savanna ecosystem restoration. *Journal of Fish and Wildlife Management*, *5*(2), 252–269. https://doi.org/10.3996/022013-JFWM-010

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. https://doi.org/10.1016/j.rse.2012.01.010

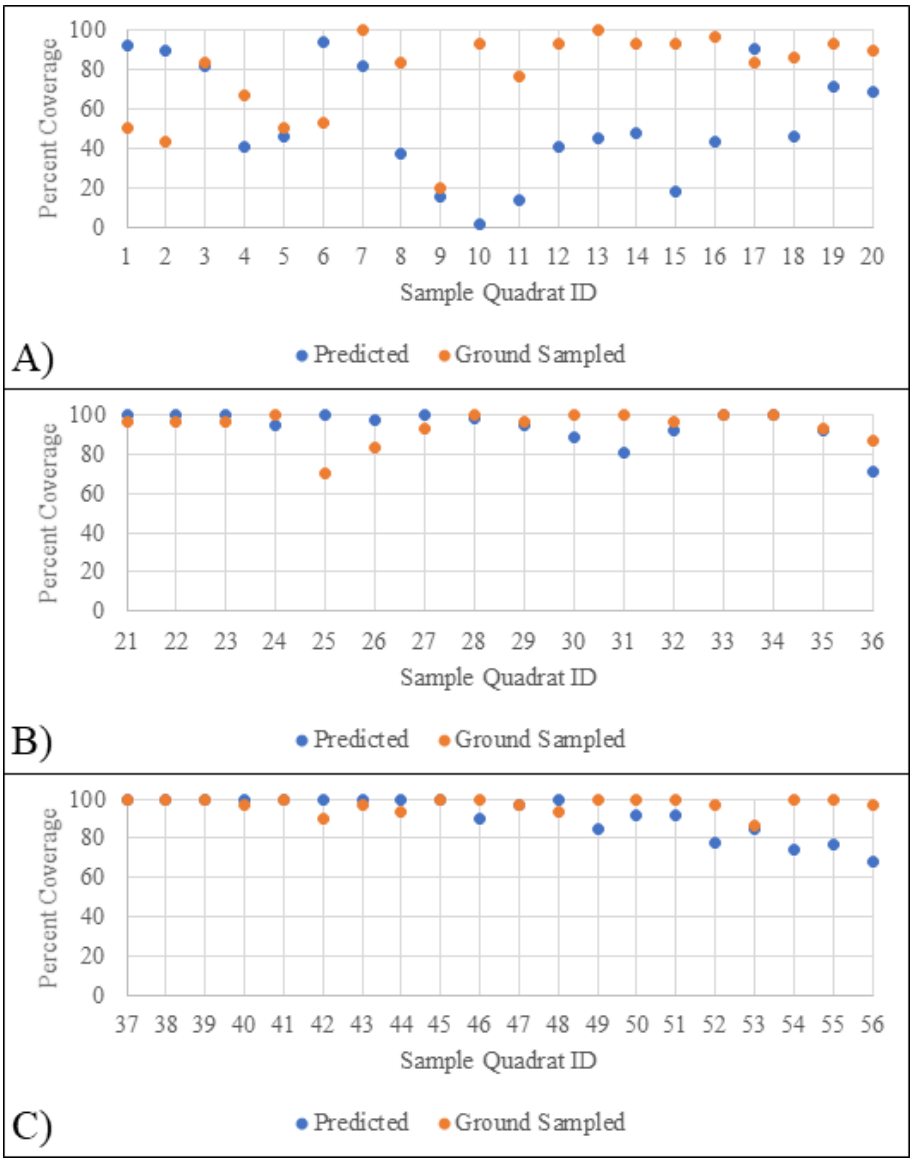
Xu, S., Xu, S., Zhou, Y., Yue, S., Qiao, Y., Liu, M., Gu, R., Song, X., Zhang, Y., & Zhang, X. (2020). Sonar and in situ surveys of eelgrass distribution, reproductive effort, and sexual recruitment contribution in a eutrophic bay with intensive human activities: Implication for seagrass conservation. *Marine Pollution Bulletin*, *161*, 111706.

Zhang, Y., Wu, B., & Wang, D. (2013). Research Dynamics of the Classification Methods of Remote Sensing Images - ProQuest. *Asian Agricultural Research*, *5*(3), 118–122. https://search-proquest-com.ezproxy.humboldt.edu/docview/1347640707/citation/8737ACBF6CD84423PQ/1?accountid=11532

Zvoleff, A. (2019). *Package 'glcm.'*

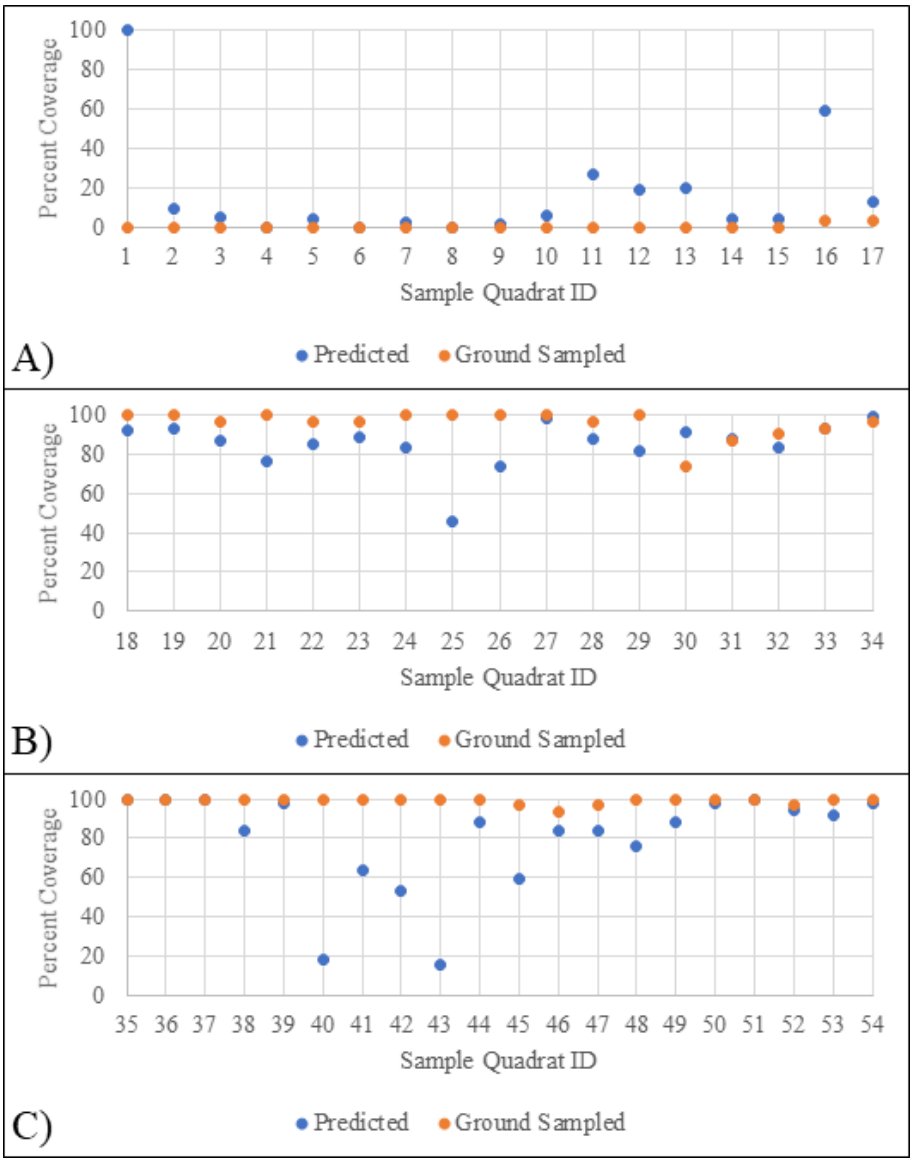
Appendix A

Predicted eelgrass percent coverage is compared with ground sampled percent cover a the SBW site for each sampling transect, including the A) high transect, B) middle transect, and C) low transect.



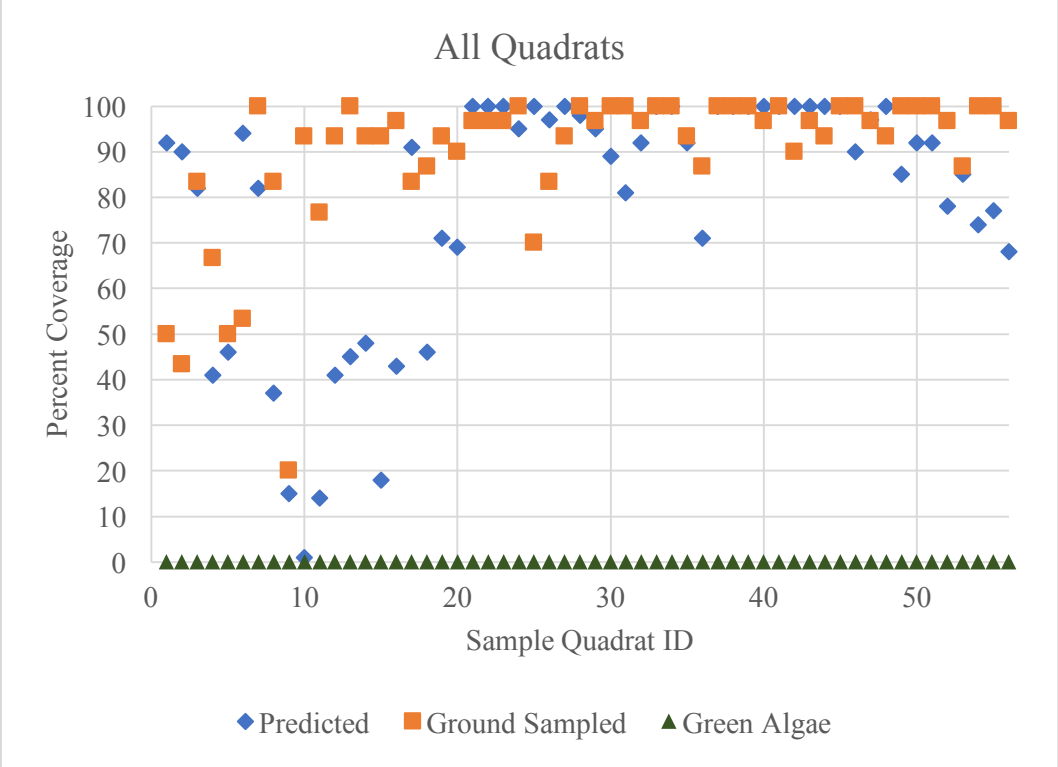
Appendix B

Predicted eelgrass percent coverage is compared with ground sampled percent cover a the MPA site for each sampling transect, including the A) high transect, B) middle transect, and C) low transect.



Appendix C

Ground sampled percent cover for eelgrass (orange, square) and green algae (green, triangle) are compared to the predicted percent cover of eelgrass (blue, diamond) at the SBW site, showing no green algae coverage at the SBW site.



Appendix D

Ground sampled percent cover for eelgrass (orange, square) and green algae (green, triangle) are compared to the predicted percent cover of eelgrass (blue, diamond) at the MPA site, showing green algae coverage is only present at the high tidal transect quadrats.