Article

UAV-Based Wetland Monitoring: Multispectral and Lidar Fusion with Random Forest Classification

Robert Van Alphen *, Kai C. Rains (), Mel Rodgers () and Rocco Malservisi () and Timothy H. Dixon ()

School of Geosciences, University of South Florida, Tampa, FL 33620, USA; krains@usf.edu (K.C.R.);
melrodgers@usf.edu (M.R.); rocco@usf.edu (R.M.); thd@usf.edu (T.H.D.)
* Correspondence: rvanalphen@usf.edu

Abstract: As sea levels rise and temperatures increase, vegetation communities in tropical and sub-tropical coastal areas will be stressed; some will migrate northward and inland. The transition from coastal marshes and scrub–shrubs to woody mangroves is a fundamental change to coastal community structure and species composition. However, this transition will likely be episodic, complicating monitoring efforts, as mangrove advances are countered by dieback from increasingly impactful storms. Coastal habitat monitoring has traditionally been conducted through satellite and ground-based surveys. Here we investigate the use of UAV-LiDAR (unoccupied aerial vehicle–light detection and ranging) and multispectral photogrammetry to study a Florida coastal wetland. These data have higher resolution than satellite-derived data and are cheaper and faster to collect compared to crewed aircraft or ground surveys. We detected significant canopy change in the period between our survey (2020–2022) and a previous survey (2015), including loss at the scale of individual buttonwood trees (Conocarpus erectus), a woody mangrove associate. The UAV-derived data were collected to investigate the utility of simplified processing and data inputs for habitat classification and were validated with standard metrics and additional ground truth. UAV surveys combined with machine learning can streamline coastal habitat monitoring, facilitating repeat surveys to assess the effects of climate change and other change agents.

Keywords: UAV-LiDAR; mangrove; machine learning

1. Introduction

Coastal wetlands and associated habitats provide many environmental services and economic benefits to coastal communities, including carbon sequestration, protection from storm surges and other erosive forces, and maintenance of biodiversity [1,2]. These valuable coastal environments are threatened by climate-induced changes to sea level, local climate, and storms. They require study and monitoring to help ensure their preservation. Coastal mangroves are common along tropical and subtropical coasts, ranging in stature from tall trees exceeding 50 m in the tropics to shrubs at the northern extent of their range [3]. Rising sea levels are contributing to their landward migration along waterways, while ocean warming is providing new habitats to the north [4,5]. Northward migration of storm tracks in the northern hemisphere combined with increasing storm intensity challenge this expansion [6]. Growth and dieback are a dynamic process, suggesting the need for monitoring. This migration will encounter anthropogenic and landscape barriers, causing a ‘coastal squeeze’, potentially reducing the extent of shoreline suitable for mangroves and requiring management in near real-time [7]. Effective monitoring of coastal forests will be key to this process.

Traditionally, wetland monitoring has relied on ground surveys that allow for a detailed understanding of the study site. However, this approach can be labor intensive, time consuming, and possibly dangerous. In response to these issues, studies taking place in the past 30 years have moved increasingly toward larger-scale airborne and satellite-based
monitoring efforts. Primarily, these efforts have been directed at wetland classification and change detection [8]. These authors also found that, in its current state, satellite-based sensing has been most successful with both convolutional neural networks (CNN) for deep learning and random forest models for machine learning. These methods are not without their challenges. Some challenges to CNN are the minimum sampling area needed for mangrove detection and spectral confusion with species common to mangrove swamps [9,10]. For random forest methods, studies have shown better performance when optical- and radar-based data are used as model inputs [11,12]. Reference [13] describes wetlands as a ‘moving target’ based on large variability between environments. This variability exists between locations and even between seasons at a single location. These and related challenges are discussed by [14,15].

An alternative approach is the use of unoccupied aerial vehicles (UAVs), which can operate at a moderate cost, facilitating repeat surveys [16,17]. UAV systems can also generate accurate digital elevation models (DEM) using LiDAR or structure from motion (SfM) [18], with higher spatial resolution compared to satellite-based data [19]. SfM is adequate for forest height and density monitoring in mangrove-dominant environments [20] but is challenged where bare earth elevation is required and vegetation is dense.

Reference [21] reviewed the current state of UAVs in wetland applications, reporting on strengths, weaknesses, and emerging trends. Among the strengths were the many ways a UAV survey can be customized. UAV-derived data have been used in biomass calculations, vegetation mapping, and four-dimensional (3D + time) change detection [22–24]. UAVs can be deployed at nearly any time, limited mainly by regulations and weather, and can be equipped with multiple sensor types for increased data gathering, with UAV-LiDAR becoming more common. However, some studies limited reporting on accuracy—for example, the number and quality of ground control points. In lieu of ground control points, real time kinematic (RTK) surveys are playing a larger role. However, processing the high-resolution data acquired by this instrumentation has proven challenging: the accuracy of RTK may be much less than the pixel resolution of the survey. At the pixel scale, heterogeneities in spectral signatures can make classification and delineation challenging. In response to this, some studies have preferred an object-based approach. This may then be complicated by the requirement to fine-tune the segmentation algorithms with appropriate variables. Reference [21] reviews these and related issues.

One emerging trend is to use UAV surveys to augment and scale up to satellite-based regional scales [25,26]. Hyperspectral data are also becoming more widely available [27–30], albeit with data processing challenges compared to simpler multispectral data with fewer bands. Machine learning has also become a tool used in many classification efforts. In part, these algorithms are employed to deal with the large number of descriptors that can be generated by UAV surveys, including UAV-LiDAR data [27,29–31].

In this study, we aim to use UAV-based methods with simplified data collection and processing protocols for wetland surveyance and classification. Our protocols highlight the utility of first-order data to monitoring efforts. We use high spatial-resolution multi-spectral data with five spectral bands, high-resolution ground elevation data derived from LiDAR, and a simple random forest machine learning classification algorithm.

2. Materials and Methods

2.1. Study Area

The study site (27.618640° North, −82.734852° East) is situated within Fort Desoto County Park (Mullet Key), along the Gulf Coast of Florida, south of the City of St. Petersburg (Figure 1). The park includes recreational facilities, paved roads and bike paths, and native plant communities typical of Gulf Coast subtropical barrier islands. A dense mangrove forest occupies the saltwater lagoon’s shoreline and forms the study site’s eastern border.

The study site is ~0.055 km² in area and includes one of twelve transects routinely ground surveyed by the Tampa Bay Estuary Program (TBEP), an EPA-funded activity
that monitors the ecological health of Tampa Bay and its watershed. The surveys assess species diversity, ecological health, habitat loss, changes related to sea level rise, and other environmental stressors (https://tbep.org/ accessed on 6 January 2020). The ground-based surveys are conducted by skilled personnel and are typically repeated every one to three years at a considerable cost. Habitat restoration activities are recommended where appropriate.

Figure 1. Study site and the surrounding landscape. (A): Map of Florida (inset) and the greater Tampa Bay area. The location of the study site is indicated by a red star. (B): Aerial photograph of the study area. The training site is indicated by the solid green box. The testing site is indicated by the dashed blue polygon. Validation survey sites in the testing site are indicated by red points.

2.2. Overview

The general workflow is summarized in Figure 2 and detailed in Figures S1 and S2 in Supplementary Materials. The overall study site contains two sections: the training site (0.016 km$^2$) to the south and the test site (0.03 km$^2$) to the north (Figure 1B). At both sites, we collected elevation data using UAV-LiDAR and ground truth with GNSS ground control and collected imagery using UAV-photogrammetry. The photogrammetry included RGB images, which were used to create orthomosaic maps [32], and multispectral images, which were used to create single-band rasters [33]. We compared UAV-LiDAR elevation data to legacy data through a digital elevation model (DEM) of difference to detect changes in elevation, including canopy height, over time. Our machine learning model was trained with a multi-band raster generated by merging five single-color bands, a normalized difference vegetation index (NDVI) [34] raster derived from the photogrammetric processing, and a digital terrain model (DTM) from the UAV-LiDAR. We collected initial vegetation data by conducting a ground survey in the training site and used principal component
analysis to assess our input features’ ability to distinguish different habitats. We applied this model to the test site and validated the habitat classification by conducting additional ground surveys.

2.3. Photogrammetry and LiDAR Surveys

We conducted two UAV surveys at the training site in January 2020 and two at the testing site in February 2021 in order to stay within the same season as the previous survey. The first survey at each site employed a DJI Phantom 4 RTK (P4RTK; DJI, Shenzhen, China) equipped with a 20-megapixel RGB camera. The second survey used a DJI Phantom 4 RTK Multi-Spectral (P4Multi; DJI, Shenzhen, China) system that housed six 2-megapixel cameras and a down-welling light sensor for image calibration, ensuring reflectance stability during the survey. Five were single-band cameras with blue, green, red, 'red-edge,' and near-infrared (near-IR) channels (Table 1), and the sixth camera was a full-color RGB camera. A DJI GNSS base station (DJI, Shenzhen, China) provided in-flight location correction to both RTK-capable UAVs. Both the base station and the rover (UAV) collected raw RINEX GNSS files for post-survey post-processing kinematic (PPK) corrections.

Table 1. Spectral Bands in the P4Multi Camera (Center wavelength and bandwidth in nm).

<table>
<thead>
<tr>
<th>Channel</th>
<th>Center Wavelength and Bandwidth (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue</td>
<td>450 ± 16</td>
</tr>
<tr>
<td>Green</td>
<td>560 ± 16</td>
</tr>
<tr>
<td>Red</td>
<td>650 ± 16</td>
</tr>
<tr>
<td>RedEdge</td>
<td>730 ± 16</td>
</tr>
<tr>
<td>Near-Infrared (NIR)</td>
<td>840 ± 26</td>
</tr>
</tbody>
</table>

In December 2022, again to stay within the winter season, we conducted a third survey using UAV-LiDAR at both field sites in order to create digital terrain and digital surface models (DTM, DSM). For this survey, we used a Hesai Pandar XT32 (Hesai, Shanghai, China; Table 2). This model provides 32-channel mid-IR range scanning lasers with 640,000 pts/s. The sensor has a 360-degree field of view, ±2 cm vertical precision, and a range of up to 120 m. The LiDAR sensor is coupled with a 24-megapixel camera. The camera provides georeferenced images for point cloud colorization.
Table 2. Metadata available for airborne LiDAR source data for 2015 DEMs compared with 2022 UAV-LiDAR.

<table>
<thead>
<tr>
<th>Acquisition Dates</th>
<th>Horizontal Accuracy</th>
<th>Vertical Accuracy</th>
<th>Scanner Wavelength</th>
<th>Scan Angle</th>
<th>Pulse Rate</th>
<th>Points per m²</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 2002</td>
<td>±0.8 m⁻¹</td>
<td>±15 cm⁻¹</td>
<td>ATM II¹ CHARTS system²</td>
<td>0°⁻³</td>
<td>2–10 kHz</td>
<td>0.1–0.2³</td>
</tr>
<tr>
<td>May–June 2006</td>
<td>80 cm at 2 sigma⁻¹</td>
<td>30 cm at 2 sigma⁻¹</td>
<td>CHARTS system (SHOALS-3000 LiDAR )²</td>
<td>0°⁻³</td>
<td>n/a⁻⁵</td>
<td>0.27³</td>
</tr>
<tr>
<td>June–August 2007</td>
<td>±116 cm⁻¹</td>
<td>9 cm RMSE⁻¹</td>
<td>Leica ALS50¹</td>
<td>29°⁻¹</td>
<td>75 kHz; 84.4 kHz⁻¹</td>
<td>1.8³</td>
</tr>
<tr>
<td>May 2015</td>
<td>1 m⁻¹</td>
<td>19.6 cm⁻¹</td>
<td>CZMIL¹</td>
<td>−21–22°⁻³,⁴</td>
<td>10 kHz⁻¹</td>
<td>1–14²³⁴</td>
</tr>
<tr>
<td>June 2015</td>
<td>1 m RMSE⁻¹</td>
<td>9.5 cm RMSE (topographic data only)¹</td>
<td>Leica HawkEye III¹</td>
<td>13–22°⁻³</td>
<td>300 kHz⁻¹</td>
<td>0.1–0.2²³⁴</td>
</tr>
<tr>
<td>UAV-LiDAR, December 2022</td>
<td>±2 cm⁻²</td>
<td>2 cm RMSE</td>
<td>Hesai Pandar XT32</td>
<td>±30°</td>
<td>3.5 MHz⁻¹</td>
<td>459³</td>
</tr>
</tbody>
</table>

¹ Provided by metadata. ² Provided by instrument documentation. ³ Extracted/Calculated from files. ⁴ Range based on tiles in the dataset. ⁵ Pulse rate not found in documentation. Scanner manufacturers in order presented in the table: NASA, Washington, D.C., USA; Teledyne Optech, Toronto, Canada; Leica Geosystems, Heerbrugg, Switzerland; Teledyne Optech, Toronto, Canada; Leica Geosystems, Heerbrugg, Switzerland; Hesai, Shanghai, China.

We flew the UAV 50 m above ground with 50 m line spacing, calibrating the inertial measurement unit (IMU) to improve location measurements, by flying the UAV in a sinusoidal or figure-eight shape after take-off and before landing. We deployed a Trimble R10 GNSS unit as an RTK base station to collect data for PPK location solutions.

2.4. Photogrammetry Processing

Initial processing results revealed the presence of distortions. Further details are provided in supplemental material. We corrected the position of both the P4RTK and P4Multi using PPK techniques. This technique uses the DJI GNSS base position, various GNSS Satellite positional parameters, and the UAV positions obtained in flight and corrects the UAV position information to cm-level accuracy.

To process the P4RTK and P4Multi images, we used standard photogrammetry procedures: feature selection, bundle adjustment, point cloud densification, and final rasterization [17,18,35]. We used Agisoft Metashape software [32] to create DEMs from the RGB images, which were only used for the creation of subsequent orthophotos (Figure 3). We used Pix4D software [33] to process multispectral images, following a predefined workflow provided by the software, with a minor modification of the addition of a step that calibrated the images via corrections from the downwelling light sensor. Pix4D produced a point cloud and raster for each color band. It also created an NDVI raster calculated from the relevant data bands. Rasters for the training site were of 1.6 m resolution and 2 cm resolution for the testing site. The six single-band rasters produced were part of the inputs used for our random forest classification.

2.5. LiDAR Processing

PCMaster is a proprietary software from the LiDAR vendor used for facilitating point cloud creation. PCMaster allows for the filtering of unwanted returns, IMU calibration, and post-processing kinematic (PPK) position correction. PCMaster performs PPK corrections during data upload. We filtered the initial point cloud to exclude points outside a 70-degree field of view. Points were also excluded if they were captured during turns in flight, minimizing errors from UAV tilting. Further cleaning was performed using CloudCompare software [36], which filters noise (e.g., clouds and power lines) and allows for manual point selection and deletion. The final point cloud has a density of 459 points/m²⁻².
To segment the point cloud we used the Cloth Simulation Filter (CSF) plugin provided by CloudCompare [37]. This algorithm divides points into ground and ‘off-ground’ points. Separation is achieved through the simulation of a ‘cloth’ on an upside-down version of the input point cloud. The parameters we used were rigidness (‘Relief,’ 2), cloth resolution (2 m), classification threshold (0.5 m), maximum iterations (500), and slope post-processing (check box selected). The values were defaults except for the slope post-processing.

After separation, we rasterized the ground and original all-points cloud using the lidR and raster packages in R (R foundation for statistical computing, Vienna, Austria). We created a DTM from the ground point cloud and a DSM from the all-point cloud at 10 cm resolution using inverse-distance weighting. From these models, we created canopy height models (CHM) by subtracting the DTM from the DSM. This normalizes all elevations by height from the ground surface. The final DSMs and DTMs were cut to match the training and testing sites. This LiDAR-derived DTM was added to the spectral bands for input into the machine learning classification detailed in Section 2.7.
2.6. Habitat Characterization

We conducted a habitat survey of the training site in order to better understand the site before the creation of a training dataset for random forest classification. We identified representative habitat types through photo interpretation followed by ground surveys along transects oriented approximately east–west and north–south across the training site. At each of the 14 survey points, we obtained photographs, identified characteristic plant species, and described notable structural characteristics, e.g., dead branches, substrate type.

There are six characteristic habitat types: mixed hardwood, water, low vegetation, road, sand, and mangrove (Figure 4). The mixed hardwood habitat type is dominated by upland and facultative palm (e.g., Phoenix spp., Sabal palmetto), often with an understory of sea grape, (Coccoloba uvifera) and is primarily located along the west border by the paved road. The low vegetation habitat type is common along the side of the road (mowed grasses) and in the interior of the training site, where it is characterized by scattered buttonwood trees, both alive and dead, (Conocarpus erectus), hardwood shrubs (e.g., Maytenus phyllanthoides), and a patchy understory comprised of salt-tolerant species, such as glasswort (Salicornia ambigua) and seacoast sumpweed (Iva imbricata). Light-colored sands are evident in breaks in the vegetation, alongside the paved road, and from the surface of the unpaved road along the southern edge of the training site (Figure 3). A mangrove forest occupies the saltwater lagoon shoreline along the study site’s eastern border. The mangroves are tree-sized (approximately 5 m in height) and form a dense, almost impenetrable thicket in this protected environment (Figure 4F). Species include red mangrove (Rhizophora mangle), white mangrove (Laguncularia racemosa), and black mangrove (Avicennia germinans).

2.7. Ground-Based Elevation Survey

To assess the accuracy of our UAV-LiDAR DSM and DTM, we compared those elevation values to data we collected during a ground-based survey using a Topcon GTS-249 (white points, Figure 5; Topcon Positioning Systems, Inc., Livermore, CA, USA) and to data previously collected from crewed-aircraft LiDAR-based DSM and DTM, hereafter aircraft LiDAR.

We conducted a ground-based elevation survey at our training site by extracting elevation data along the road, traveling from south to north (Figures 5 and 6). Each raster location matched the ground-based survey data, correcting for any offset in the UAV-LiDAR model.

The aircraft LiDAR DEMs were created from data collected between 2002 and 2015 (Table 2); [38–42]. This multi-year dataset was combined using weighted averages to create the models used here. DTMs and DSMs were created and provided at 2 m resolution by colleagues at Texas A&M-Corpus Cristi. We refer to these data as 2015 DTM or 2015 DSM.

We used the 2015 DTM and 2015 DSM to produce a 2015 aircraft LiDAR CHM and compared it and the 2022 UAV-LiDAR CHM at points sampled at 2 m intervals (Figures 5 and 7). The 2 m interval corresponds to the resolution of the 2015 elevation models.

To create a DEM of difference between the 2022 UAV-LiDAR CHM and the 2015 aircraft LiDAR CHM, we subtracted the 2015 CHM values from the 2022 CHM. The DEM of difference illustrates the change in canopy surface elevation between 2015 and 2022.

Figure 5. The digital terrain model results based on the 2022 UAV-LiDAR. This model was included as one of the classification features and to create the canopy height models. White points correspond to points in Figure 6. Black points correspond to points in Figure 7.
2.8. Machine Learning Habitat Analysis

2.8.1. Random Forest Models

Random forest is an ensemble machine learning algorithm that takes a training set of data across multiple features, fitting hundreds of decision trees [43,44]. Classifications are decided from a consensus across these decision trees [44]. Random forest has been extensively used as a classifier and regression tool. Its performance has been compared to other ensemble-based and supervised machine learning algorithms, showing that it is one of the top performers [45,46]. It has also been used for vegetation classification and identification [47,48]. We chose a random forest approach for its ease of use, rapid processing, and ability to handle large datasets with moderate computing power [49,50]. We implemented the random forest method with the Python package sci-kit learn [51].
2.8.2. Training and Testing

Our initial model used the five single-band rasters from the multispectral photogrammetry, (Table 1), the NDVI raster, and the 2022 UAV-LiDAR DTM, which was resampled to match the 2 cm resolution of the other rasters (Figure 5). We chose the bare earth DTM raster since mangroves and wetlands are sensitive to slight changes in ground elevation and fused all rasters into a single multi-band raster for ease of processing.

To create the training data, we defined polygons where sections of land use were known to be uniform based on our initial field site survey and photo interpretation of the orthophoto imagery. Cumulatively, these polygons only represent a small area of the entire training site (Table 3). There were seven land cover classes, six determined during our initial habitat characterization plus a shadow class (Figure 4). Shadowed subjects in multispectral and hyperspectral imagery give different spectral signatures, so this class was created to minimize misclassification in dimly lit locations, following [28]. We sampled randomly from the pixels within each polygon to create the training sets. A pixel-based approach was used to fully utilize the high-resolution data gathered during the UAV surveys. We gathered two sets of samples for each class—one set of 2000 pixels per class and the other of 5000 pixels per class. These amounts were chosen to limit computing time, allowing us to quickly investigate multiple models with different input data sets. Principle component (PC) analysis was performed on each model to assess artifacts in the training process.

Table 3. Classification categories with their associated number of polygons for extracting training data and the percentage of training area occupied by the polygons. Polygons were only created where habitats were uniform with little ambiguity.

<table>
<thead>
<tr>
<th>Habitat</th>
<th>Number of Polygons</th>
<th>% of Training Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Hardwood</td>
<td>9</td>
<td>6.6</td>
</tr>
<tr>
<td>Water</td>
<td>1</td>
<td>2.4</td>
</tr>
<tr>
<td>Low Vegetation</td>
<td>9</td>
<td>1.6</td>
</tr>
<tr>
<td>Road</td>
<td>3</td>
<td>4.6</td>
</tr>
<tr>
<td>Sand</td>
<td>8</td>
<td>1.1</td>
</tr>
<tr>
<td>Mangrove</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Shadow</td>
<td>22</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>59</td>
<td>23.8</td>
</tr>
</tbody>
</table>

Before model fitting, 20% of the training data were held back for model validation. Models were fitted with default parameters, except with ‘oob_score’ set to true for calculation of the out-of-bag error. We created two models, one for each training set, a 2000-pixel model and a 5000-pixel model. After the initial model fit, we performed a grid search optimization, refitting the model to obtain one with optimum estimators. The estimators we optimized for were the number of decision trees and the number of features included in each decision tree. Feature importance graphs were used to better understand the final optimized random forest models for each training set. Feature importance was calculated using the mean decrease in impurity method. This method assigns importance values to features based on the number of classes in branching nodes that branch from that feature in a decision tree. This means the fewer the classes in a node, the more “pure” it is, and thus, the more discriminatory or important a feature is.

2.8.3. Classification Map Filtering

We generated the testing site dataset in the same manner as the training site dataset, described in Sections 2.3–2.5 (Figures 2 and S1). We then classified the testing site with each of the random forest models, 2000- and 5000-pixel models, resulting in two classified habitat maps, one for each model.
Both results had pixels classified as shadow based on the scheme described in Section 2.8.2 To improve information content, we reclassified shadow pixels into one of the other classes based on a nearest-neighbor analysis. We applied a moving window majority filter with a window size of 11 pixels. The majority filter includes pixels in the $11 \times 11$ window around the shadow pixel, reclassifying the target pixel to the most common class in that window. This procedure resulted in two additional rasters, which we refer to as either 2000- or 5000-pixel (no shadow) habitat maps.

We applied a 123-pixel window majority filter to the 2000- and 5000-pixel (no shadow) maps to create a second pair of maps with a resolution that reflected the approximate size of tree tops at the field site (6 m$^2$). We will refer to these maps as 2000- or 5000-pixel smoothed habitat maps. These two filtering processes gave us four habitat maps for validation.

2.8.4. Validation Habitat Characterization

To assess the habitat map accuracy, we conducted ground surveys in the testing site (January 2023 and December 2023) at 54 validation points. Forty-two validation points were pre-selected randomly using functions built into the QGIS software (QGIS Association). Twelve points were added opportunistically in the field to ensure adequate representation of habitat types initially underrepresented in the validation point dataset. At each validation point, we obtained photographs, identified characteristic plant species, and described notable structural characteristics, e.g., dead branches and substrate type.

2.8.5. Validation Metrics

We calculated the user’s accuracy and the producer’s accuracy for each habitat type. Producer’s accuracy is defined as the number of validation points correctly assigned by the model to a particular category divided by the total number of validation points in that category [52]. User’s accuracy is defined as the number of validation points correctly assigned by the model to a particular category divided by the total number of validation points the model assigned to that category [53]. Thus, the producer’s accuracy is a measure of the quality of a classification category, while the user’s accuracy is a measure of how well the classification maps to the real world. Other metrics include balanced accuracy and kappa score. Balanced accuracy is computed by computing the accuracy per class and then averaging across these values [54]. The kappa coefficient was computed for the whole dataset. Kappa coefficient is a measure, from –1 to 1, of how well the classifier compares to random chance [55].

3. Results

3.1. Elevation and Canopy Height Analyses

Initial elevation ground-truthing revealed a 1.6 m vertical offset between our points measured during the ground-based total station survey and those of the UAV-LiDAR. This offset occurs in the reference frame conversion from WGS84 to NAD83/NAVD88 and is compatible with observations by NOAA’s National Geodetic Survey [56]. After correction, the ground-based survey elevations were plotted against both the corrected UAV-derived elevations and the 2015 DTM elevations (Figure 6). The UAV-LiDAR has centimeter levels of correlation with the ground-based survey, giving a vertical RMSE of 1.9 cm. The 2015 DTM has a slightly larger scatter with a vertical RMSE of 3.7 cm.

The two LiDAR sources, 2022 UAV-based and 2015 aircraft-based DEMs, were measured at separate times and hence allowed some assessment of change in land use or canopy height. Here the elevation ground-truthing process becomes vital. We observed minor (cm-level) variation in the unmoving section (e.g., the paved road) of each data set. This allows us to separate real changes in the habitat from georeferencing errors. Our DEM of difference (Figure 8A) shows pockets of elevation loss and widespread growth in parts of the tree canopy. We acknowledge that aircraft LiDAR tends to overestimate canopy cover, due in part to scan angle, point density, etc. [57,58]. The metadata provided for the aircraft LIDAR source data did not describe these parameters; however, we validated the
majority of the elevation loss. We accomplished this by field observations as described in Section 2.5 and through Google Earth imagery. The Google Earth satellite imagery from 2015 depicts a more continuous canopy throughout the field area (Figure 8C). In 2022, a more discontinuous living canopy is evident, with whitish-gray patches of dead vegetation (Figure 8B). Based on our ground-truth observations, we suggest that most of the elevation losses reflect the die-off of individual buttonwoods, while most of the elevation gains reflect canopy growth. Areas maintained by the park, roadways, and adjacent lawns show minimal to no overall growth, as expected.

Figure 8. Comparison of the DEM of difference with color imagery. (A): DEM of difference between the 2015 aircraft-LiDAR and the 2022 UAV-LiDAR canopy height models, both normalized to the 2020 ground surface. (B): 2020 orthophoto from UAV photogrammetry. (C): Satellite image from 2015 of the training site. Blue circles highlight some areas of negative change. Imagery in (C) taken from Google Earth.

3.2. Machine Learning Habitat Classification

The results of the principal component (PC) analysis reveal several key points, best seen by plotting the input features in the 3-D PC space (Figure 9). When plotting the input pixels by classification within the PC space, we observed that most classifications occupied discrete areas. The water classification was especially distinct, as expected. The most overlap was observed between mixed hardwood and mangrove PCs (Figure 9). Minor overlaps occurred between adjacent pixels near the margins of the low vegetation and shadow classes. These two classes also appear adjacent to mangrove and mixed hardwood (Figure 9).
Each of the seven model features shows good variability through the space. Spectral features are grouped by wavelength proximity. Red, green, and blue form one cluster, while near-IR and red-edge form a second cluster. NDVI and DTM show clear separation from all other groups, indicating their importance for classification.

Initial models (both 2000-pixel and 5000-pixel) showed high model accuracies (>95%) on the 20% of data that was held back. These initial models also had low OOB errors (Table 4). The final grid search optimized 2000-pixel model has an accuracy of 98%; the grid search optimized 5000-pixel tuned model was comparable at 97% (Figure 3). Most of the model misfit was between the mixed hardwood and mangrove classifications, consistent with the PC analysis. In the 2000-pixel model, the lowest classification accuracy was a tie between mixed hardwood and mangrove classes at 95%. In contrast, the 5000-pixel model’s lowest score was 92% accuracy for the mixed hardwood classification. Once the best-fit models were obtained, feature importance was calculated (Figure 10). These results indicate the DTM and NDVI are the most diagnostic parameters, while blue and green bands were the least important.
Figure 10. Feature importance graphs. (A): Feature importance calculated for the 2000-pixel model. (B): Feature importance calculated for the 5000-pixel model.

Table 4. Results of the random forest grid search and subsequent model fitting analyses performed on the 2000- and 5000-pixel models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Number of Estimators</th>
<th>Max Features</th>
<th>Out-of-Bag Error</th>
<th>Tuned Model Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000-point</td>
<td>100</td>
<td>5</td>
<td>0.27</td>
<td>98</td>
</tr>
<tr>
<td>5000-point</td>
<td>300</td>
<td>4</td>
<td>0.2</td>
<td>97</td>
</tr>
</tbody>
</table>

We applied the 2000- and 5000-pixel models to the testing site to generate habitat classification maps. The classifications from these models produced visually similar results (Figures 11 and 12). Some obvious differences are present west of the road. Looking at the filtered classifications, where each map was down-sampled to 6 m resolution, we see the large-scale trends preserved while smoothing out noise in the finer-resolution map (Figure 12).
Figure 11. Habitat classification maps with shadow pixels removed (testing site). (A): The 2000-pixel (no shadow) habitat map. (B): The 5000-pixel (no shadow) habitat map.
Incorporating higher-resolution data (i.e., 5000 pixels versus 2000 pixels) into the model did not increase map accuracy. The accuracy metrics were relatively similar across all four maps, but the map with consistently higher accuracy ratings was the 2000-pixel smoothed habitat map (Table 5). The 2000-pixel smoothed habitat map identified the water and the mangrove habitats best with a producer’s accuracy of 100% and 85.7% and a user’s accuracy of 100% and 80%, respectively (Table 6). From a producer’s perspective, mixed hardwood
and sand habitats performed the worst (54.5% and 71.4%), while from a user’s perspective, the road and low vegetation were the worst-performing (50% and 71.4%) (Table 6).

Table 5. Accuracy metrics for each of the final habitat maps.

<table>
<thead>
<tr>
<th>Model</th>
<th>Balanced Average (%)</th>
<th>User’s Accuracy Average (%)</th>
<th>Producer’s Accuracy Average (%)</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000-point, no shadows</td>
<td>75</td>
<td>79.3</td>
<td>76.2</td>
<td>0.70</td>
</tr>
<tr>
<td>2000-point, filtered</td>
<td>78</td>
<td>80.8</td>
<td>79.2</td>
<td>0.70</td>
</tr>
<tr>
<td>5000-point, no shadows</td>
<td>77</td>
<td>80.2</td>
<td>79.1</td>
<td>0.68</td>
</tr>
<tr>
<td>5000-point, filtered</td>
<td>78</td>
<td>77.3</td>
<td>76.5</td>
<td>0.68</td>
</tr>
</tbody>
</table>

Table 6. Accuracy metrics by habitat for the 2000-pixel smoothed habitat map.

<table>
<thead>
<tr>
<th>Habitat</th>
<th>User’s Accuracy (%)</th>
<th>Producer’s Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Hardwood</td>
<td>100</td>
<td>54.6</td>
</tr>
<tr>
<td>Water</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Low Vegetation</td>
<td>71.4</td>
<td>83.3</td>
</tr>
<tr>
<td>Road</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>Sand</td>
<td>83.3</td>
<td>71.4</td>
</tr>
<tr>
<td>Mangrove</td>
<td>80</td>
<td>85.7</td>
</tr>
</tbody>
</table>

4. Discussion

Our UAV methods provide a range of benefits compared to the 2015 DEMs derived from crewed aircraft. These include greatly improved spatial resolution and temporal sampling at the expense of limited aerial coverage. Incorporation of LiDAR shows benefits relative to photogrammetric methods, reflecting the latter’s inability to penetrate the canopy [27,59]. Flying low and slow with overlapping scans allowed us to successfully penetrate dense vegetation.

With its higher spatial resolution, UAV surveys can better match remotely sensed features with ground surveys (down to the level of individual plants) compared to aircraft or satellite surveys, tracking both small and substantial changes between time steps. This is best seen in the DEM of difference, which can be used to detect changes in canopy height while highlighting areas with significant (>1 m) elevation change (Figures 7 and 8A). Colorized maps can be used in the field to connect these topographic changes with local vegetation (Figure 8B,C).

4.1. Assessment of Limiting Features and Pixel-Based Classification

With only two UAV flights, one for the P4Multi and one for the UAV-LiDAR, we were able to capture six model features (NDVI being derived from two of the six). Many studies of this type utilize tens to hundreds of features, at least initially, for machine learning-based classifications [12,60–62]. These features are usually whittled down using various methods to find only the most important, typically some combination of spectral indices, positional information, and optical data [8]. Our work reached similar conclusions using only the first-order data products, saving time and effort, albeit with some loss of accuracy. Further trade-offs exist when approaching classification from an object-based approach, requiring more pre-processing. This adds complexity with appropriate variables to delineate each image segment [61,63]. Our methods minimize these complexities and trade-offs, offering a process that is more approachable for park managers and other non-specialists. As technology progresses in UAV lift capacity and sensor miniaturization, multiple data types can be collected on a single flight, but this will increase data processing complexity while not necessarily increasing utility.

4.2. Model Efficacy

Using just a small fraction of available data (23.8% of the training area) we achieved accurate classification models with minimal parameter tuning. Both the 2000- and 5000-pixel models achieved this with less than 400 estimators using roughly half of the available
features. These models could be made even more efficient by swapping lower-importance features with an additional vegetation index or positional model such as a DSM which has been found to increase model accuracy [26].

Both the 2000- and 5000-pixel habitat maps performed well overall. For individual habitats, the 2000-pixel smoothed map performed the best. Our results suggest that smoothing reduces edge effects, allowing a consensus of neighboring pixels to define the class. This has been found in at least one other study [63], which acknowledges that post-processing of pixel-based models can provide similar generalized maps compared to those created using an object-based approach.

Overall, these results show that more pixels in the training data do not necessarily give better results. One explanation for this could be the difference in important features of the 5000-pixel model (Figure 10). This model relies heavily on the NDVI and gives a higher importance to the NIR feature. These differences could result in the overfitting of certain classes.

Comparing the frequency histogram table (Table 7) for each habitat map allows us to compare results quantitatively. We see agreement between habitat maps that share random forest models. This suggests that our smoothing operations did not distort the classification. In some cases, smoothing clearly improved the classification. This suggests that an object-based approach would yield similar results, as it is based on larger groups of pixels like the smoothed maps. The largest change in frequency is between the 2000- and 5000-pixel low vegetation classification. In the 2000-pixel (no shadow) habitat maps, it was classified in 21.8% of the pixels, but in the 5000-pixel (no shadow) map, it comprised ∼15% of the pixels, an approximate 7% reduction (Table 7). In contrast, mixed hardwood’s frequency in the 2000-pixel habitat map is ∼8%, while in the 5000-pixel maps, it is ∼15%, a 7% increase in frequency. The mangrove classification was unchanged. This suggests that our models were well-defined in most of our classifications, with some refinement needed between the low vegetation and mixed hardwood classes. In future studies, this might be accomplished by adding a canopy height parameter.

Table 7. Numerical breakdown of frequency histograms for each habitat map. Values in each column sum to 100%.

<table>
<thead>
<tr>
<th>Habitat</th>
<th>2000-Point (%)</th>
<th>5000-Point (%)</th>
<th>2000-Point, Filtered (%)</th>
<th>5000-Point, Filtered (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed Hardwood</td>
<td>8.09</td>
<td>14.8</td>
<td>7.1</td>
<td>14.8</td>
</tr>
<tr>
<td>Water</td>
<td>19.4</td>
<td>20.6</td>
<td>19.4</td>
<td>20.6</td>
</tr>
<tr>
<td>Low Vegetation</td>
<td>21.8</td>
<td>14.9</td>
<td>22.5</td>
<td>14.2</td>
</tr>
<tr>
<td>Road</td>
<td>7.6</td>
<td>6.3</td>
<td>7.6</td>
<td>6.5</td>
</tr>
<tr>
<td>Sand</td>
<td>8.7</td>
<td>10</td>
<td>8.4</td>
<td>10.1</td>
</tr>
<tr>
<td>Mangrove</td>
<td>34.2</td>
<td>33.3</td>
<td>34.8</td>
<td>33.6</td>
</tr>
</tbody>
</table>

4.3. Challenges and Future Work

There were several limitations to this survey which came from the use of UAVs as the mode of data collection. As discussed above, the field area for most UAV surveys is limited. This comes from the amount and capacity of the available batteries. Further, the instrumentation chosen can decrease the flight time based on the UAV’s lift capacity. These challenges can be overcome with proper planning. The timing of the photogrammetry surveys needs to be an additional consideration. For both the multispectral and RGB images, flights closer to solar noon result in the best quality of images by limiting shadows. Additionally, multispectral images need to be corrected based on solar reflectance. The standard method uses some form of calibrated reflectance panel; here, we instead rely solely on the downwelling light sensor integrated into the UAV. This could have introduced some error in our spectral data, possibly resulting in some of the misclassifications seen above. Future studies using these methods should conduct their respective surveys on the same day or otherwise close in time. Here, we used data collected in the same season for consistency. The next step for this research is to expand in size and scope, gathering
data over a larger area using a long-endurance UAV such as a fixed-wing model [64] and gathering data throughout the year to further verify this methodologies’ results.

5. Conclusions

The proliferation of hardware and easy-to-use software has simplified geodetic and ecological surveying. UAVs have lowered the cost of data collection, enabling higher-resolution surveys compared to space-based satellite and high-altitude aviation, albeit with much less spatial coverage.

Landscape classification based on machine learning has also improved to the point where non-specialists can apply the technique. We used a Python-based machine learning package providing all the algorithms needed for machine learning model generation, data preparation, and analysis. This allowed us to create a streamlined workflow where we generated models and reviewed these in minutes on low-cost workstations.

Combining topographic and multispectral raster layers allowed us to study and characterize a mixed-use subtropical lowland at the mouth of Tampa Bay, Florida. Simple on-site workflows for UAV deployment facilitated rapid and spatially accurate data collection. Several machine learning-based classification models were tested using the Random Forest algorithm. Despite the small training area, comparison with ground-based validation data indicates that the 2000-pixel smoothed models are 80% accurate. A smoothing process based on a consensus of neighboring pixels improved classification accuracy.

For a coastal wetland site, adding bare earth elevation data to multi-spectral optical imagery improves classification. With the growing interest in wetland habitats as markers of climate change, detailed mapping will be needed. Our approach offers a cost-effective solution, with sparse parameter input (five spectral bands plus LiDAR-based elevation) and minimal processing. These methods should work in other environments where habitats are partitioned particularly by ground elevation, but this will require additional research.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/drones8030113/s1. The supplemental material describes further the exact steps used in the material and methods section of the manuscript. It also discusses the difference between RTK and PPK georeferencing results for photogrammetric data. We performed PPK methods with data from CODEs using the precise orbits and the rtklib software [65–67]. The data processing used functions from [68]. Figure S1: Workflow for processing the LiDAR and photogrammetry surveys to create a multi-band raster; Figure S2: Workflow for the random forest classification; Figure S3: Comparison of modeled DEM derived from RTK and PPK processing with ground truth; Figure S4: UAV position and elevation comparison between RTK and PPK processing with residuals.


Funding: This research was supported by a grant from the National Science Foundation (SPOKES project, #1762493, F. Muller-Karger PI); grants to T.H.D. from NASA (80NSSC22K1106), NOAA’s Office of Coast Survey (COMIT Cooperative Agreement NA20NOS4000227), and the USACE Engineer Research and Development Center (Cooperative Agreement W9132V-22-2-0001) for final data analysis and refinement of methodology for expanded applications.

Data Availability Statement: Data are available on request from authors.

Acknowledgments: We would like to thank Texas A&M for supplying us with the aircraft LiDAR, Josephina Reyman for aiding in fieldwork, and Edgar Guerron-Orejuela for assistance in the creation of the flowcharts. We also thank the staff at Ft. DeSoto County Park, particularly Park Manager David Harshbarger, for allowing us access to the field area and giving us permission for UAV operations in the park.
Conflicts of Interest: The authors declare no conflicts of interest.

References


26. Pricope, N.G.; Minei, A.; Hallis, J.N.; Chen, C.; Wang, Y. UAS Hyperspatial LiDAR Data Performance in Delineation and Classification across a Gradient of Wetland Types. *Drones* 2022, 6, 268. [CrossRef]


32. Agisoft LLC. *Agisoft Metashape Pro*, Version 1.8.2; Agisoft LLC: St. Petersburg, Russia, 2022.


41. OCM Partners. 2015 USACE NCMP Topobathy Lidar: Florida Gulf Coast. 2015. Available online: https://www.fisheries.noaa.gov/inport/item/49720 (accessed on 11 February 2024).


47. Burnett, M.; Mccauley, D.; Leo, G.A.D.; Micheli, F. Quantifying coconut palm extent on Pacific islands using spectral and textural analysis of very high resolution imagery Biogenic carbonates on sea and land View project Public Health and decision making View project. *Int. J. Remote Sens.* **2019**, *40*, 7329–7355. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.