Final report

Unmanned Aircraft System Lidar and Imagery in the National Estuarine Research Reserve System

Evaluating the Effectiveness of Unmanned Aircraft System Sensors and Platforms for Multi-Purpose Mapping of Marshes and Beaches in the National Estuarine Research Reserve System Sentinel Site Network

June 2018

National Oceanic and Atmospheric Administration
National Ocean Service
Office for Coastal Management

Principal Investigator:
Kirk Waters[1]

Co-Investigators in alphabetical order:
Sue Bickford[2]
Jamie Carter[3]
Nina Garfield[1]
Andrea Habeck[4]
Nate Herold[1]
Jared Lewis[5]
Jonathan Pitchford[6]
Melissa Rosa[3]

Supported by a grant from the NOAA Unmanned Aircraft Systems Program

[1] NOAA Office for Coastal Management
[2] Wells National Estuarine Research Reserve
[3] The Baldwin Group at the NOAA Office for Coastal Management
[4] Jacques Cousteau National Estuarine Research Reserve
[5] San Francisco Bay National Estuarine Research Reserve
[6] Grand Bay National Estuarine Research Reserve

This page intentionally left blank

TABLE OF CONTENTS

1 INTRODUCTION

Our intent with this project was to evaluate the effectiveness of unmanned aircraft system (UAS) platforms to produce multiple mapping data and products for elevation and vegetation mapping in marshes and dune systems. We sought a UAS solution that could fly multispectral and lidar elevation instruments sequentially on the same platform. We contracted to the private sector (Quantum Spatial, Inc., and PrecisionHawk) for UAS data collection and conducted the ground truth ourselves (National Estuarine Research Reserve System and NOAA staff members). We used multiple National Estuarine Research Reserve System sentinel sites as test beds. While we intended to conduct repeat collections to determine relative and absolute elevation change mapping abilities in different ecosystems and to evaluate the cost–benefit of acquiring multi-season data for vegetation mapping, events beyond our control prevented us from acquiring multi-season data.

The reserve system and other natural resource stakeholders have a constant need for accurate digital elevation models (DEMs) and habitat maps to support a diversity of applications. Applications include supporting sea level rise research and management and flood forecasts; evaluating the impact of specific vegetation management practices on elevation in marsh microenvironments; assessing beaches after storms for damage assessment and restoration purposes; and identifying high priority invasive and sensitive vegetation. Data from multiple high-resolution multispectral sensors and lidar elevation were acquired for three reserve sites: Jacques Cousteau, New Jersey; Grand Bay, Mississippi; and Rush Ranch in San Francisco Bay, California. The data were evaluated on their ability to meet specifications, primarily positional accuracy and resolution, and their potential to improve habitat mapping.

The technology readiness level (TRL) at the beginning of this project was approximately TRL 5 (concept validated in relevant environment). We proposed a prototype demonstration at pilot sites and evaluation of capabilities within the coastal and salt marsh environments to move to TRL 7. In the final analysis, we believe we have moved to TRL 8, system demonstration in an operational environment.

This work addresses two National Ocean Service Roadmap priorities: coastal intelligence and place-based conservation. Observations acquired through a UAS would contribute to a strong foundation of geospatial data available for oceans and coasts. This project will evaluate the potential for UAS to provide multi-purpose ecosystem mapping products in a timely, cost-effective manner. Ultimately, these products could improve the ability of resource managers to understand and evaluate the current state and dynamics of the environment they manage and enable them to make informed decisions for ecosystem protection, conservation, and management.

1.1 NATIONAL ESTUARINE RESEARCH RESERVE SITES

1.1.1 Jacques Cousteau National Estuarine Research Reserve

The Jacques Cousteau Research Reserve is in New Jersey just north of Atlantic City and encompasses areas around Great Bay and Little Egg Harbor. We wanted to use the reserve as our example of a dune ecosystem with the primary focus on the lidar elevation capabilities. We expected this to be the easiest ecosystem for lidar collection, with significant areas of bare sand and generally sparse vegetation. The originally proposed area for collection was the Edwin B. Forsythe refuge with a target collection in November 2016. This area is one of the very few dune systems within the reserve system nationwide.

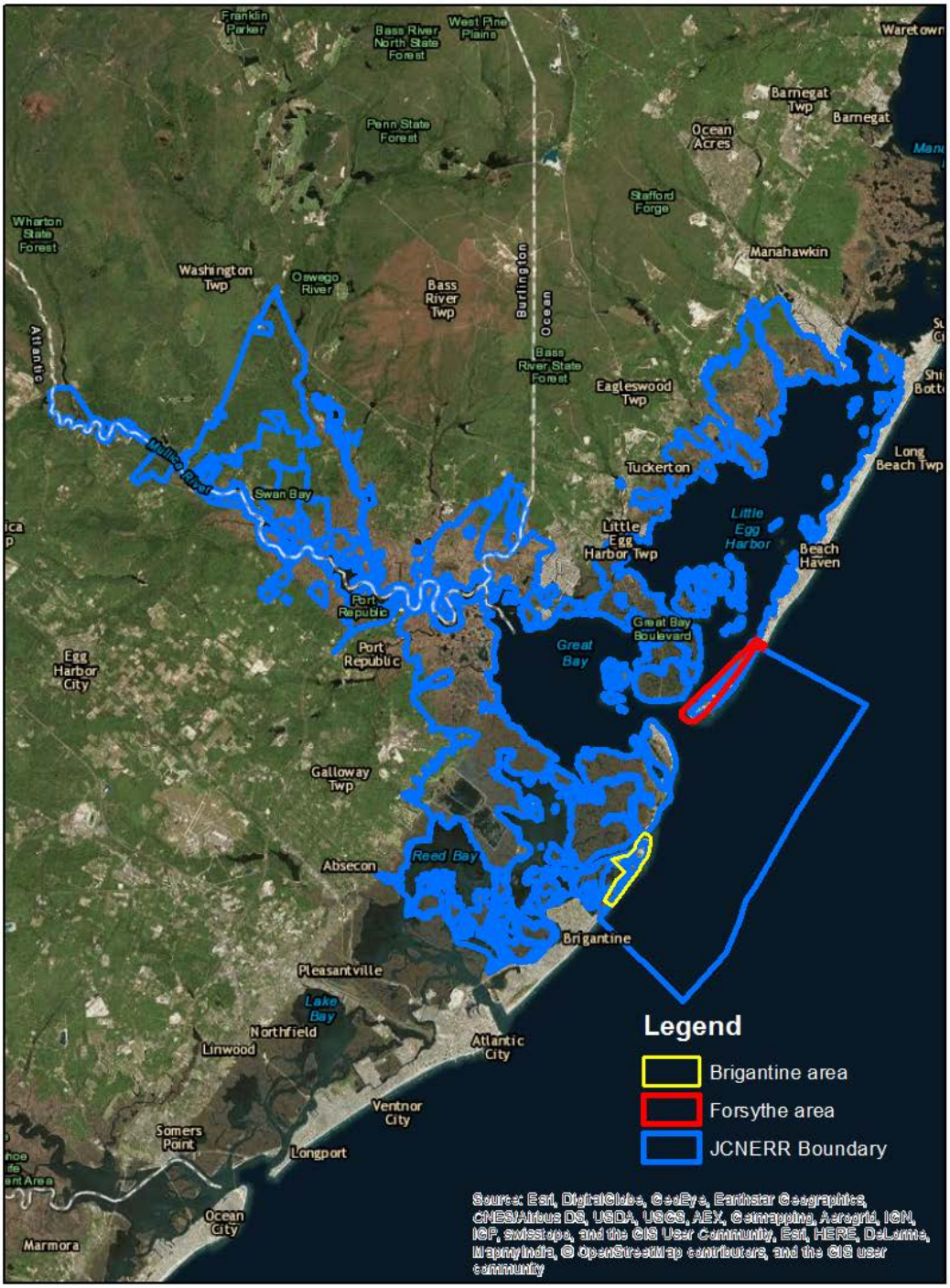


Figure 1. Jacques Cousteau Research Reserve boundary, with the Edwin B. Forsythe Wildlife Refuge in red and Brigantine State Park in yellow

As with all the reserves, the Jacques Cousteau Research Reserve does not own the land. The U.S. Fish and Wildlife Service manages the Edwin B. Forsythe Refuge and must grant permission to fly within the refuge. An application was submitted to the refuge with the available details on flight duration and patterns. The proposed UAS system determines the best flight pattern based on the winds at the time of flight, and the total time needed to fly the area

always depends on the weather conditions. The flight uncertainties led the Fish and Wildlife Service to deny the application.

With the Forsythe Refuge unavailable, we had to look for other options quickly. The only other dune ecosystem within Jacques Cousteau Research Reserve is at Brigantine State Park, managed by the State of New Jersey. An application was submitted in December 2016 with the intent to fly in March 2017. This was a narrow window, since the state requires three months to review applications, and the Brigantine area closes for bird nesting in mid-March. Fortunately, permission was granted in early March, though the window to get the job done was narrow, with a hard end date regardless of weather issues.

Figure 2. Photos at Brigantine State Park illustrating the land cover types and conditions

1.1.2 Grand Bay National Estuarine Research Reserve

Grand Bay Research Reserve, near Moss Point, Mississippi, comprises approximately 18,000 acres and contains pine savannas, salt marshes, salt pannes, bays, and bayous, as well as terrestrial habitats that are unique to the coastal zone. Four areas were selected within the reserve, representing the different ecosystems for the sentinel site. These areas had been previously flown with UAS-based imagery, providing an opportunity to compare the private-sector data to data collected through a university. We expected this area to be significantly more difficult for the lidar to penetrate compared to Jacques Cousteau Research Reserve.

Obtaining permission to fly at Grand Bay Research Reserve was straightforward and required communicating our plans to the U.S. Fish and Wildlife Service as managers of the area. Since this area has seen a number of UAS flights in the past, they were more comfortable with the effort, and our request was quickly approved. This site is also far enough away from airports that the UAS contractor's blanket certificate of authorization was sufficient for the flights.

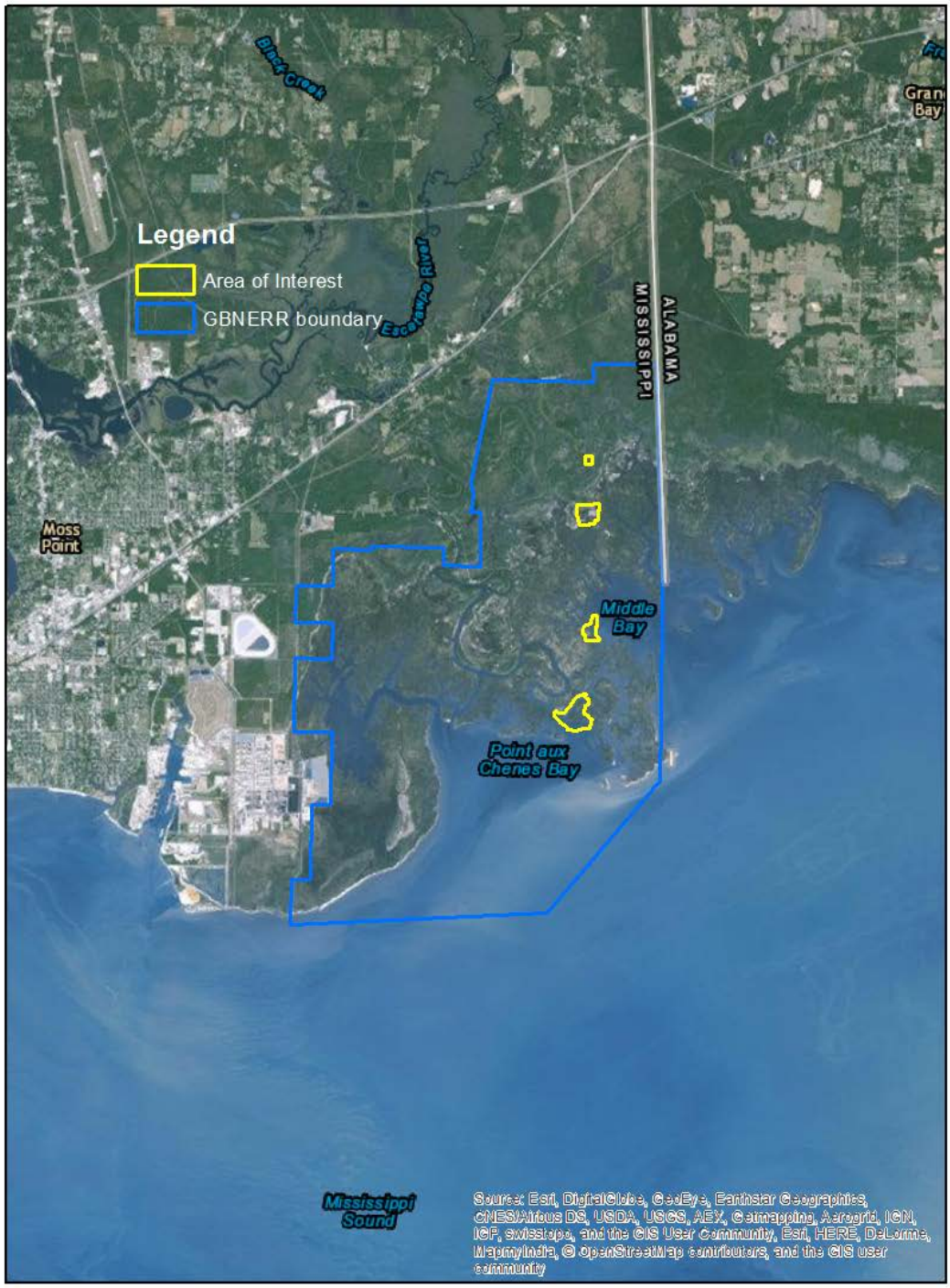


Figure 3. Grand Bay Research Reserve boundary in blue, with areas flown with UAS outlined in yellow. Areas are numbered 1 through 4 from North to South.

Figure 4. Ground view of the Grand Bay Research Reserve areas of interest. In clockwise order from the upper left, areas of interest 1, 4, 2, 3.

1.1.3 San Francisco Bay National Estuarine Research Reserve (Rush Ranch)

Rush Ranch is one of the two areas of the San Francisco Bay Research Reserve and consists of tidal marsh and cattle grazed upland near Suisun City, California. Our original plan was to survey the marsh area twice: once in the fall of 2016 and once in the late spring or early summer of 2017. Rush Ranch is within 5 miles of Travis Air Force Base and required an additional site-specific Federal Aviation Administration (FAA) certificate of authorization to

fly, unlike the other two reserves, which were covered by PrecisionHawk's blanket certificate of authorization.

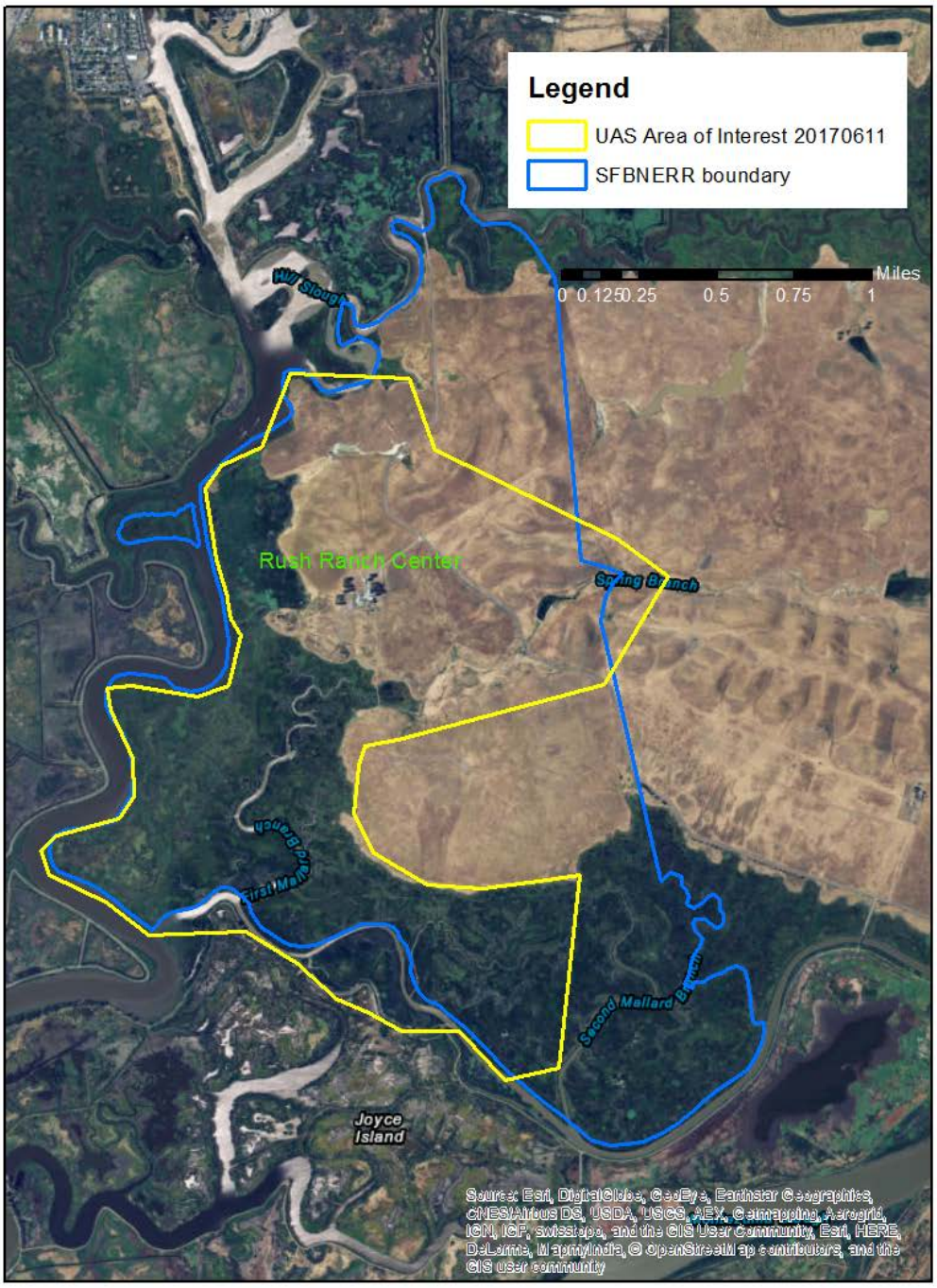


Figure 5. San Francisco Bay Research Reserve's Rush Ranch area shown in blue, with the area flown by UAS in yellow

PrecisionHawk applied to the FAA for a certificate of authorization in May 2016, shortly after the project was funded. We planned flights for September 2016. The certificate approval was delayed, but we went ahead with fieldwork plans in September in the hopes that the FAA would approve the certificate of authorization by early October, before the vegetation state changed. October came and went without FAA approval, and we missed the

flying season. Discussions with the San Francisco Bay Research Reserve director suggested we should target June 2017 for our first flights to avoid black rail nesting season.

The FAA did not approve the certificate of authorization until December 2016. The FAA transitioning from the Section 333 exemptions to the Part 107 rules may have partly caused the delay. When they did approve the certificate, it was only valid through April 2017, when PrecisionHawk's Section 333 exemption expired. PrecisionHawk had transitioned to the simplified Part 107 rules and certifications, and an extension had to be requested under Part 107. PrecisionHawk requested the extension, and they were told in February 2017 that it should be approved in a few weeks. The final approval came in May, just in time for the June flights.

Meanwhile, back at Rush Ranch, instead of the marsh being accessible in June as expected, it would stay closed for endangered birds until September 1, 2017. We made plans to fly the upland areas in June instead. Shortly before mobilizing, we decided to do one set of flights of both the marsh and upland areas in September instead of flying the uplands in June and the marsh in September. This reduced mobilization costs for all parties. We already had no way to get multi-date data in the marsh, so that decision did not affect our research.

The marsh at Rush Ranch is very dense. While standing in the marsh, a person frequently cannot see any ground. The *Salicornia* (pickle weed) grows densely as much as a foot tall, making it one of the easier vegetation types to travel through. Higher patches of *Juncus*, *Schoenoplectus americanus*, *Typha*, and *Grindelia* were more difficult. All vegetation types varied in density, and the possibility exists to get lidar points to the ground in the sparser instances. In general, we expected Rush Ranch marsh to be the hardest area to penetrate to ground; this area also has the greatest variability in ground cover. The uplands typically had dry grasses (*Bromus* and star thistle) with some areas mowed or allowing cattle grazing (example vegetation is shown in Figure 6).

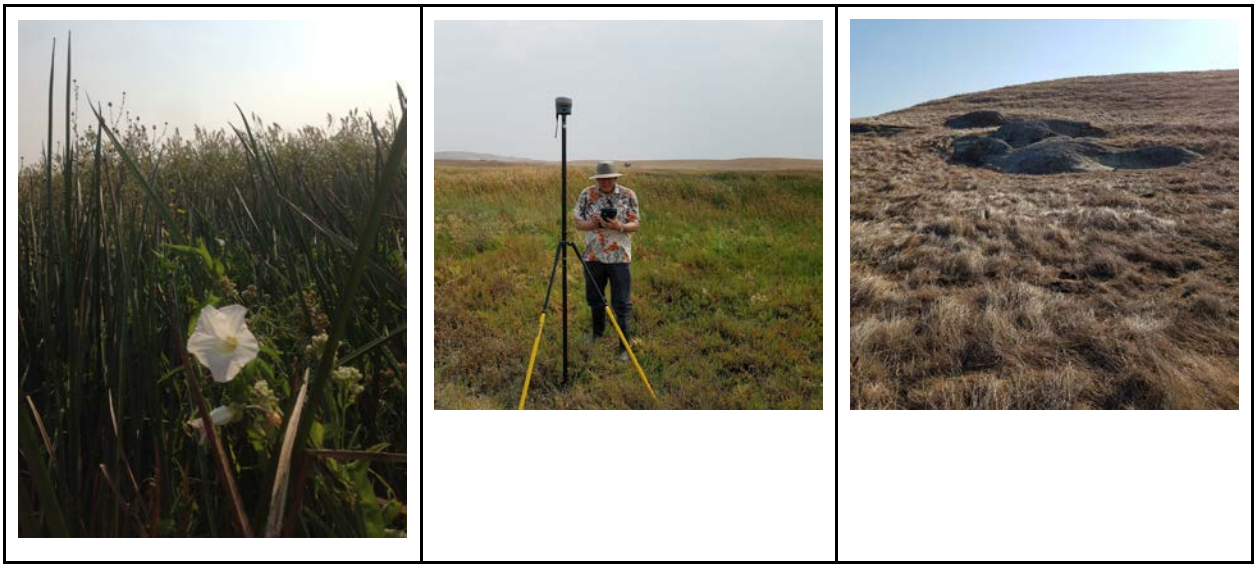


Figure 6. Example land covers at San Francisco Bay Research Reserve's Rush Ranch

1.2 CONTRACTING

The contract for UAS services was done as a task order on the Coastal Geospatial Services Contract administered through the NOAA Office for Coastal Management and the Eastern Administrative Support Center (EASC). This is a Brooks Act contract vehicle, in which awards go to the best-qualified company and are not selected through a bidding competition. With little data available for contracting UAS work, we selected a prime contractor (Quantum Spatial) with a solid record of accomplishment in metric mapping and a sub-contractor (PrecisionHawk) with extensive experience with UAS systems and working with the FAA.

The costs for each site were higher than originally anticipated in our proposal to the NOAA UAS Program. This was primarily due to the processing work needed for the lidar data, the extent of which was not fully appreciated during proposal generation. To fit the budget and best meet the objectives of the proposal, we scaled back the flights to do multi-date flights at only one of the sites. Since Rush Ranch (San Francisco Bay) was the most ecologically complex and therefore most likely to benefit from multiple seasons of data, we chose it for multiple dates. As noted earlier, this plan did not work out because of nesting restrictions.

During the planning phases of this project, we anticipated that PrecisionHawk would use its Lancaster 5 fixed-wing platform and hot swap the sensors. Technology is evolving rapidly in this area, and we quickly had a choice between staying with old technology for comparability between sites and moving to new technology to evaluate current capability. We chose to move to current capability. The changes made will be detailed in the results, but we note it here because private-sector contracting allowed us to take advantage of those changes instead of using what we could have bought at the start of the project.

Based on observations of the effort the contractors put into the job and the progress they made in improving their mapping capability through the course of this project, we believe that their costs exceeded what we paid. We mention this here because the costs per square mile for this UAS contract look very high compared to the typical costs for manned flights, and it might be thought that the contractor was overcharging. The reverse is true. The project made a good public-private partnership where each contributed resources to advance the technology readiness level for UAS-based metric mapping.

1.3 SPECIFICATIONS

The specifications for data products mirrored typical manned data specifications with some modifications for UAS capabilities (Table 1). For lidar, the 10 centimeter vertical uncertainty is typical of modern large-area collections and is part of the standard 3D Elevation Program specification. The 30 points per square meter density is higher than normally found for manned lidar and reflects the greater density expected from low altitude. Four-band imagery is also very typical for manned collections, but 3 centimeter pixels reflects what we believed should be attainable from the lower altitude. The 15 centimeter horizontal accuracy is also tighter than a typical manned imagery collection, but reflects a similar distance in terms of pixels.

Table 1. Imagery and lidar specifications

Lidar	
Vertical accuracy	0.10 meters **RMSEz** (root mean square error for *z*)
Density	30 points per square meter
Imagery	
Number of bands	4 (red, green, blue, NIR)
Ground sample distance/pixel size	0.03 meters
Horizontal accuracy	0.15 meters RMSE (root mean square error)

2 JACQUES COUSTEAU NATIONAL ESTUARINE RESEARCH RESERVE

2.1 METHODS

In preparation for UAS lidar and imagery flights at Jacques Cousteau Research Reserve, we placed spatially distributed targets in the North Brigantine State Natural Area. Targets were either 11 or 12 inches on a side with an "iron cross" pattern. We constructed the 11-inch targets with a laser printed pattern affixed to a heat sensitive foam board. The 12-inch targets were floor tiles spray-painted with the design. We held the foam tiles in place with long wire stakes inserted into the sand. The ceramic floor tiles were heavy enough to stay in place, and we used them in the more exposed areas to minimize potential movement. We surveyed target locations with a Trimble R8 real time kinematic (RTK) GPS. For six location, we also surveyed the targets with rapid static GPS for 20 minutes as a check on the accuracy of the RTK. We surveyed 33 photo-id targets, including four of existing features in the environment (concrete corner, piling, etc.). We placed and surveyed the targets on March 6, 2017.

We conducted elevation surveys for validation at the photo-id targets and along profiles. The profiles followed the standard beach profile methodology used by the research reserve system.

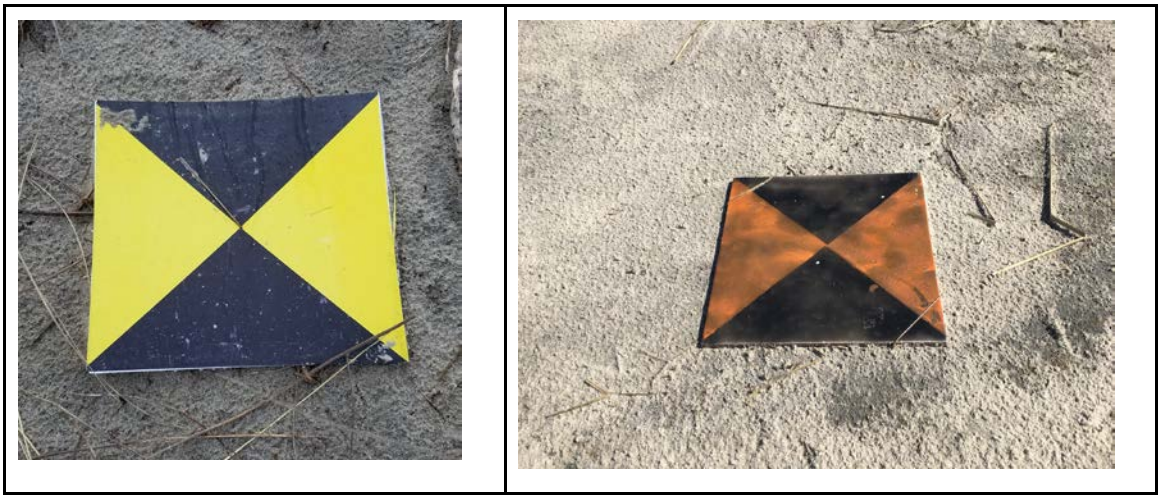


Figure 7. Examples of targets used. On the left is a foam-backed target and on the right is a floor-tile target. All targets had the potential for coverage by sand, which may have made their identification harder.

The flights for UAS lidar used a PrecisionHawk Lancaster 5 UAS with a Velodyne Puck VLP-16 based lidar system and commenced on March 7, 2017 (specification in Table 2). The Lancaster 5 is a fixed-wing platform

with a dual-frequency GPS. At a nominal survey altitude of 50 meters, the pulse density was projected to be 30 pulses per square meter. PrecisionHawk pilots conducted the flights as a sub-contractor to Quantum Spatial. Windy conditions were a problem throughout the week, frequently gusting over 20 mph (Figure 8). Lidar collections were carried out when the wind conditions allowed on March 7, 8, and 10.

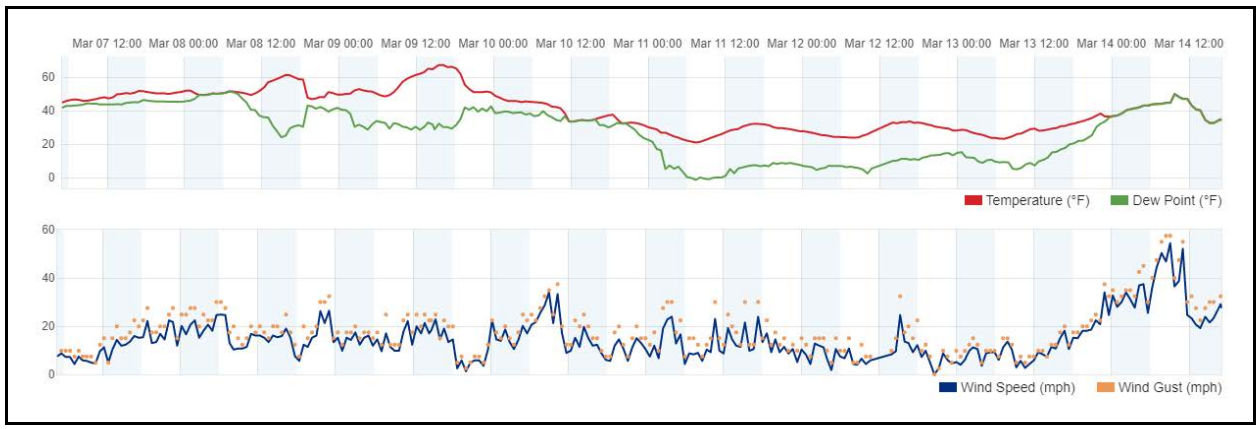


Figure 8. Weather conditions from March 7 to March 14, 2017, from the Steeleman Bay KNJBRIGA14 station

Table 2. Specifications for the lidar survey at Jacques Cousteau Research Reserve

Lidar Survey Settings and Specifications	
Acquisition Dates	March 7, 8, and 10, 2017
Aircraft Used	PrecisionHawk Lancaster Rev 5 Dual-Frequency GPS
Sensor	Velodyne Puck VLP-16
Maximum Returns	2 (Strongest/Last Return)
Nominal Pulse Spacing	18 centimeters
Nominal Pulse Density	30 pulses/meter^2
Survey Altitude (AGL)	50 meters
Target Speed	27 knots
Field of View	110 degrees
Scan Frequency	5-20 hertz
Pulse Rate	300 kilohertz
Pulse Duration	6 nanoseconds
Pulse Width	15 centimeters
Wavelength	903 nanometers
Pulse Mode	Single Pulse in Air
Beam Divergence	3 milliradians
Swath Width	143 meters
Overlap	50%

Imagery collection followed when wind conditions allowed on March 12, 2017, with a 5-band MicaSense sensor on the fixed-wing Lancaster (specifications in Table 3). The combination of the Lancaster airspeed and the MicaSense minimum frame rate of 1 hertz (Hz) required flying at an altitude resulting in a ground sample distance more than twice the contract specification (6.5 centimeters instead of 3 centimeters). The intent was to fly two 3-band sensors that did meet the resolution specification after the MicaSense flights; however, the weather prevented those flights from happening before the area was closed for bird nesting (piping plover).

A miscommunication between the sub and prime contractors resulted in no ground control being set for the lidar and imagery collections. This seriously affected the spatial accuracy of the products. A subsequent effort was made to improve the results using five of the NOAA validation points, though this proved inadequate.

Table 3. Specifications for imagery collection at Jacques Cousteau Research Reserve

Digital Orthophotography Specifications	
Equipment	MicaSense RedEdge
Spectral Bands	Blue, green, red, NIR, red edge
Ground Sample Distance	8 centimeters per pixel at 120 meters
HFOV (horizontal field of view)	47.2 degrees
Frame Rate	1 frame/second
Final Project Resolution	6.5 centimeter pixel size
Image	12-bit geoTIFF
Flight Altitude	100 meters

2.2 RESULTS

2.2.1 Coverage

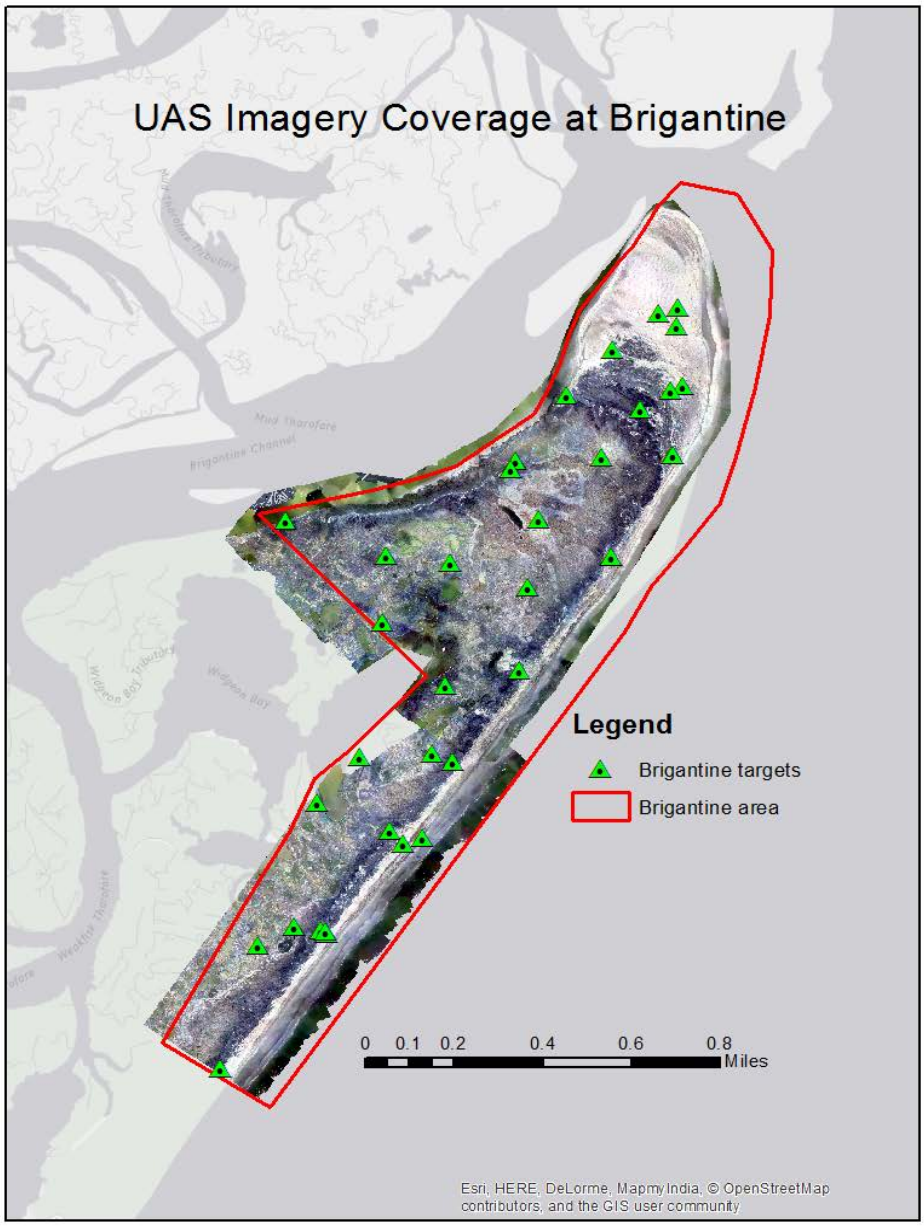


Figure 9. Imagery coverage at Brigantine State Park with photo-id targets shown as green triangles and the area boundary shown in red. The imagery collection was not extended over water to match the boundary by mutual agreement.

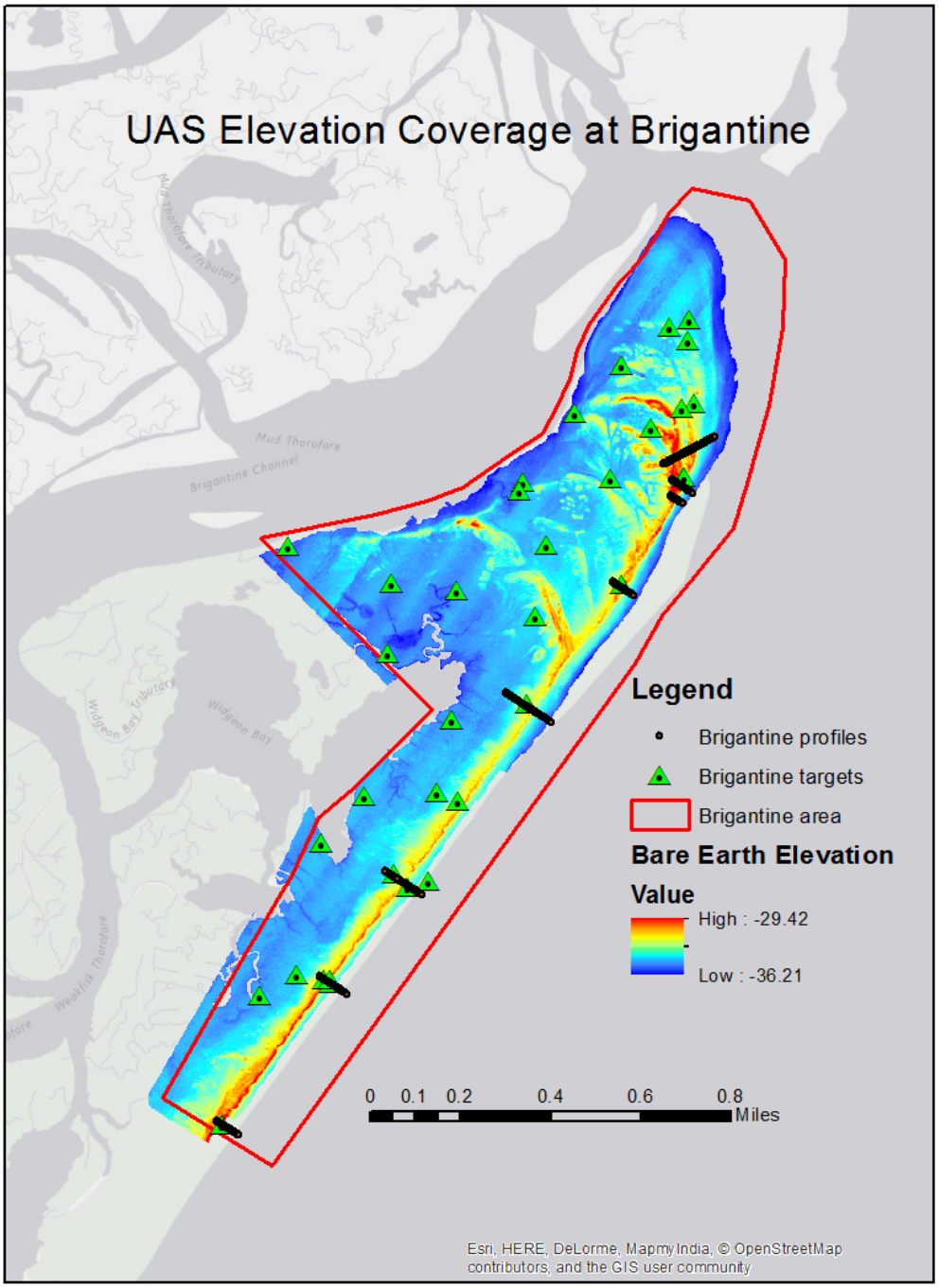


Figure 10. Coverage by lidar is shown for Brigantine State Park. Elevation checkpoints are shown as green triangles.

The imagery and lidar coverage are shown in Figure 9 and Figure 10, respectively. Part of the area of interest was over water, and the flights did not cover that area by mutual agreement since there was no need to risk the aircraft for data that we would not use.

2.2.2 RTK versus Rapid Static

We compared GPS results using RTK and rapid static observations for five points. Though it has limited statistical significance, we did it primarily as a validation that we were getting comparable results from the RTK such that we could trust the RTK as ground truth for further comparisons. The two methods had a mean difference of 2 centimeters horizontally and 1.1 centimeters vertically. We considered this sufficient validation of the RTK to use it as our ground control. The RTK observations were for less than 15 seconds, and the rapid static observations were for 20 minutes or more.

Table 4. Comparison between Static GPS and RTK

Static GPS (X90D)					RTK (Trimble R8)				Horiz Diff [m]	Vert Diff [m]
SiteID	Northing [m]	Easting [m]	Ellipsoid Height [m]		Site ID	Northing [m]	Easting [m]	Ellipsoid Height [m]		
JC01	4365198.788	557167.264	-32.555		t6	4365198.813	557167.26	-32.571	0.025	0.016
JC02	4363699.582	556080.667	-31.27		t27	4363699.622	556080.66	-31.273	0.041	0.003
JC03	4366007.492	557725.93	-31.687		t16	4366007.482	557725.92	-31.685	0.016	-0.002
JC04	4365624.971	557504.828	-32.078		t14	4365624.974	557504.83	-32.063	0.005	-0.015
JC05	4364596.078	556695.371	-32.941		t2	4364596.064	556695.37	-32.921	0.015	-0.020

2.2.3 Elevation

Nineteen of the twenty elevation-suitable validation points fell within the elevation coverage area. The 20th point did not have sufficient lidar coverage. We used the lascontrol tool from LAStools to create a triangulated irregular network (TIN) of the bare-earth lidar points around each validation point and interpolate the elevation at the validation point. The RMSEz was 1.46 meters with a standard deviation of 0.36 meters and an average error of negative 1.42 meters. This indicates that most of the error is in the bias, which is reasonable given the lack of control during collection. However, even if we remove the bias, the standard deviation is far higher than the 10 centimeter RMSEz specification.

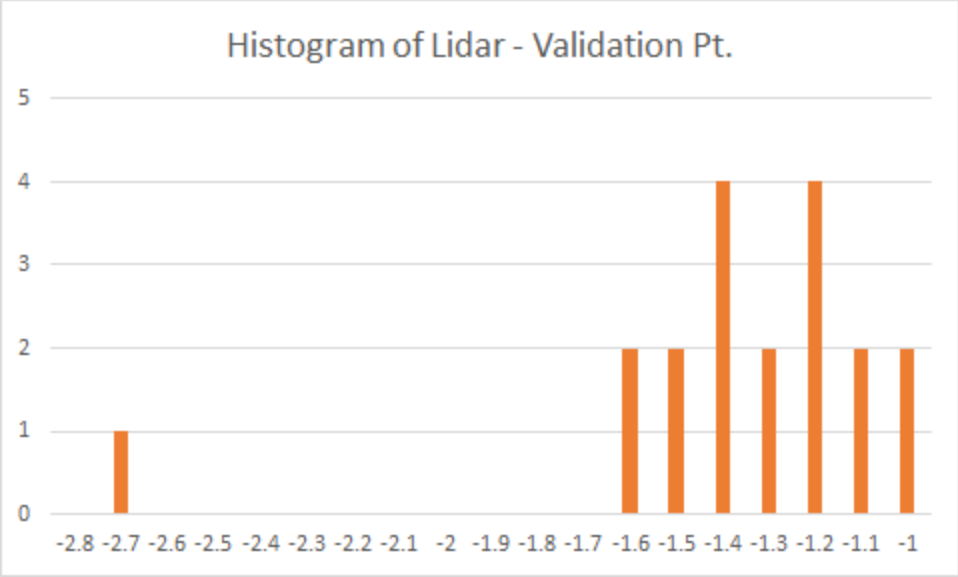


Figure 11. Histogram of errors for lidar at Jacques Cousteau

We also compared the RTK points from the profiles to the lidar. These points are spatially restricted to the beach/dune region (see coverage map) and may be in vegetation, although the vegetation would not have been dense. The RMSEz for the 472 points that lascontrol found near enough to lidar points is lower at 1.35 meters, although the standard deviation is higher at 0.97 meters.

The contractor (PrecisionHawk) had made significant improvements to its lidar processing between the time of the Jacques Cousteau Research Reserve data delivery (May 2017) and the flights at San Francisco Bay Research Reserve (September 2017). We gave them five of our control points to use in reprocessing the data in late September. The hope was to improve the data to the point where it would be useful to Jacques Cousteau Research Reserve. Reanalysis of the data indicated that there was no way to improve it and maintain scientific integrity of the data. PrecisionHawk is planning (as of June 2018) to do another flight within Jacques Cousteau during the fall of 2018 with improved platform and sensors, but that flight will not be in time for this report.

2.2.4 Imagery

2.2.4.1 Positional Accuracy

In general, we had difficulty finding the photo-id targets in the imagery. Because the image resolution was coarser than anticipated, there were fewer pixels covering a target than originally designed. Instead of a target covering approximately 10 x 10 pixels, they were only about 5 x 5 pixels. However, rarely did the targets show any of the design. The majority appeared as either a black square or a white square despite their actual design. This made it very difficult to be certain we were really looking at the target and accurately determining the center point. Some examples of the targets in the imagery are shown in Figure 12.

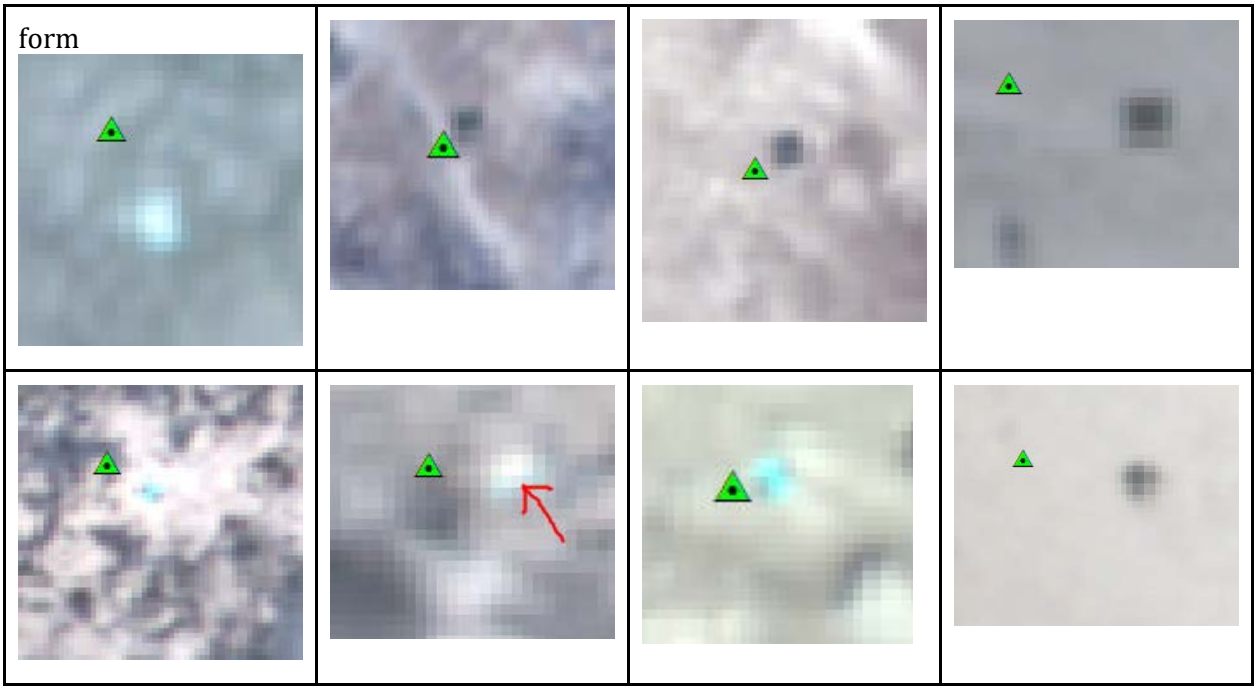


Figure 12. Examples of the targets seen in the imagery with no stretch. Green triangles show the position surveyed. Only a minority of the targets showed a noticeable pattern.

With 22 targets identified to the best of our ability, the radial RMSE was 0.60 meters, resulting in a National Standard for Spatial Data Accuracy (NSSDA) horizontal accuracy at 95% confidence of 1.04 meters. This is as expected given the lack of ground control used in the processing and is an important result illustrating the need for ground control when doing metric mapping. While obvious for those frequently engaged in metric mapping, many UAS operators may not have experience in this type of mapping if they come from fields where precise geopositioning is not important.

The contractor-provided report indicated that significant processing needed to be done to the imagery to get this level of accuracy, including non-standard practices used because of the positioning problems, as indicated by the following taken from the report, where PH is PrecisionHawk and QSI is Quantum Spatial, Incorporated.

> The initial mosaics delivered by PH had significant offsets when compared to base imagery from the NOAA National Geodetic Survey (NGS) Sandy Coastal Mapping project. The pixel sizes were also inconsistent between the two mosaics; therefore, QSI provided additional support to the orthoimagery portion of the project by resampling imagery to a pixel size of 6.5 centimeters to ensure consistency. To correct the inaccuracy, the PH-delivered mosaics were georectified to the NOAA NGS Sandy imagery using photo-identifiable features from each set of imagery. This resulted in a more accurate and useable dataset, but is considered a non-standard image processing practice and would not be used if correct GPS/IMU processing resulted in accurate initial positioning.

2.2.4.2 Image quality

The imagery was delivered as two 5-band orthorectified mosaics in geoTIFF format, one for the northeast end and one for the southwest end. Bands are in the order red, green, blue, near infrared (NIR), and red edge. The seamlines between images in the mosaics look very good. The contractor technical report states:

> Spectral artifacts exist in the imagery mosaics as well. This is likely due to a combination of issues including, but not limited to: 1) Inconsistent cloud cover, 2) Saturation, 3) Color balancing, and 4) Relative positioning of frames. PH is aware of these artifacts and is working to better calibrate the imaging system moving forward. Additionally, some of the imagery positioning would not solve, which manifested as gaps in the final mosaic.

The lack of distinct colors in the landscape makes it difficult to verify the band order of the data. It appears that it should be blue, green, red, NIR, and red edge, contrary to the order listed in the contractor's report. The images in Figure 13 show some of the spectral artifacts that may make the imagery difficult to use. Images are false color

infrared (RGB = NIR, red, green) and true color, depending on which band order is correct. The left images assume order blue, green, red, NIR, red edge. The right images assume red, green, blue, NIR, red edge.

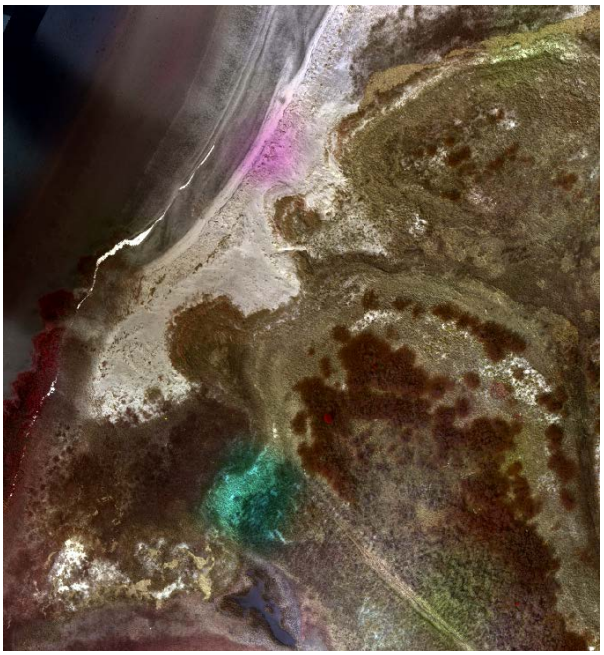
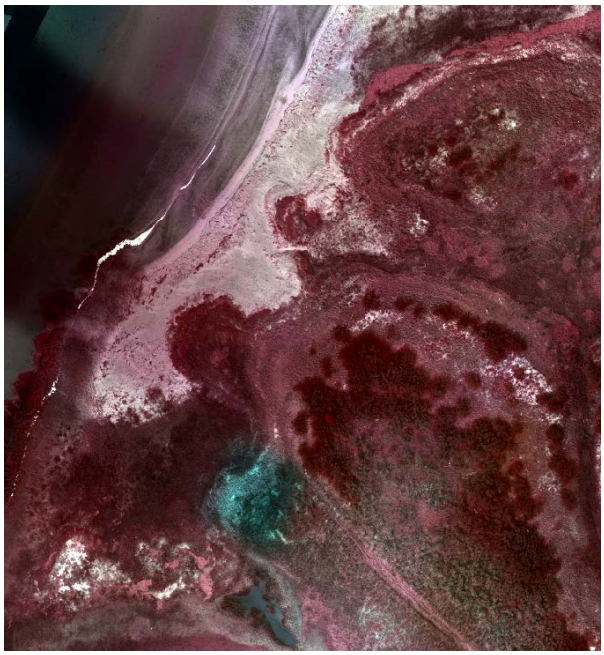
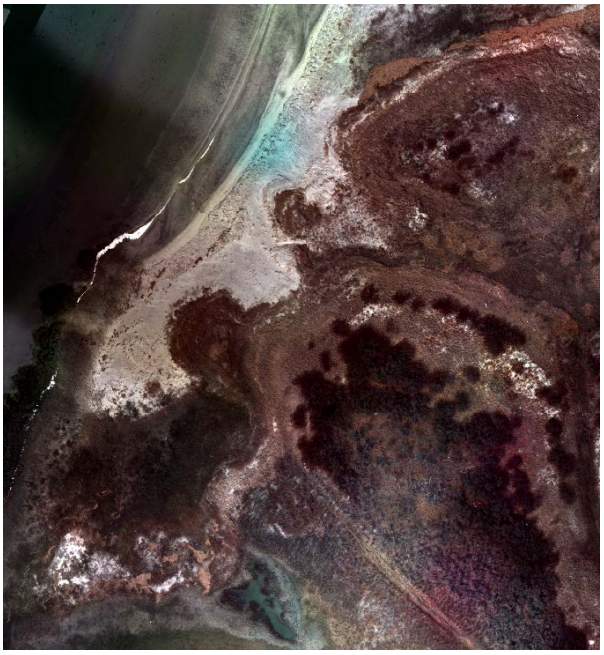
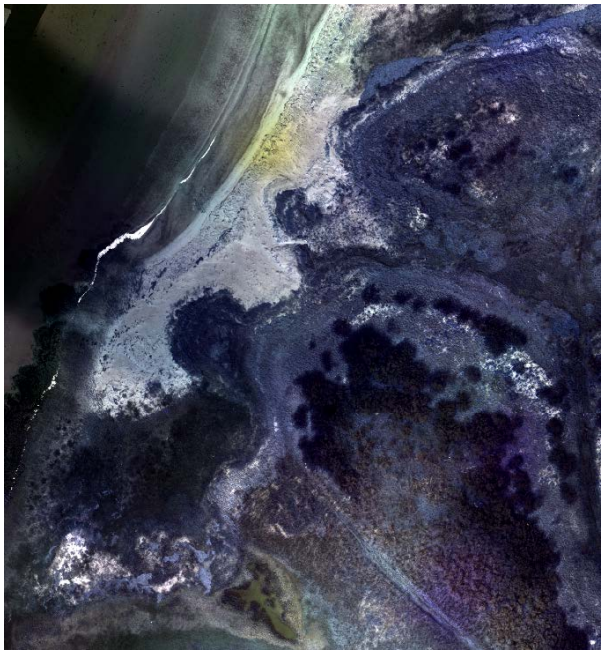
	
Area along the estuary side of the park shown with bands 4, 3, and 2. Possible false color infrared	Area along the estuary side of the park shown with bands 4, 1, and 2. Possible false color infrared
	
Possible true color using bands 3,2,1	Possible true color using bands 1,2,3

Figure 13. Comparison of imagery to determine the band order. Several spectral artifacts can also be seen in the data.

For comparison, the landscape in the field was generally sand or brown vegetation. Most water bodies were blue. There were a few evergreen trees and they pop out as expected in both possibilities of the false color infrared. Photos showing field conditions are shown in Figure 14. Consultation with the contractor confirmed that the band order is blue, green, red, NIR, red edge.

Figure 14. Photographs of the landscape illustrating the color balance and palette expected in the imagery

In general, these data are not suitable for analysis where spectral band consistency is important, such as image segmentation or deriving land cover by automated processes. It may still be useful for visual interpretation of features, though the poor spatial accuracy precludes integrating the imagery with other high-precision data sets such as lidar and field surveyed features. Because of these factors, we did not carry out the planned segmentation of the imagery, since we did not expect it to give useful results.

3 GRAND BAY NATIONAL ESTUARINE RESEARCH RESERVE

3.1 METHODS

We placed photo-id targets in each of the four areas for Grand Bay. The targets consisted of rectangular pieces of siding material measuring approximately 12 inches by 24 inches as shown in Figure 15. Twenty-one (21) targets were visible in the imagery

Figure 15. Photo-id target used at Grand Bay Research Reserve

and could be used for accuracy assessment. The targets were distributed in the four areas of interest as shown in the target placement images (Figure 16) where yellow triangles represent each target and the red polygon is the area boundary. Targets were surveyed using RTK GPS before the collection of imagery. Note that a hole was drilled in each target (not shown) and a ¾" PVC pipe was used to secure each target while allowing it to float when the tide came in.

The flights for UAS lidar used a PrecisionHawk Lancaster 5 fixed-wing unmanned aerial vehicle (UAV) with a Velodyne Puck VLP-16 based lidar system as in Jacques Cousteau Research Reserve. The imagery was acquired using the Zenmuse X5 digital camera manufactured by DJI, Inc., mounted on a DJI Matrice 100 quad-copter UAV platform. This was a change in both sensor and platform compared to Jacques Cousteau. Imagery was flown to produce a 3 centimeter ground sample distance from 50 meters above ground level. The camera provides a 16 megapixel 8-bit GeoTiff image at up to seven frames per second. Weather was favorable and the flights were carried out from May 9, 2017, to May 11, 2017.

3.1.1 Target Placement

Areas of Interest 1 - *Cladium*

Areas of Interest 2 - Upper

Areas of Interest 3 - MB

Areas of Interest 4 -SPAL

Figure 16. Images showing the placement of photo-id targets within each area of interest as yellow triangles. The imagery backdrop is not from the UAS.

3.1.2 Ground-Truth Comparisons

We determined the spatial accuracy for the imagery by locating the photo-id targets in the imagery and comparing the location based on the image georeferencing with the surveyed location of the target. For the lidar data, the lidar points classified as ground were used to create a TIN. For each ground-truth location, a value was interpolated

from the TIN to obtain the height at that point. The difference between the surveyed ground-truth and the interpolated lidar was then calculated.

3.2 RESULTS

3.2.1 Coverage

Coverage for the imagery is shown in Figure 17 for Grand Bay. During collection, we noted that area of interest 4 had a significant area of water that probably was not of interest and that capturing it added extra risk in the event of a flight problem. The principal investigator and the UAS contractors agreed not to collect data over the water and some additional terrestrial data were collected. Other than this acceptable difference, the imagery covered the areas of interest, including a buffer.

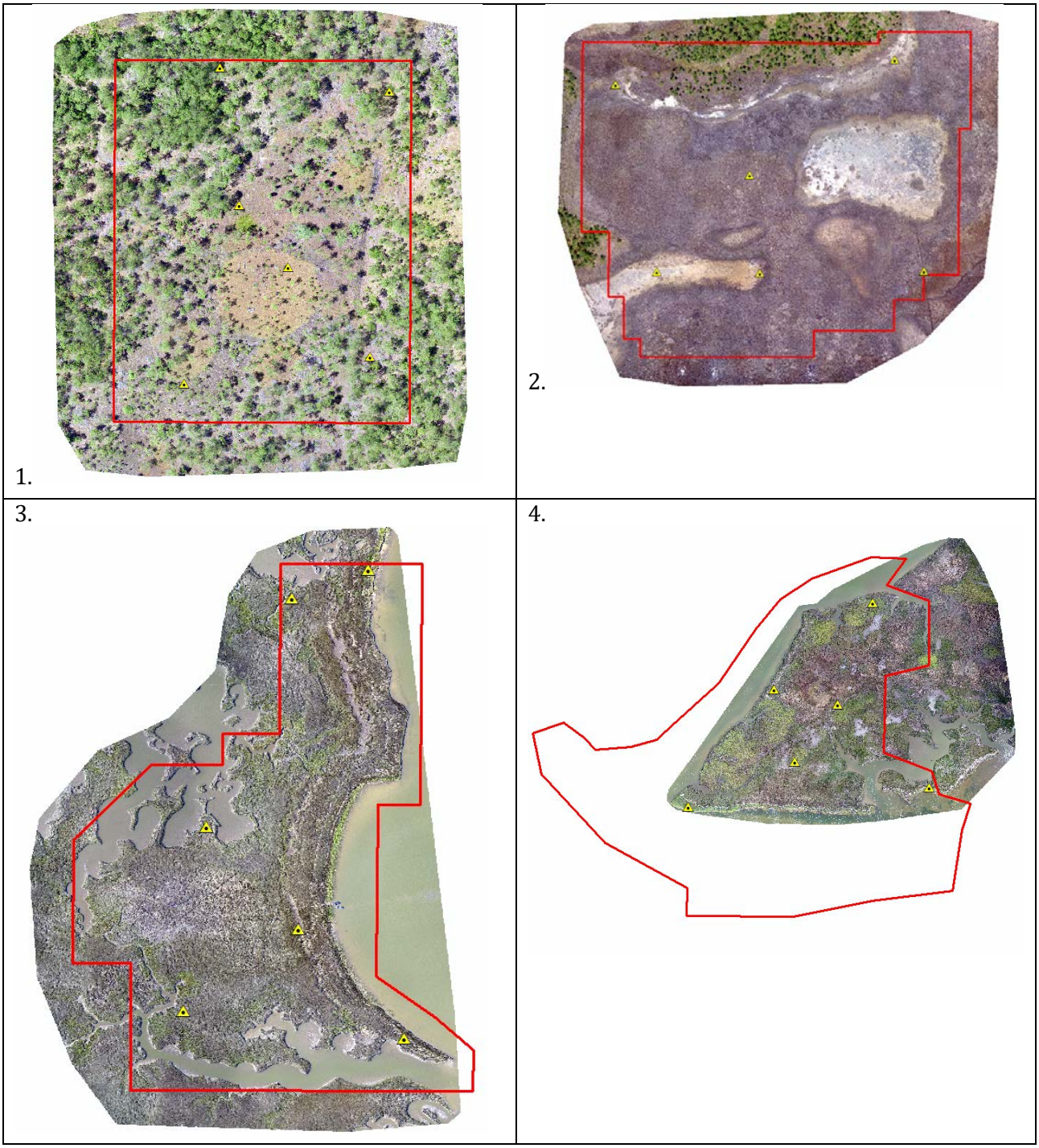


Figure 17. UAS imagery coverage for Grand Bay Research Reserve. Areas over water were not covered by mutual agreement in the field.

Lidar collection coverage is very similar to the imagery collection (Figure 18). The coverage shown in the images is the bare-earth derived DEM. Areas covered by water, such as streams, are shown as blank, illustrating that water areas have been properly classified.

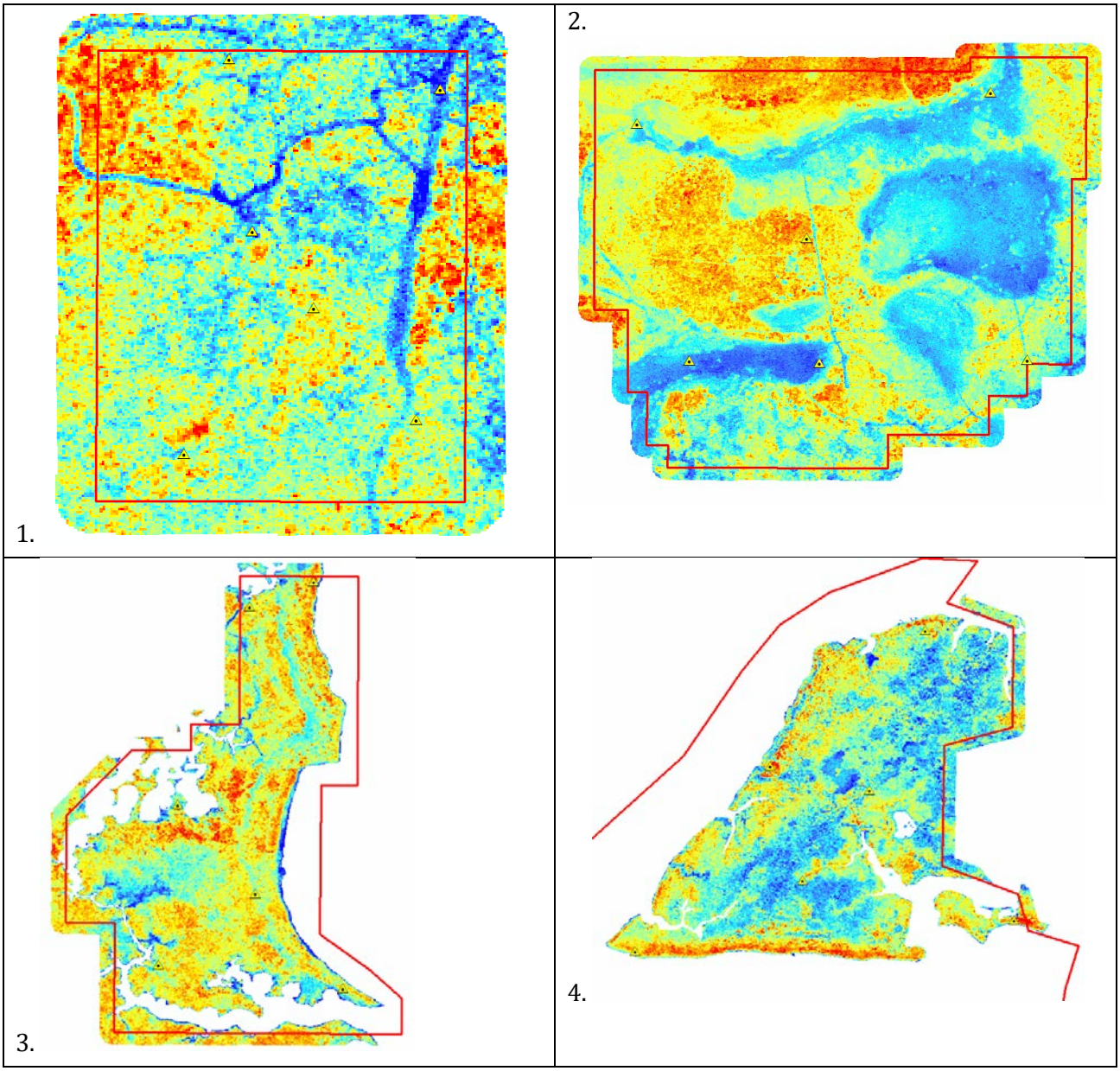


Figure 18. Lidar coverage in Grand Bay Research Reserve, with area of interest boundaries shown as a red outline

3.2.2 Imagery

The quality and accuracy of source imagery plays a key role in usability and in determining the final accuracy of imagery-derived mapping products. Through this project, we assessed the acquired UAS imagery for both spatial accuracy and image quality in order to determine its effectiveness for habitat mapping. Spatial accuracy refers to the positional agreement between real-world coordinates and image-derived coordinates of photo-identifiable features. Image quality is a subjective metric. We tested it by using the imagery as a component of a habitat classification process. Image quality refers to criteria such as effective spatial resolution (i.e., size of the smallest object that can be resolved on the ground), radiometric resolution (i.e., smallest change in intensity level that can be detected by the sensing system), and geometric consistency (e.g., relief displacement, scale, color balancing).

Lightweight UAVs are more susceptible to wind buffeting than traditional aircraft, which in turn can increase image-to-image scaling issues. These errors become readily apparent in orthomosaics, and they have a direct impact on image processing and map production. The maximum operating ceiling for small UAS's of 400 feet also imposes challenges by greatly increasing the number of individual nadir frames needed to assemble a seamless orthomosaic of a scene.

3.2.2.1 Spatial Accuracy

We tested spatial accuracy following standard positional accuracy guidelines for digital geospatial data. A minimum of 20 independent horizontal checkpoints for each project area were established to test the accuracy of the imagery. According to the contract specifications for this project, the UAS imagery was to have a resolution of 3 centimeters or better and a horizontal accuracy of 15 centimeters or better at the 90% confidence level. These specifications required the independent checkpoints to be surveyed to a horizontal accuracy of 5 centimeters or better, and in fact they were surveyed to 2 centimeters or better.

The photo-id targets used at Grand Bay Research Reserve were larger than those used at Jacques Cousteau, and the imagery had a smaller ground sample distance. The pattern on the targets showed well enough to identify the center. Examples are shown in Figure 19 where the green triangles are the surveyed center of the target displayed in the image coordinates.

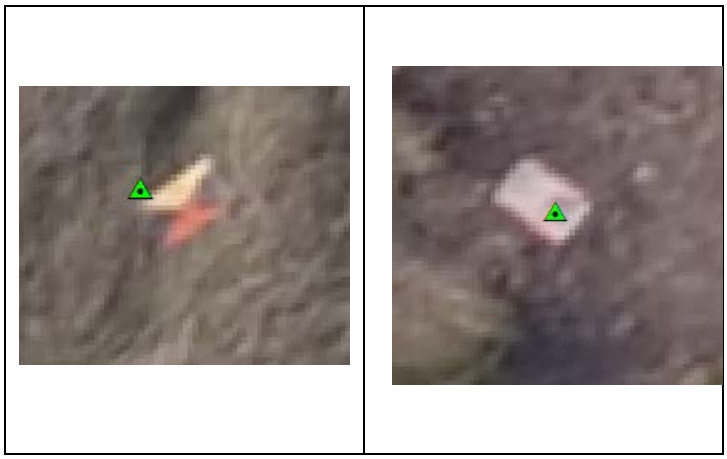


Figure 19. Examples of photo-id targets in the UAS imagery with the surveyed location of the target shown as a green triangle. If the imagery georeferencing were perfect, the green triangle would be in the center of the target.

With only 21 targets, the spatial accuracy was assessed for the entire project instead of by individual area of interest. However, it was noted that one area of interest was significantly different than the others. Area 2 (Upper) appeared as though no adjustment had been made to account for the contractor's ground control. The RMSE (n=6) was 0.84 meters, significantly higher than the specification. For all areas together, the RMSE (n=21) was 0.47 meters while excluding area 2 resulted in an RMSE (n=15) of 0.14 meters, which is within the specification. The contractor was asked to check if a mistake had been made with area 2. The contractor found that area of interest 2 was in the wrong datum (ITRF96). After correction, the overall horizontal spatial accuracy was 0.15 RMSE, which meets the contract specification.

3.2.2.2 Image Quality

The quality of the delivered color-balanced orthomosaics was generally quite good. The spatial resolution, or ground sample distance, of the UAS imagery for the Grand Bay Research Reserve mission was 2.9

Figure 20. Image illustrating the detail in the imagery

centimeters. This is equivalent to the highest resolution orthoimagery commonly acquired by aircraft. At this resolution, features that are larger than approximately 10 centimeters can be easily discerned from the background or adjacent features as long as they are spectrally differentiable (Figure 20). The Grand Bay imagery was delivered as 8-bit unsigned integer data with 256 possible pixel values per band. This contrasts with most metric mapping imagery used for classification work, which is typically 12- or 16-bit, with up to 65,536 possible pixel values.

The 4-band imagery was generated by flying the area twice, once with an RGB camera and once with a green-red NIR camera. While the registration of these is generally very good, certain features within the scene caused spectral inconsistencies and horizontal displacement problems. Objects that moved during the time between the RGB and green-red-NIR acquisitions were imaged in two different locations, causing *ghosting* effects. An example of this can be seen in Figure 21 where the boat moved between the flight times. The NIR channel (shown as red) has the boat to the right of the green and red channels (shown as blue and green).

Figure 21. Ghosting from combining imagery acquired at different times can be seen with the boat.

Similarly, objects that have significant vertical structure (e.g., trees and shrubs) caused issues in the compiled 4-band imagery. The tops of these features often appeared in different locations in the RGB and NIR-green-blue images. In most cases this was because of differences in the sensor locations and look angles between the two acquisitions. Another significant issue was the shadows cast by the tall vegetation. Due to the resolution of the imagery, shadows were prominent and distinct, and they appeared in different locations in the two sets of imagery. This impacted the classification results quite significantly, to the degree that we added an additional step to classify the shadows separately in order to model the true underlying habitat class using contextual information in subsequent steps (Figure 22). Finally, while wind was not a factor during the Grand Bay flights, it has the potential to cause motion-induced object displacement problems.

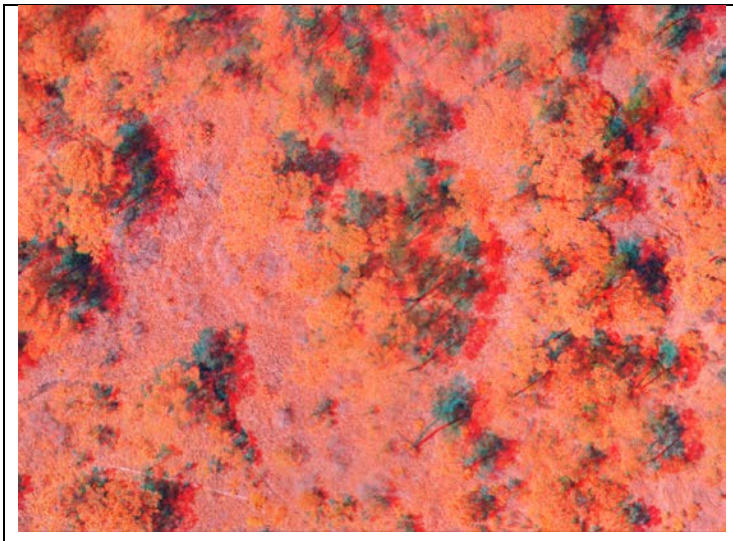


Figure 22. Sun angle change between collections affected the shadows of the combined band product.

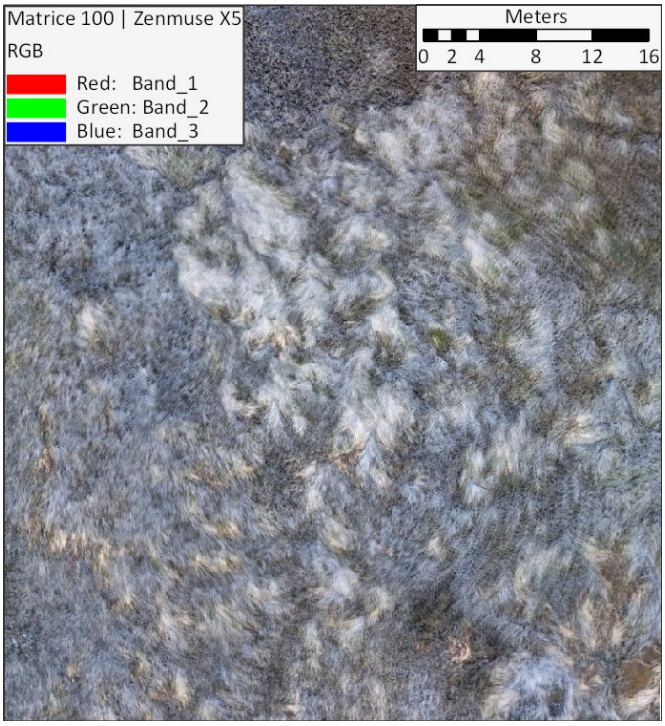


Figure 23. Smearing and blurring in the imagery that may be a result of the mosaicing and orthorectification process

UAS imagery is often characterized by artifacts that result from mosaicking the thousands of frames required to cover areas of interest. While mosaicking and color-balancing algorithms are robust and able to generate high quality image mosaics, the sheer number of frames and seamlines required inevitably lead to features appearing smeared, warped, multiplied, and omitted. In extreme cases, segments of the Grand Bay Research Reserve imagery exhibited high-frequency, radial distortions that caused significant blurring of the target features and reduced the image quality so that its effective spatial resolution was probably far less than the actual resolution of 2.9 centimeters (Figure 23).

3.2.2.3 Classification

There are aspects of marshes that make them easier to map than other environments, most notably the distinct spectral differences associated with wet versus dry features and vegetated versus non-vegetated, and the structural characteristics of tall, woody vegetation versus shorter, non-woody vegetation. Conversely, mapping specific habitats or plant species within the marsh platform (including low, mid, and high marsh classes) can be very challenging because of the similar spectral and structural characteristics of those features. Certain measures can be taken to help improve mapping results, such as using multiple dates of imagery and lidar and using imagery acquired when plants are most differentiable spectrally (i.e., leveraging knowledge of plant phenologies).

In order to assess image quality from a classification perspective, we performed a series of tests using the UAS imagery as an input to a habitat classification process. Rather than completing a final habitat map for publication, we focused on the largely automated steps that comprise the initial steps of a habitat classification process. This approach helped minimize the subjectivity associated with this type of work. The first step was to generate image segments from the orthomosaics. Image segmentation recursively groups pixels until homogeneity criteria are met, producing relatively homogenous polygons. We were then able to associate quantitative characteristics with each unique segment (e.g., length/width, spectral range and variance). These multi-dimensional "image objects" inherently characterized the imagery based on spatial and spectral patterns. Quality-controlled, field-based habitat

data were then used to classify training samples that were then employed to perform a supervised classification to generate a draft habitat map. The actual classification method included a variety of thresholding and random-forest classifiers. These steps were repeated using the WorldView 3 satellite orthoimagery previously used to generate a draft habitat map for the Grand Bay Research Reserve. This approach (i.e., using two different source images) allowed us to apply standard image processing and classification workflows in parallel using the same in situ training data to assess the trade-offs between UAS imagery and more typical satellite imagery. A comparison of imagery specifications is shown in Table 5 and example imagery is shown in Figure 24.

Table 5. Comparison of satellite and UAS imagery specs at Grand Bay Research Reserve

	Worldview 3	**UAS Zenmuse X5**
Resolution	1.2 meters	3 centimeters
Bands	8	4 (3+1)
Date acquired	May 2015	May 2017

In a typical habitat classification workflow, a draft map would be subjected to additional processing in the form of localized spatial modeling and manual editing in order to generate a final habitat map. These steps were not performed as part of this assessment. Instead, the UAS and Worldview orthoimagery-derived draft maps were compared to each other and to existing habitat maps. Areas of agreement and disagreement were determined, and a set of stratified random samples were generated for field-based investigation to allow us to compute a variety of accuracy assessment statistics. The results of these tests informed us about the quality of the UAS imagery and whether the UAS data can produce a higher quality draft map than one generated from manned aircraft orthoimagery using standardized, automated processes.

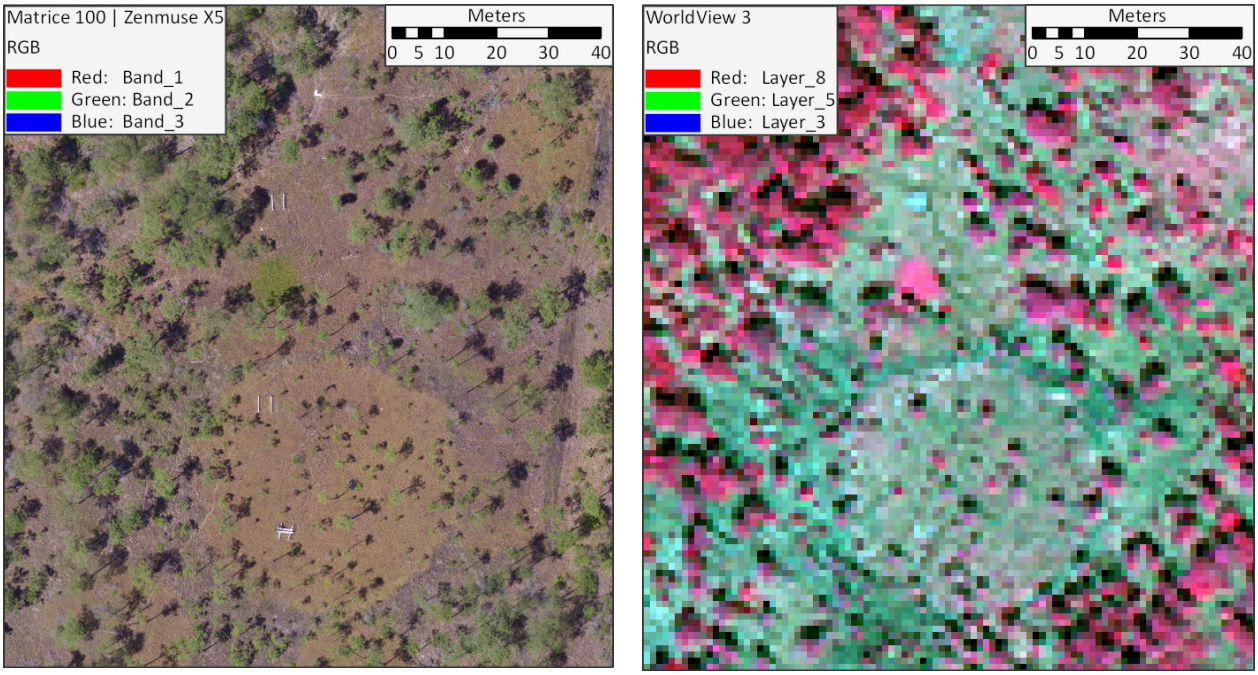


Figure 24. Comparison of UAS imagery (left) and WorldView 3 imagery (right) for the same area

3.2.2.3.1 Image Processing

Habitat classification within the four areas of interest at the Grand Bay Research Reserve was performed using a combination of image segmentation and classification routines available in Trimble's eCognition software. Segments were created using scale parameters that maximized the effective resolutions of both the UAS imagery and the WorldView 3 satellite imagery. Effective resolution was determined by qualitatively assessing the size, shape, and configuration of image segments and their use in discerning unique habitat features. This resulted in the generation of 19,662 segments for the UAS imagery and 1,341 segments for the WorldView imagery (Table 6). The large difference was driven primarily by the difference in image resolutions.

Table 6. Segment comparison between UAS and satellite imagery

Sensor	Resolution	Scale Parameter	Number of Segments
UAS Matrice 100 | Zenmuse X5	2.9 centimeters	200	19,662
WorldView 3	1.2 meters	25	1,341

In addition to the number of segments generated, the boundaries of the segments also varied (Figure 25). The difference in image resolution drove this at least partially, with the WorldView segments appearing significantly coarser when compared to the UAS segments. However, the differences extended to the actual delineation of features as well. Some boundary differences were attributed to temporal changes between 2015 and 2017 (e.g., different tide stages, transient piles of wrack, and management impacts such as controlled burns), and other boundary differences were attributed to illumination differences. The higher resolution UAS imagery exhibited much more spectral variation for given target features because of variable sun angles and differences in vegetation stature on the marsh (e.g., upright, leaning, or flattened due to wind, waves, weight, and wrack).

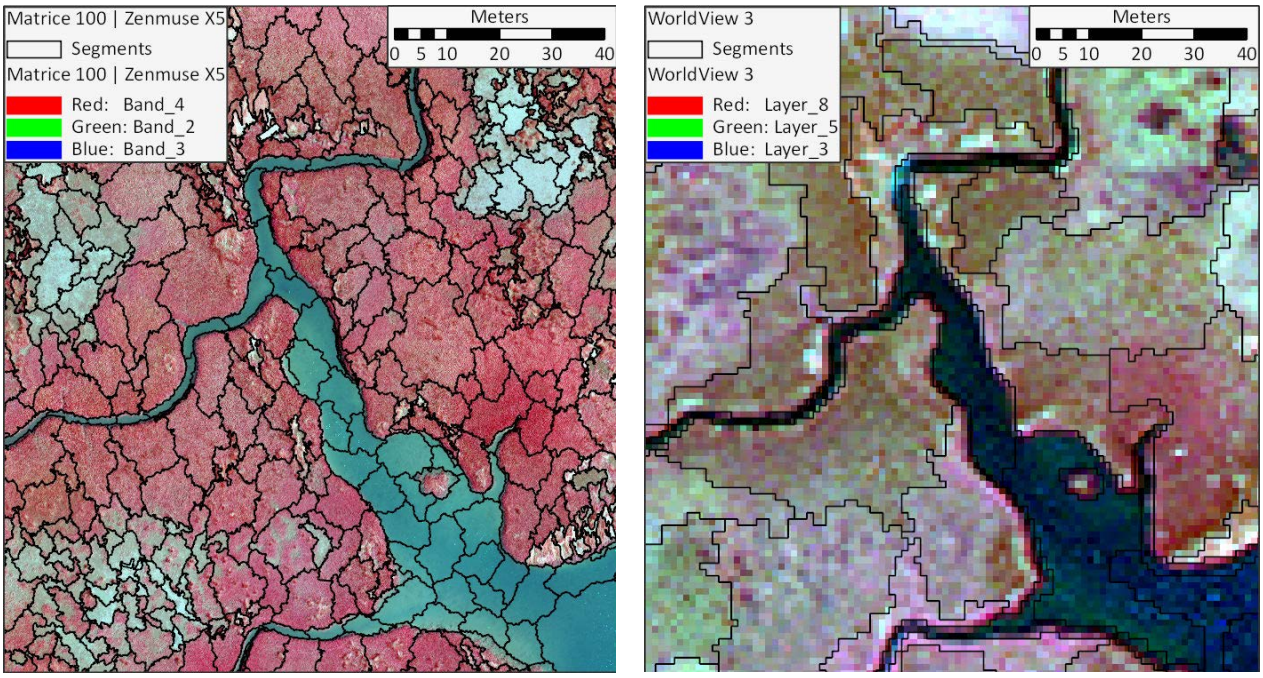


Figure 25. Comparison of segments generated shown on the imagery

The greater detail in the UAS imagery allowed much smaller features to be segmented and subsequently classified. At the scale shown here (Figure 25; 1:1,000) it is clear that the UAS segments follow distinct boundaries between

vegetated and non-vegetated areas. It is more difficult to discern the boundaries between different types of vegetation using just the spectral information. The color-infrared rendering of the WorldView imagery highlights the breaks between different types of vegetation more distinctly, despite the coarser resolution. Textural information within the UAS-derived segments may actually provide better discriminating characteristics because the imagery has the potential to capture a vertical profile of different vegetation types. For example, we noted that *Juncus roemerianus* (mid marsh) often exhibited randomly oriented cross-hatch patterns owing to the fairly rigid stems and upright-to-somewhat-leaning stature. Conversely, we noted that *Spartina alterniflora* (low marsh) often exhibited linear patterns that were fairly uniform within a segment but varied between segments. This information can be harvested using textural characterization features to help inform classifications.

Once the image segments were generated, we ran a supervised classification using a series of thresholding and random-forest classifiers to generate a largely automated, draft habitat map. We adjusted a few of the threshold parameters to better accommodate the different spectral and radiometric resolutions of the two different sets of imagery. Habitat classes were selected based on previous research reserve mapping efforts and are aligned with the National Estuarine Research Reserve System habitat classification standard (Table 7), though we only used classes that were present in the project area.

Table 7. National Estuarine Research Reserve System habitat classification table

Class ID	Class Name	Indicator Species
0	background	
1	bare (developed)	
2	beach	
3	impervious (roads)	
4	impervious (buildings)	
5	salt panne	*Salicornia virginica*
6	low marsh	*Spartina alterniflora*
7	mid marsh	*Juncus roemerianus*
8	high marsh	*Spartina patens, Distichlis spicata*
9	low shrub	
10	mid shrub	*Baccharis halimifolia*
11	high shrub	*Iva frutescens*
12	grassland	*Cladium jamaicense*
13	shrub non-marsh	*Pinus elliottii*
14	pine	*Pinus elliottii*
15	water	
16	shadow	

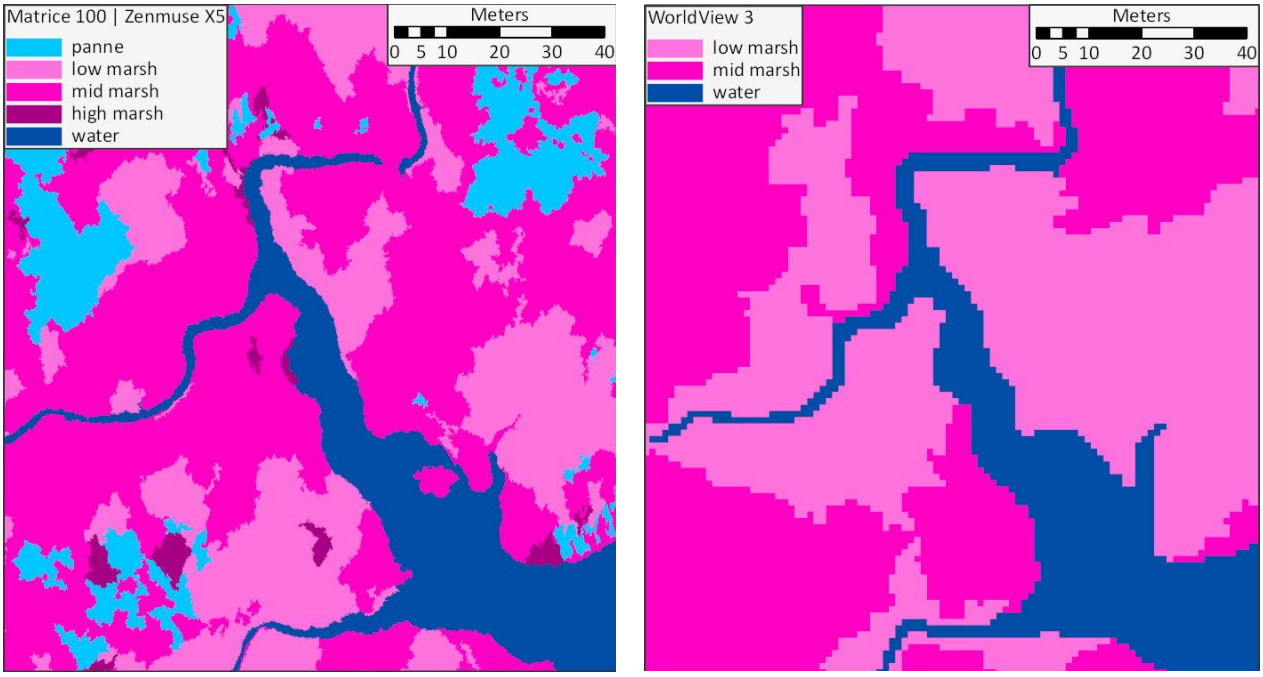


Figure 26. Comparison of initial UAS and satellite classifications

After the draft maps were prepared, we used a set of field validation points to generate confusion matrices and compute accuracy statistics to help interpret the results and understand where the two imagery sources provided value. We should note that some of the validation points were located in areas that fell below the minimum mapping area of the WorldView image segments. It is clear just from looking at the two maps (Figure 26) that a certain amount of confusion exists within the low and mid marsh classes and salt pannes. While not shown on these maps, there was also confusion between the shrub classes and marsh and grass classes. This was primarily due the lack of sufficient training data for the shrub classes.

The overall accuracy for the UAS map was 71% (kappa statistic = 0.63), and the overall accuracy for the WorldView map was 60% (kappa statistic = 0.52). These values compare quite favorably to the overall accuracy of the existing Grand Bay Research Reserve habitat map. which was 33% (kappa statistic = 0.22), although the existing map was based on 2007 imagery and we noted several significant landscape changes had occurred between 2007 and 2017.

Water mapped very well, and pine mapped fairly well but was subject to errors associated with validation points that were not directly coincident with single-crowns and instead fell on shrub or grassland classes. Interestingly, the WorldView imagery exhibited considerably more confusion between pine and shrub than the UAS imagery, but did not exhibit the confusion between pine and grassland that the UAS imagery demonstrated. This disparity can be explained by the inclusion of a shadow class in the UAS-derived map. Had the shadow class been addressed through additional modeling, many of those areas would have flipped to grassland or shrub and the observed error would be less. The emergent marsh classes appeared to map fairly well, though there was certainly confusion between low and mid marsh, and high marsh was simply underrepresented. We observed this same pattern with the WorldView map, which points to the need for more training data for these classes and the use of more aggressive classifiers, possibly including more textural and contextual information. The marsh shrub classes were also underrepresented in both the training data set and the map, leading to poor accuracy results. It should be noted that there was some confusion between upland shrub and pine and grass. This is because the upland shrub class was generated using structural characteristics and no training data were available for the class. Salt pannes were an interesting feature. The user's accuracy (false positives metric) was 98%, but the producer's accuracy (false negatives metric) was 66%. Upon further investigation, most of the confusion was between pannes and mid marsh, and was attributable to segments that wrapped around small pannes but were primarily *Juncus*. In some

cases, the validation point clearly fell in a mid-marsh segment. This could indicate horizontal displacement within the image mosaic that exceeded the measured amount of displacement.

Table 8. Kappa table for Grand Bay Research Reserve

	Ground Truth													
Classified Map	Class Value	C_3	C_5	C_6	C_7	C_8	C_10	C_12	C_13	C_14	C_15	C_16	Total	User Accuracy	Kappa
	C_3	0	0	0	1	0	0	0	0	0	0	0	1	0.00	0.00
	C_5	0	57	0	1	0	0	0	0	0	0	0	58	0.98	0.00
	C_6	0	5	51	17	1	0	1	0	0	0	0	75	0.68	0.00
	C_7	0	21	53	134	0	10	1	0	0	0	0	219	0.61	0.00
	C_8	0	1	0	0	4	0	0	0	0	0	0	5	0.80	0.00
	C_10	0	0	0	2	0	2	0	0	1	0	0	5	0.40	0.00
	C_12	0	3	0	0	1	4	53	0	8	0	0	69	0.77	0.00
	C_13	0	0	0	0	0	1	6	0	5	0	0	12	0.00	0.00
	C_14	0	0	0	0	0	0	8	0	39	0	0	47	0.83	0.00
	C_15	0	0	0	0	0	0	0	0	0	28	0	28	1.00	0.00
	C_16	0	0	0	0	0	0	0	0	1	0	0	1	0.00	0.00
	Total	0	87	104	155	6	17	69	0	54	28	0	520	0.00	0.00
	Prod Accuracy	0.00	0.66	0.49	0.86	0.67	0.12	0.77	0.00	0.72	1.00	0.00	0.00	0.71	0.00
	Kappa	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.63

Interestingly, the WorldView map exhibited significant confusion between low marsh and water, whereas the UAS map did not confuse these classes. The reason appears related to ability of the UAS imagery to resolve the fringing, semi-submerged *Spartina alterniflora,* while the water absorption in the NIR bands of the WorldView imagery overpowered the vegetation signal.

3.2.2.3.2 Classification conclusions

UAS imagery appears to provide some useful benefits for classification purposes, but introduces challenges that are not often encountered using traditional, publicly available orthoimagery. The increased detail is the primary benefit, but the ability to acquire imagery at opportune times to maximize phenological differences is also a potential benefit. Very high-resolution imagery (≤ 10 centimeters) has the potential to yield extremely detailed segments, especially when compared to the size of the mapped features. However, segments that are smaller than the resolution of incorporated DEMs exhibit variations in first surface elevations that are not consistent with the inherent spectral characteristics of the affected segments and can lead to classification errors.

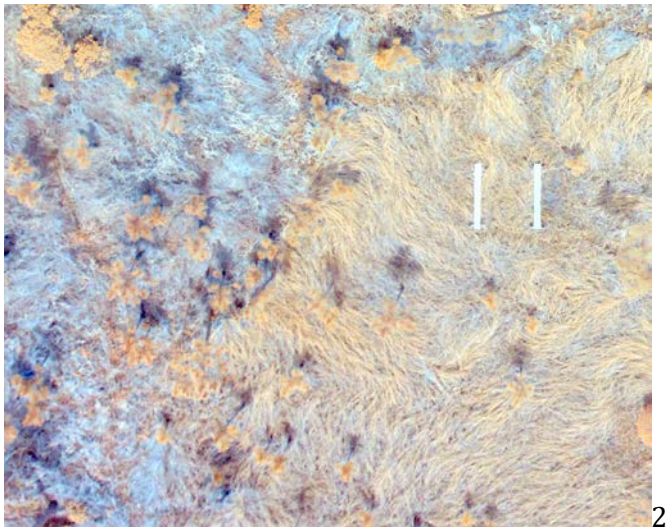
The Grand Bay Research Reserve image classification effort revealed several challenges. Acquiring the 4-band imagery using two sensors at different times was one problem and can be resolved in the future using a single multispectral sensor. We noted many instances of segments being influenced by the inclusion or exclusion of band 4 (the NIR band), especially where vertical structures (e.g., trees) caused horizontal image displacement, which reshapes segments because of shadows. This artifact introduces a great amount of error in layers that rely on the NIR band in band ratio algorithms. The normalized difference vegetation index, which relies upon the ratio between the NIR and red bands at coincident locations and is extremely useful for characterizing different types of vegetation, is skewed considerably in shadowy areas. In order to accommodate shadows, classification routines need to break them out separately and then model them to conform to the true underlying habitat. In order to generate a more accurate habitat map using this imagery, we would need to process 3D structural cover types excluding NIR band and then merge the classes in a subsequent step.

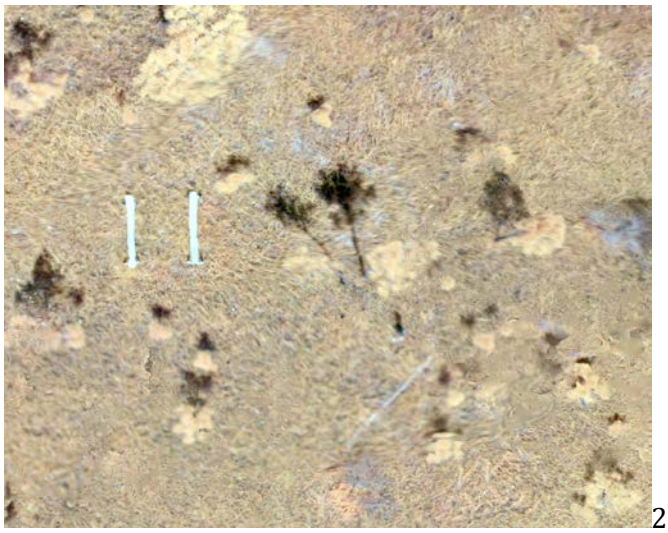
3.2.2.4 Comparison with other UAS imagery

One of our objectives was to evaluate the potential advantages of using private-sector mapping firms compared to more opportunistic acquisitions with UAS. Project partners, cooperative institutes, or any other connection could operate these opportunistic acquisitions. Obviously, we could not evaluate all possible sources of UAS imagery data. However, Grand Bay Research Reserve has had these same areas flown with imagery multiple times through a mechanism other than the private sector. To our knowledge, these data have no claim for any particular spatial accuracy and were not flown with an accuracy objective. Their accuracy should be typical of results from an experienced operator flying without ground control.

We examined data sets from 2015 and 2016, comparing them to our 2017 data. We focused our comparison on the sediment elevation table (SET) benches and boardwalks, because they are features that should be stationary in the environment and would not have seasonal changes. These also provided features of known shape and properties that help to detect imagery issues. We did not do an exhaustive examination of the imagery looking for faults but simply compared the areas around some of the known features.

In Figure 27, we can see the benches around a SET in the imagery from three consecutive years. All of the benches should be in the same place if the images were accurately georeferenced. However, we see that they are not. Since we did photo-id points and measured the accuracy of the 2017 data, we have faith that it is in the right place. For 2015 and 2016 data, we see that they do not agree with the 2017 data, and they do not agree with each other by an even greater degree. The misalignment of the data is such that automated processing for change detection would be very difficult to impossible unless a way was found to correct it. Note that we did not do anything to try to correct or re-mosaic the earlier images. They are simply as delivered to Grand Bay Research Reserve.


2015


2016


2017

Figure 27. UAS imagery from 2015, 2016, and 2017 showing the same benches around a SET. Image coordinates are aligned such that these images have the same corner coordinates.

We extracted locations for identifiable points on the boardwalks to get a more quantified look at the differences, with some points in each of the four areas. All year-pairs were compared and the offsets were plotted in x and y directions (Figure 28). While there are not sufficient points to make many conclusions, it does appear that there

are large differences (typically meters) between positions and the degree of difference varies greatly within the images for any given pair of years.

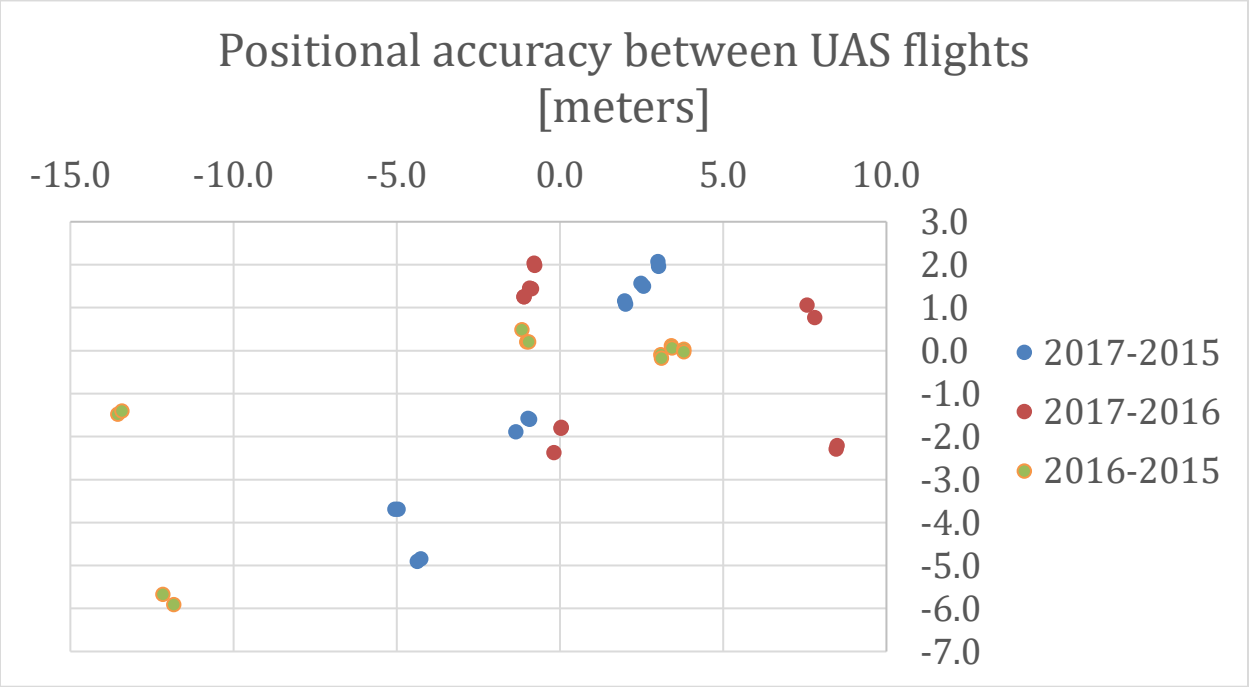


Figure 28. Positional offsets of identifiable points on the boardwalks

We also looked for evidence of issues with the mosaicking process. We see discontinuities in the boards for the 2015 and 2016 data. We did not try to determine if the cause of these image problems was from the original image positioning or the mosaicking process itself. Figure 29 shows a couple of examples of the artifacts in the 2015 and 2016 imagery. The boards, which should be straight, have discontinuities and replication. Smearing is seen in several areas of the images. We also see some color issues, such as the blue area at the north end of the 2015 board. In contrast, the 2017 imagery primarily had problems related to flying the NIR band at a different time of day, but discontinuities were not evident.

Figure 29. Images with artifacts from 2015 (left) and 2016 (right)

Overall, the private-sector flown imagery appears to be a superior product. It was better in terms of both image quality and positional accuracy.

3.2.3 Lidar

Lidar accuracy for non-vegetated areas had an RMSEz of 0.072 meters using 106 points (Figure 30). This meets the contract accuracy specification of 0.10 meters RMSEz. Errors are not normally distributed, in contrast to our expectation for non-vegetated areas. The scatterplot of the lidar and survey NAVD88 elevations shows a very narrow range of less than a meter. This is to be expected for the low-lying marsh. The histogram of errors shows a skew to the high side (Figure 31).

The point density for lidar was very high, as expected. There were 105 points per square meter on average for the area and slightly over 3 points per square meter were classified as ground. However, wet or dark ground had lower return rates. This is consistent with manned lidar systems using NIR lasers. Materials that are dark at the frequency of the laser are much less likely to return a signal sufficient to be counted as a return. NIR lasers typically have trouble getting a return signal from wet surfaces because the water has low reflectivity and high absorption, reducing return signal strength. In Figure 32 you can see such an area of dark or wet ground. On the left are the lidar returns colored by elevation. The white areas have few returns. On the right is the imagery for the same area where you can see that the low return areas correspond to the dark ground.

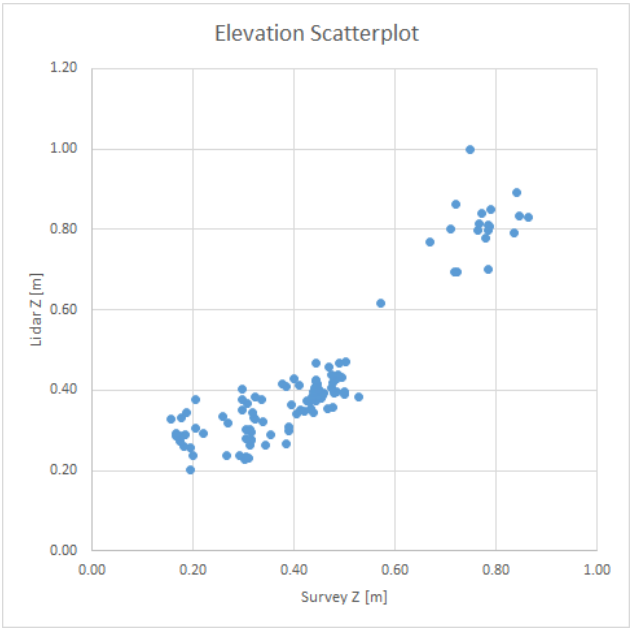


Figure 30. Ground truth versus lidar elevations in Grand Bay Research Reserve

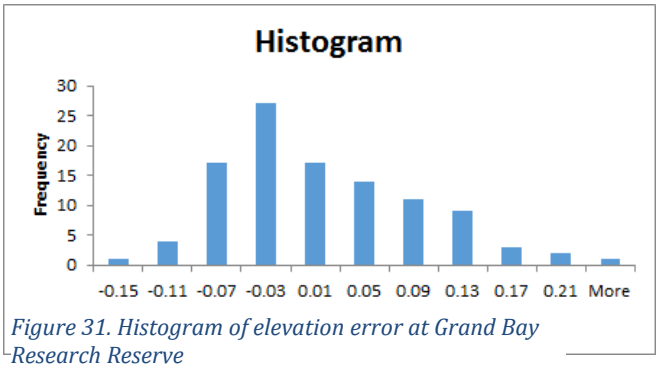


Figure 31. Histogram of elevation error at Grand Bay Research Reserve

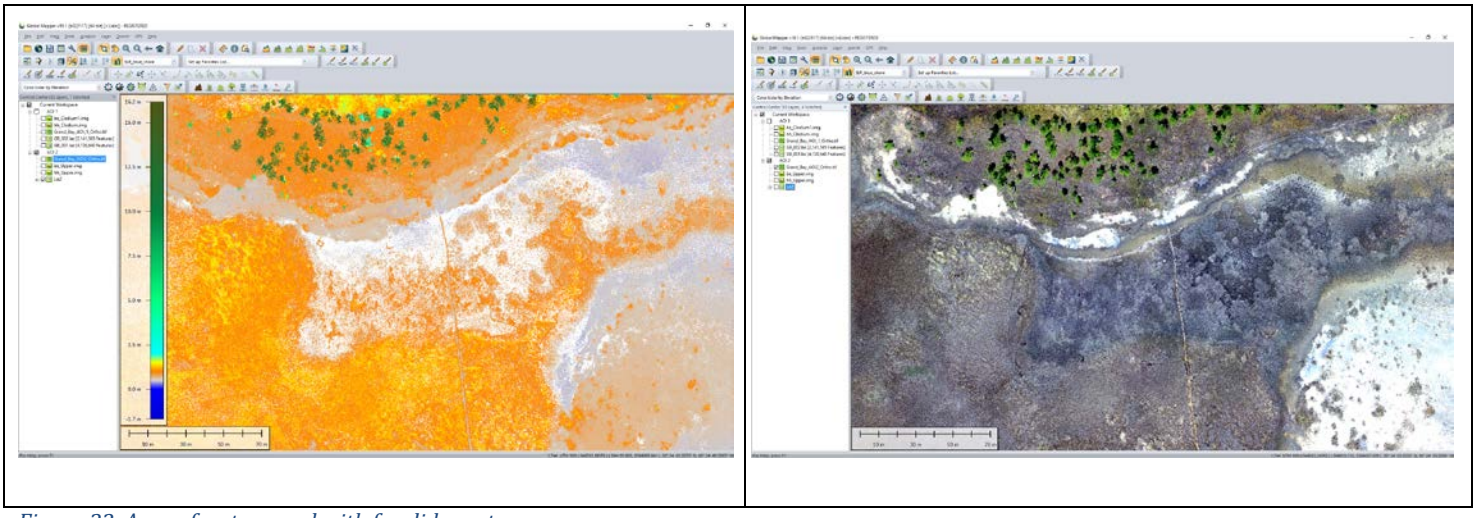


Figure 32. Area of wet ground with few lidar returns

In the vegetated areas, the lidar accuracy results can be split into vegetation classes (Figure 33). We generally see the pattern in the RMSEz that we also see with manned aircraft. The thicker marsh grasses have RMSEz values in the neighborhood of 20 to 25 centimeters.

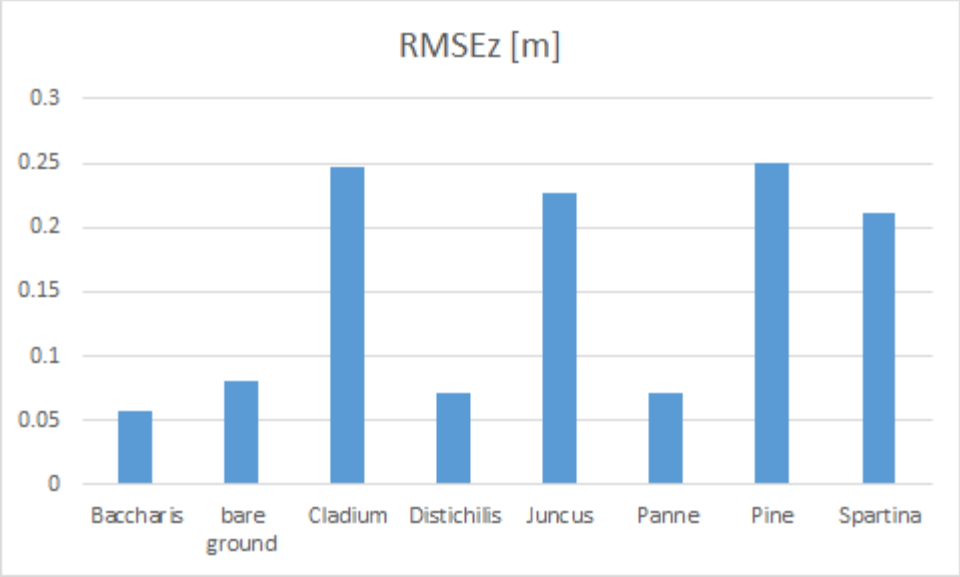


Figure 33. Lidar vertical root mean square error in different vegetation types (Grand Bay Research Reserve)

The bias in thick marsh grasses is a known issue with manned lidar. We hypothesize that returns from the grass within about 30 centimeters of the ground result in the system being unable to respond to the actual ground return. A combination of pulse length and system response and recovery time preclude the ability to resolve elevation differences at that scale. The lidar returns in the grass are often such that the lowest returns look like ground to the automated classifiers and are misclassified. One of the tests for this research was whether the smaller spot size on the ground would allow the UAS lidar to penetrate to the ground where there were holes in the cover. Looking at the mean error by vegetation type (Figure 34), we see that the RMSEz error is primarily driven by the mean error for the marsh grasses (*Cladium*, *Juncus*, and *Spartina*).

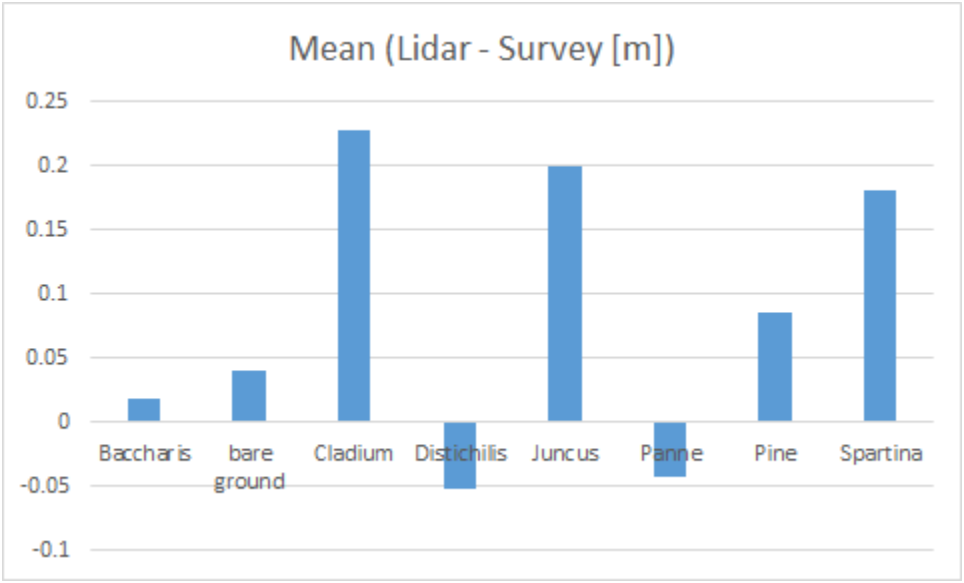


Figure 34. Lidar mean error in different vegetation types

The horizontal accuracy of lidar is generally more difficult to determine. We can use some of the boardwalks in the imagery to make some estimate though. We have already found the imagery to have a radial RMSE of 0.15 meters. Comparing the locations of the boardwalks in imagery and lidar suggests approximately 0.7 meters shift in the north-south direction and 0.4 meters shift in the east-west direction (Figure 35).

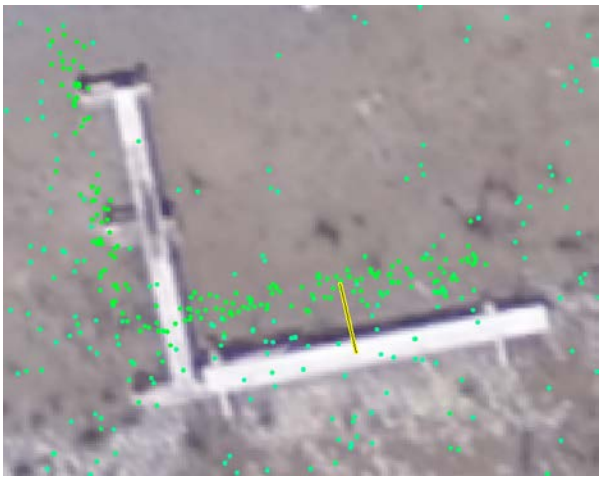


Figure 35. Boardwalk in area of interest 2. Returns from the wet mud around the planks made returns (green dots) scarce and the boardwalk more apparent in the lidar.

3.2.3.1 Comparison with manned lidar

Lidar covering Grand Bay Research Reserve was flown from manned platforms in 2015 and 2005. The two data sets from 2005 are before or shortly after Hurricane Katrina, and it is likely that there have been significant changes in the landscape since that time. Therefore, we will only compare results to the 2015 lidar data collected for the Mississippi Department of Environmental Quality and the U.S. Geological Survey. The lidar was collected and processed to meet a maximum nominal post spacing (NPS) of 0.7 meters. The NPS assessment is made against single-swath, first-return data located within the geometrically usable center portion (typically about 90%) of each swath. The data were collected using a Leica ALS70 500 kHz Multiple Pulses in Air (MPiA) lidar sensor and a Chiroptera II sensor. The ALS70 sensor collects up to four returns per pulse, as well as intensity data, for the first three returns. The Chiroptera II sensor provides 35 kHz bathymetric data and up to 500 kHz topographic data.

The 2015 data were tested for accuracy using 60 non-vegetated survey points and 50 vegetated survey points. The non-vegetated accuracy was 0.166 meters at the 95th percentile (0.085 meters RMSEz) and the vegetated accuracy was 0.168 meters at the 95th percentile. It is unlikely that any of the validation points were in the Grand Bay Research Reserve areas of interest because the data set covers almost 6,000 square miles.

While the non-vegetated accuracy results for the UAS and the 2015 lidar are very similar (0.72 and 0.85 meters, respectively), we are most interested in whether the UAS has an advantage in the vegetated areas. We do not have information on the vegetation types used in the original 2015 vegetated accuracy analysis, and it is unlikely that they represent marsh habitat well. Instead, we will use the UAS ground-truth points for both, with some caveats. Table 9 and Figure 36 show the RMSEz comparison.

Table 9. Manned versus UAS lidar in vegetation

Veg Type	RMSE 2015	RMSE UAS	Mean 2015	Mean UAS	Samples
Baccharis	0.23	0.06	0.20	0.02	17
bare ground	0.37	0.08	0.33	0.04	55
Cladium	0.57	0.25	0.56	0.23	97

Distichlis	0.12	0.07	0.10	-0.05	6
Juncus	0.36	0.23	0.35	0.20	155
Panne	0.05	0.07	0.03	-0.04	51
Pine	0.24	0.14	0.22	0.11	66
Spartina	0.19	0.21	0.17	0.18	98

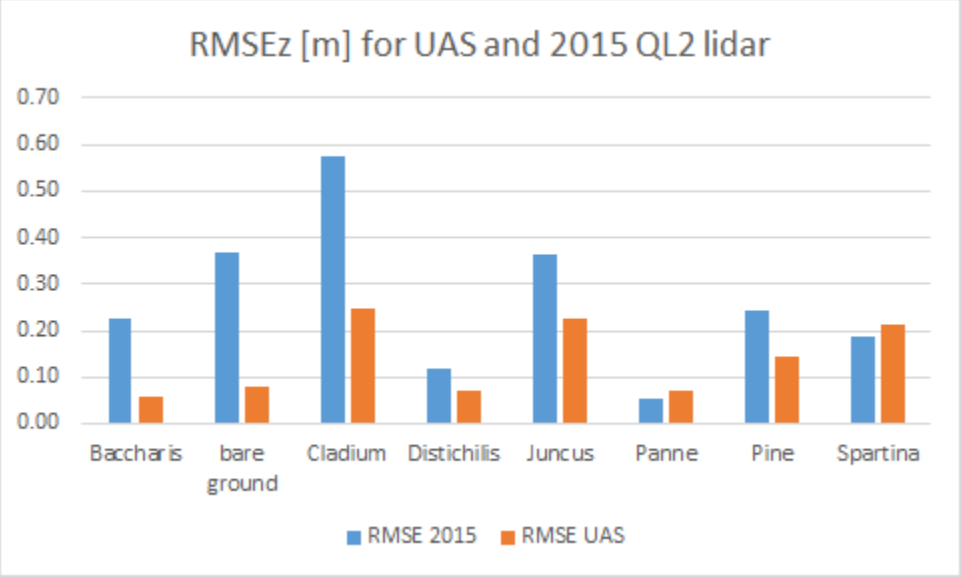


Figure 36. Manned (2015) and UAS lidar error in different vegetation types

In general, we see the UAS has a lower RMSEz than the 2015 lidar. One difference that stands out and needs explanation is the high RMSEz for bare ground in the 2015 lidar, far higher than the 2015 ground-truth survey estimates. An examination of imagery from 2015 and 2017 indicates that many of the 2017 bare ground points were not bare ground in 2015 or may be a result of lower point density in 2015. We also see consistent mean bias values for *Cladium*, *Juncus*, and *Spartina*, each showing roughly 20 centimeters too high. Removing 20 centimeters from the lidar ground points in those classes results in RMSEz values of 10-11 centimeters.

From these findings, we conclude that for the type of marsh found in Grand Bay Research Reserve, the UAS-based lidar does provide an advantage over traditional manned lidar. The RMSEz values are significantly reduced, and it appears there may be a bias that could be removed from the ground points using an appropriate habitat map.

3.2.4 Structure from Motion

The contractor generated a point cloud from the imagery using structure from motion (SfM) techniques. The point cloud created via SfM is always a first surface model and only represents the ground in bare areas. We wanted to look at two aspects of the SfM output: 1) comparison with the lidar first surface; and 2) if more consistent than lidar, whether we could use it with the lidar bare-earth to generate better canopy heights.

Preliminary examination of the SfM point cloud in area of interest 1 (Figure 37) indicated that it had spurious artifacts. In particular, there were impossibly low points that were also away from the area of interest. After

removing those points, the point cloud still had significant gaps and many points that were much higher than expected (Figure 38).

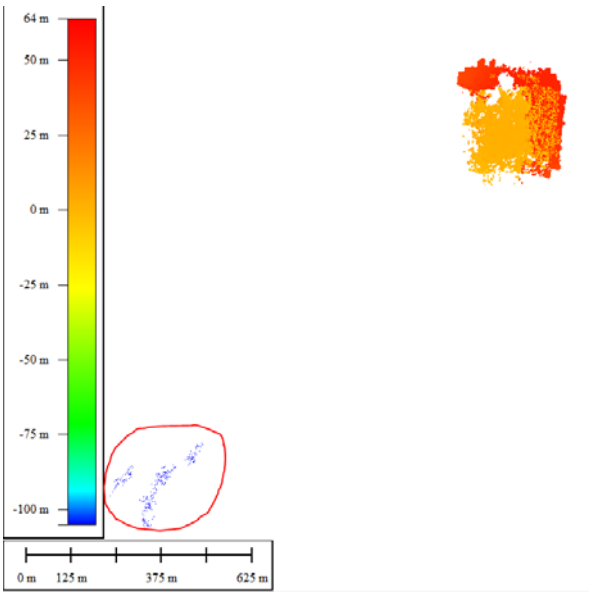


Figure 37. SfM point cloud for area of interest 1. Noise points are circled in the lower left.

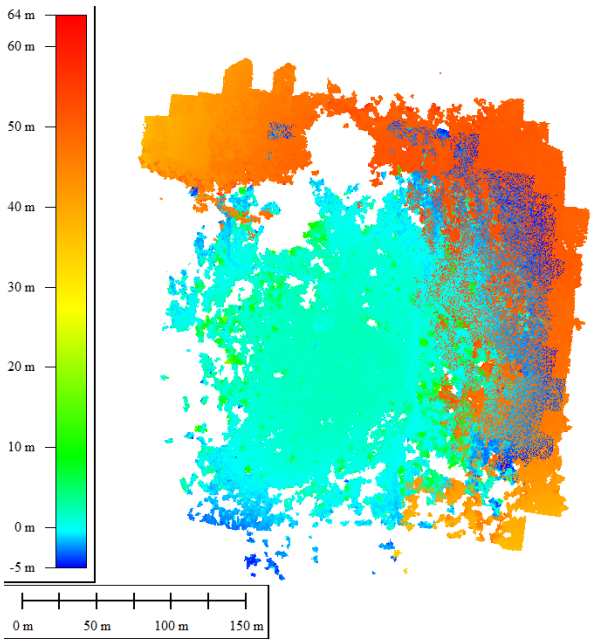


Figure 38. SfM point cloud for area of interest 1 after removing points below -5 meters

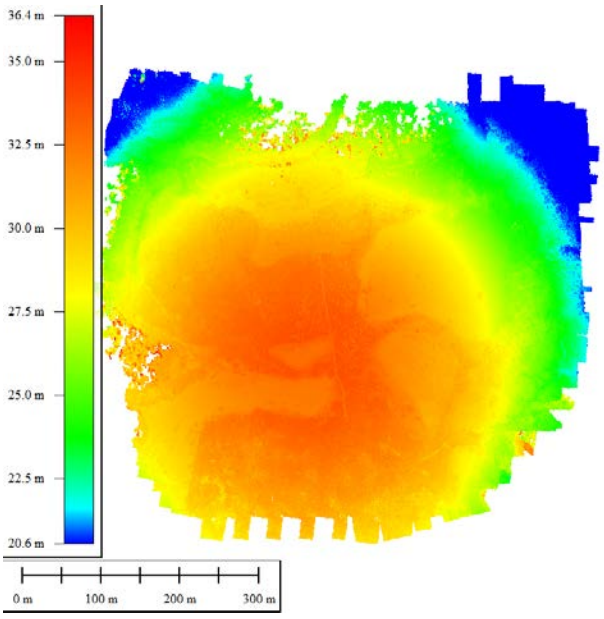


Figure 39. Area of interest 2 SfM didn't have the noise problem, but the pattern and the elevation range (20.6 to 36.4 meters) didn't make sense.

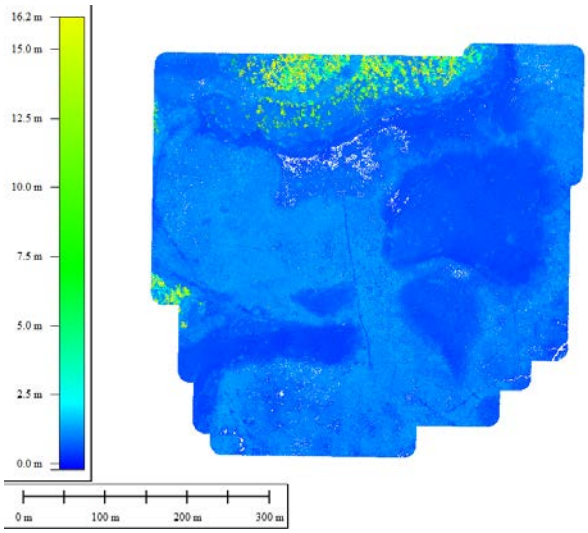


Figure 40. Lidar for area of interest 2 with all points. Some similar structures can be seen toward the center. The elevation range (in NAVD88) makes much more sense.

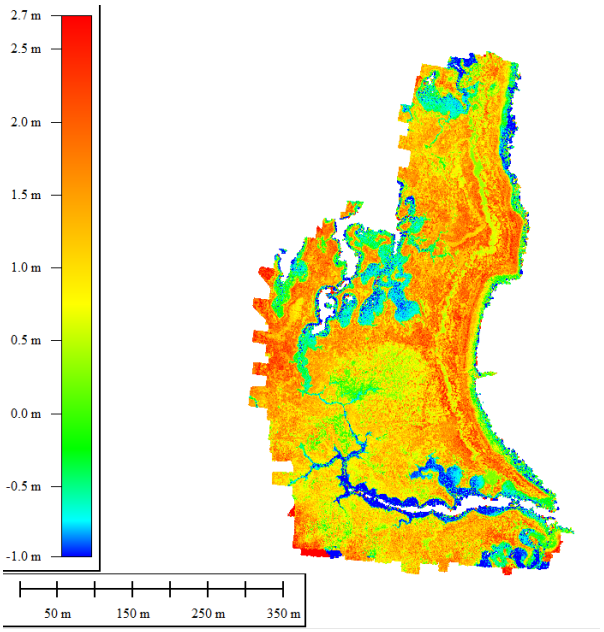


Figure 41. Area of interest 3 SfM. In this area dominated by marsh grass, the elevations are more sensible than area of interest 2, but still high since nothing was above our heads.

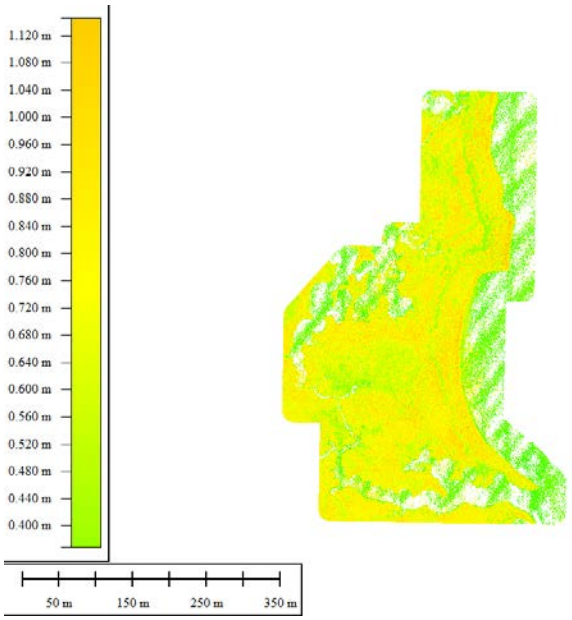


Figure 42. The lidar for area of interest 3 shows a smaller range than the SfM data. These values are in good agreement with our field experience.

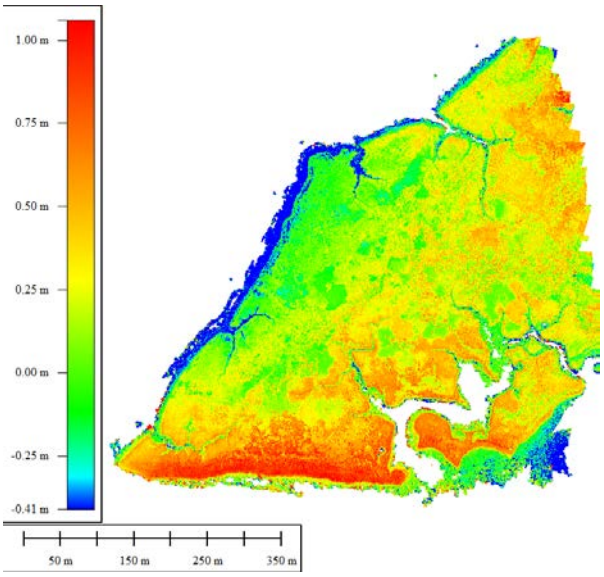


Figure 43. The SfM for area of interest 4 appears the most reasonable of the areas. In this area, MLLW is 11 centimeters above 0 NAVD88, thus -0.41 is unlikely and much of the area in green is shown as below mean sea level.

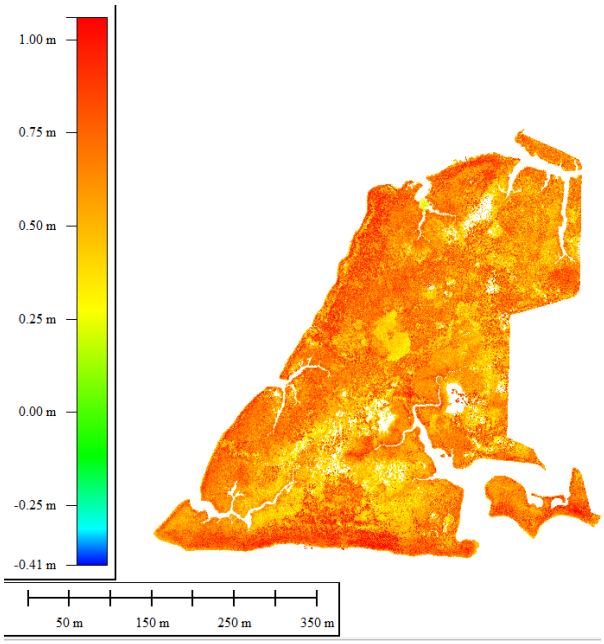


Figure 44. Lidar elevations for area of interest 4 shown on the same scale

The differences seen between the lidar and the SfM data (Figures 37-44 above) indicate that attempts to create a canopy model between the two would not be fruitful. In general, the SfM data were not believable for the terrain mapped. The issues with the SfM were different for each area of interest, suggesting the problems are related to the landscape and not the processing approach.

4 SAN FRANCISCO BAY NATIONAL ESTUARINE RESEARCH RESERVE (RUSH RANCH)

4.1 METHODS

For the final field campaign at Rush Ranch, adjustments were made based on what we had learned at the Jacques Cousteau and Grand Bay Research Reserves. The photo-id targets were made significantly larger as 18" square pieces of ⅛" thick yellow foamed PVC (Figure 45). Black Duck tape was used to create a checker pattern. Holes were drilled in the corners to allow staking the targets; a hole at the center was drilled to set the point of the RTK survey rod. Thirty new photo-id targets were distributed in the area of interest, approximately half in the upland and half in the marsh. One of the old targets was also deployed.

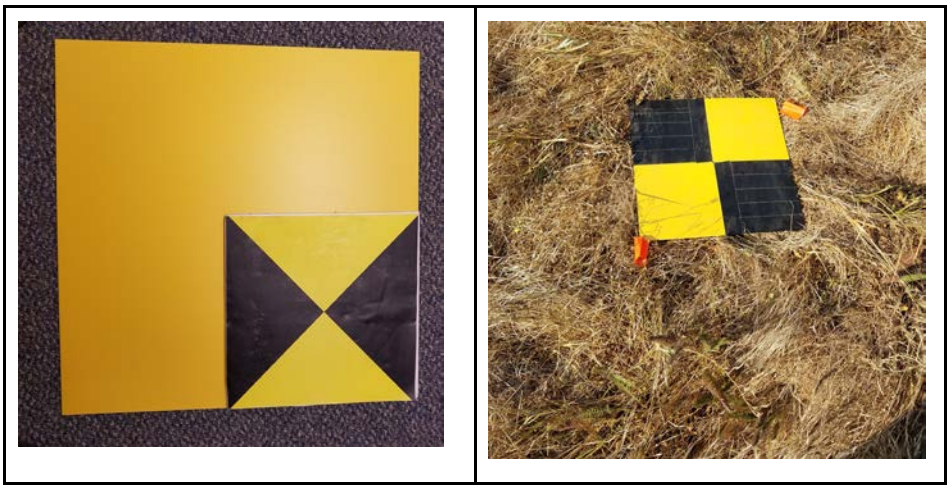


Figure 45. The image on the left shows the size of the targets used at Jacques Cousteau compared to the new targets before Duck tape patterning. The image on the right shows the final targets as deployed in the uplands at Rush Ranch.

The vegetation at Rush Ranch generally prevented affixing a target flush to the ground without damaging the marsh. Therefore, we anchored the targets horizontally, but floated on top of the lowest vegetation we could find. We did not use them for lidar elevation assessment. The distribution of the targets is shown below (Figure 46). There were areas in the marsh that were not accessible due to large water channels, particularly in the southern area. In addition, we could only enter the western side at one area near the middle (where three points are clustered). This resulted in a very long hike in 114°F heat. To mitigate the potential danger of having the contractor also make that hike, we donated one of the photo-id points to them. It is the orange pentagon in the target placement image.

An additional test element was added at Rush Ranch to evaluate the resolving power of the imagery. While the ground sample distance provides the size of the pixels, it does not necessarily indicate the size of objects that can be differentiated. We printed a 1951 USAF resolution test chart on an 4-foot x 4-foot vinyl banner and staked it to the ground. Unfortunately, it appeared that the largest lines were approximately 1 inch wide and unlikely to be resolved. We added strips of black fabric above the chart to expand our resolution range. The fabric strip sets were 4, 8, and 12 centimeters wide. The deployed chart is shown in Figure 47.

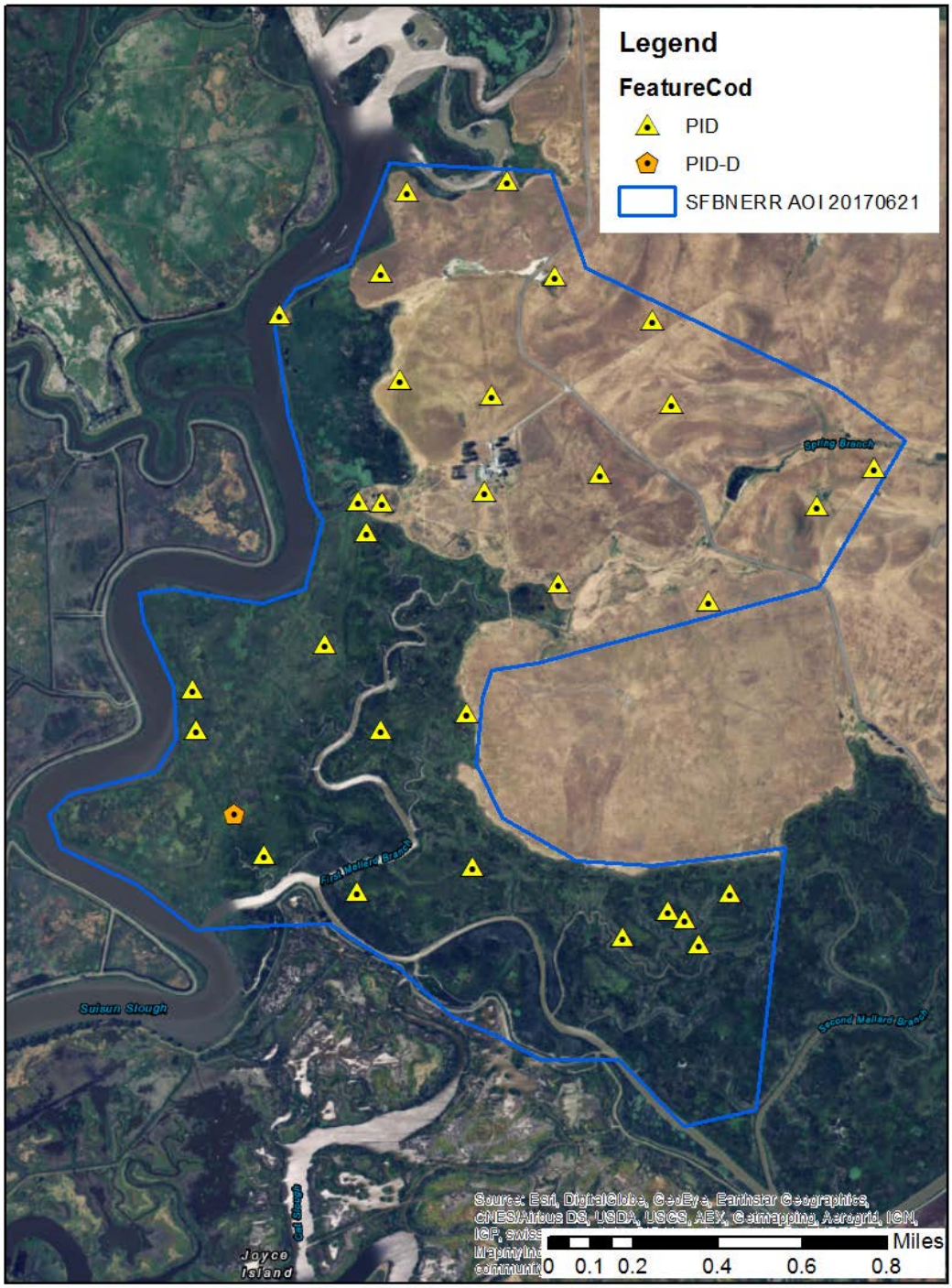


Figure 46. Photo-id placement at Rush Ranch

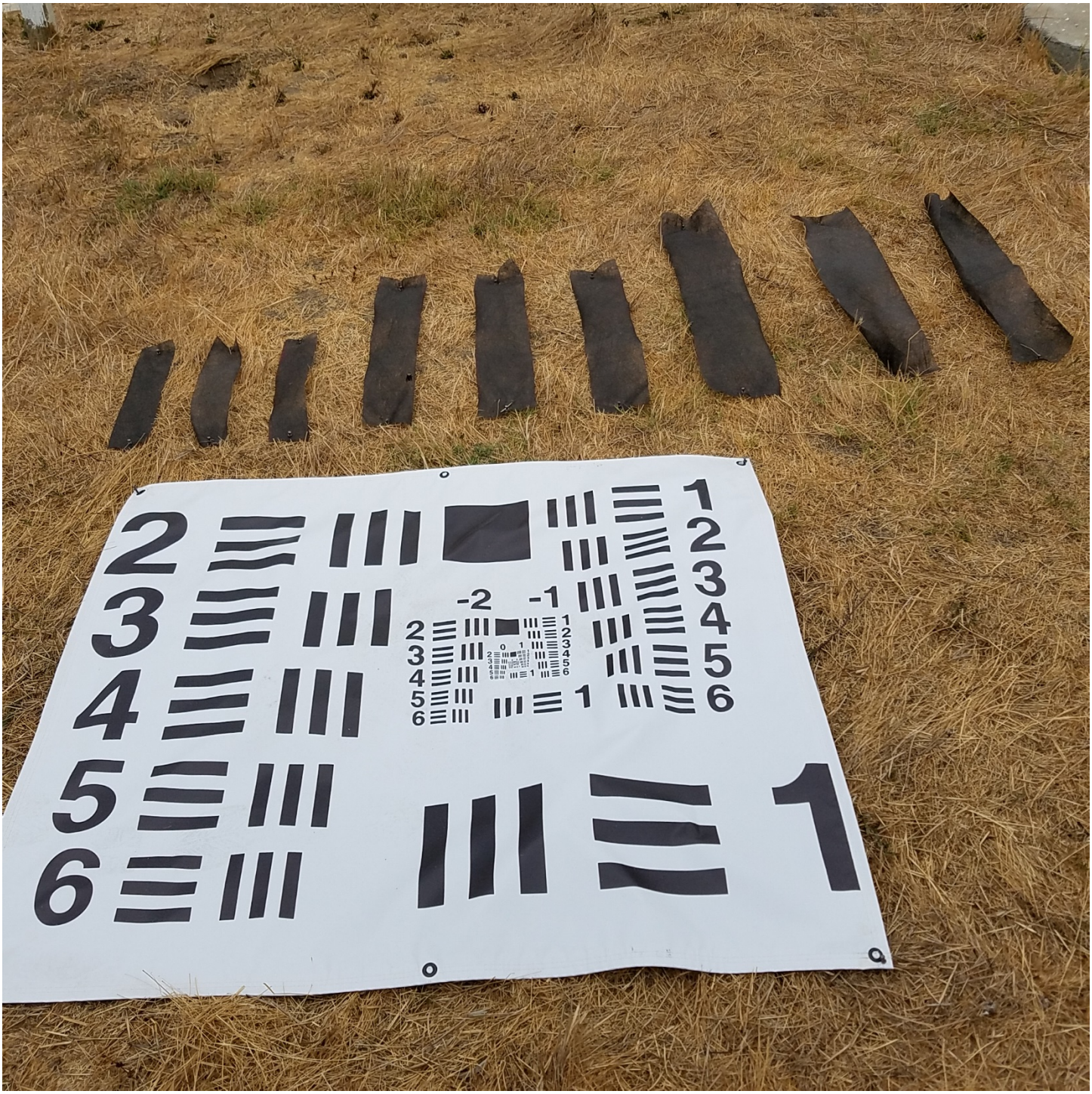


Figure 47. USAF-1951 resolution chart with additional resolution strips at 4, 8, and 12 centimeter widths

4.2 FLIGHTS

PrecisionHawk changed platforms and sensors for the Rush Ranch collection. The platform for both imagery and lidar was the DJI Matrice 600 hexacopter. The imagery sensor was the MicaSense 5-band used at Jacques Cousteau, and the lidar sensor was a YellowScan systems integration of a Velodyne Puck. With a flight ceiling of 200 feet imposed by the FAA certificate of authorization, the imagery collection was expected to take much longer than the lidar collection. PrecisionHawk brought two platforms and flew them simultaneously, one with lidar and one with imagery. Once the lidar was complete after three days, both hexacopters flew imagery to complete the mission faster. Flights began September 5 and completed September 10, 2017.

Over 380,000 images were collected, four times more than originally planned because of the airspace restriction.

We were able to keep the drones in view up to 1.5 kilometers away. Typical flight times were about 20 minutes in duration with about 5 minutes between flights for battery change and data USB swap. The high heat during the flights required extra fire precautions as the take-off and landing sites were in the dry uplands. They kept the generators for charging the batteries in the back of a pickup or placed them on plywood to reduce the risk of fire.

Using the MicaSense camera instead of the ZenMuse was a decision made by the NOAA and National Estuarine Research Reserve System team. The MicaSense would give 5 bands collected simultaneously, but would not meet the resolution specification (4 centimeter instead of 3 centimeter ground sample distance). We deemed this preferable to combining images from two 3-band cameras with non-simultaneous data as seen in Grand Bay.

4.3 RESULTS

4.3.1 Coverage

Lidar coverage is shown below as a density map for all points and for ground classified points. The entire area of interest was covered. Water areas have lower point density as expected. The lidar data were very dense, averaging over 500 points per square meter and exceeding 1,000 points per square meter in places. The density of ground-classified points shows a high number of points in the uplands, in the 20 to 30 points per square meter range, with lower penetration in the marsh. However, there were patches in the marsh that also approached 20 points per square meter, and areas where no points were deemed to have penetrated to the ground.

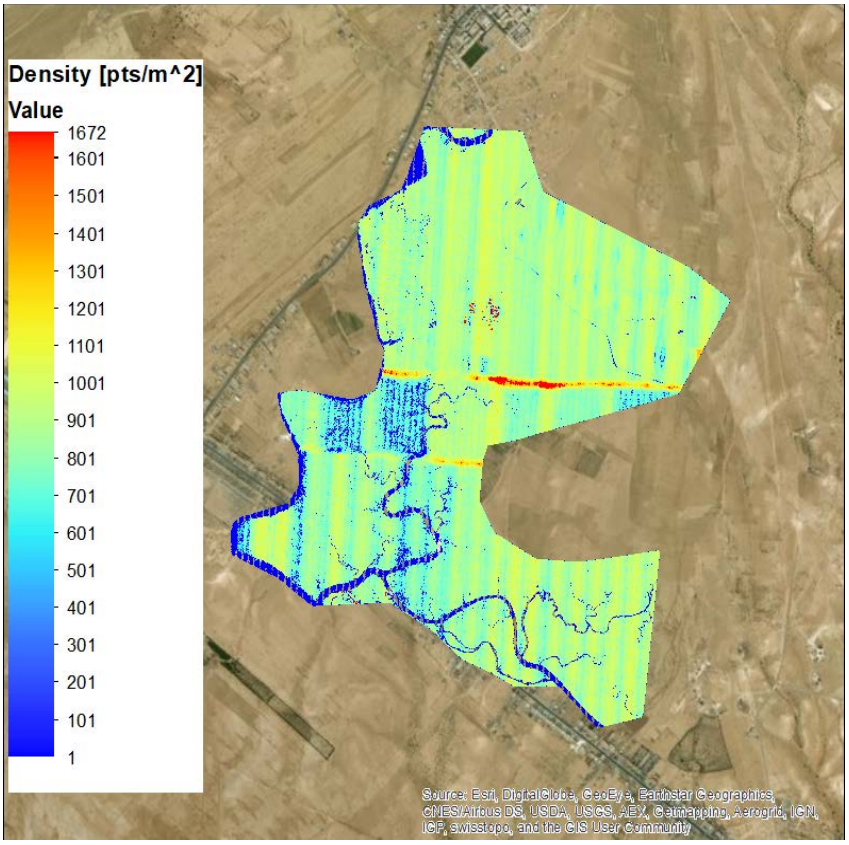


Figure 48. Density map for all lidar points

Imagery coverage extended beyond the bounds of the area of interest but had some holes in it. The holes were in the hilly areas and the rapid relief may have led to the problems. This was most noticeable in the north. The river on the west side also had holes, but it is outside the area of interest. The NoData value was not designated in the files but appears to be a zero in all bands. Real data did not conflict with this.

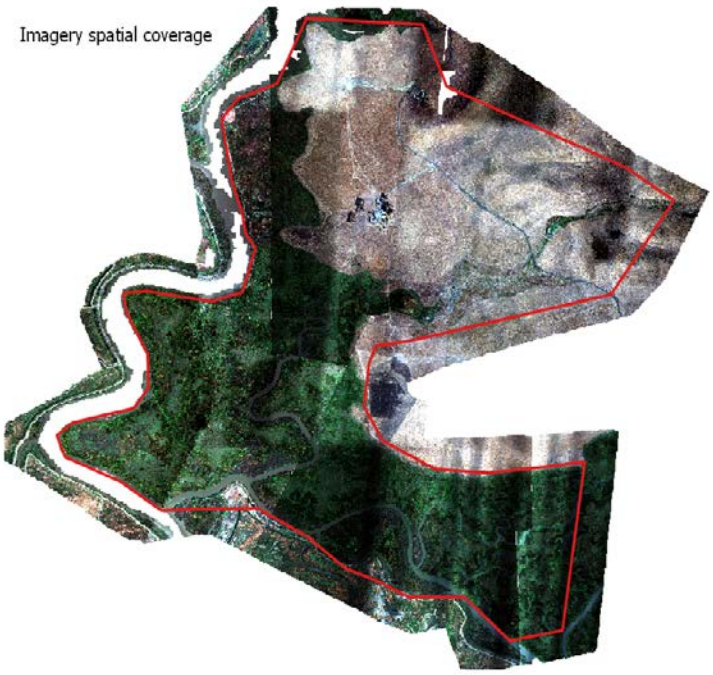


Figure 49. Imagery spatial coverage at Rush Ranch

4.3.2 Imagery

The final imagery collection for Rush Ranch was considerably different from originally planned by NOAA and agreed upon by the contractors. The areas originally submitted with the statement of work consisted of three smaller areas within the marsh totaling approximately one square mile of area with a plan for repeat coverage. At that point, we anticipated that it could be flown at an altitude of 400 feet. Consolidation of areas and the loss of an opportunity to get multiple seasons of coverage within the marsh changed this to a single area of approximately two square miles. The certificate of authorization also restricted the flight altitude to 200 feet. Combined, this changed an estimated 2-3 day collection of disconnected areas to a 6-day collection.

Radiometric consistency was directly affected by these changes. Rather than three separate areas, each collected within a limited amount of time and thus limited light variability, the collection was spaced over six days, with all types of light conditions, and without any separation of areas. It is reasonable to expect that for the original three areas of interest, the radiometric values within each area would have been consistent and any variation between areas caused by different time and lighting conditions would have been less glaring because they were independent of each other. Low collection altitude compounded this problem by dramatically increasing the collecting period, and the smaller image footprints made mosaicking and color balancing more challenging.

The volume of imagery data and the changing light conditions inherent in collecting all day for six days resulted in many processing issues. PrecisionHawk used the Amazon Cloud for the processing but still had to break the processing up into blocks. PrecisionHawk set and surveyed 44 targets in this project area, which for the area seemed at the time to be professionally adequate. They drew their image-processing boundaries in a way that maximized the number of ground control points in each block and tried to overlap control points along seams, while trying to group by illumination. However, even with their best efforts, some of the blocks only had 3-4 ground control points. These data were processed in Agisoft PhotoScan, which is a robust photogrammetric software but in this case was unable to combine the large number of images and low number of ground control points into a horizontally accurate data set. The only way around this would have been to increase the ground control from 44 points to around 110 points.

Mosaicking all the blocks together to make a color-balanced final image failed to produce a good product. The learning curve PrecisionHawk had to go through to handle over 380,000 images delayed the delivery of the imagery by many months. Final delivery of 11 blocks was received on April 27, 2018, almost eight months after the flights completed in early September. With an area approximately six times larger than the originally anticipated

individual areas and half the flying height, PrecisionHawk had to mosaic approximately 24 times more images together than they had planned, and those images had much greater illumination variability. This illustrates a potential limiting factor for UAS collection areas that cannot be collected rapidly due to their size. In contrast, a manned flight with a large format digital camera might need less than 100 frames to cover the area with standard stereo overlap and could be covered with relatively constant illumination, though the turns would greatly reduce efficiency.

The variability in illumination made the imagery impossible to color balance, resulting in the approach of processing by blocks of similar illumination. Even with this approach, there are color balance issues within blocks. Mosaicking the blocks, as with the coverage image above, clearly shows the color balance issues at the block boundaries. Although the imagery clearly has faults, the Quantum–PrecisionHawk team did everything they could under difficult circumstances with a task that was well beyond the original plan and agreement. They contributed significant amounts of their own resources, likely contributing more money to the project than NOAA did.

4.3.2.1 Positional Accuracy

The horizontal positional accuracy was measured by the Quantum–PrecisionHawk team and by the NOAA–National Estuarine Research Reserve System team independently with results of 1.626 meters and 2.09 meters (n=33) RMSExy, respectively. This is far worse error than the specification of 15 centimeters. The spatial distribution of the error is non-random (see Figure 49), with the highest errors in the hills. It is not clear from the contractor's technical data report that they accounted for elevation during the orthorectification process. We asked for clarification, and the contractor did verify that elevation was used in the orthorectification process. However, it was elevation derived from the imagery (see "Structure from Motion" section), and this may have been the cause of the problem.

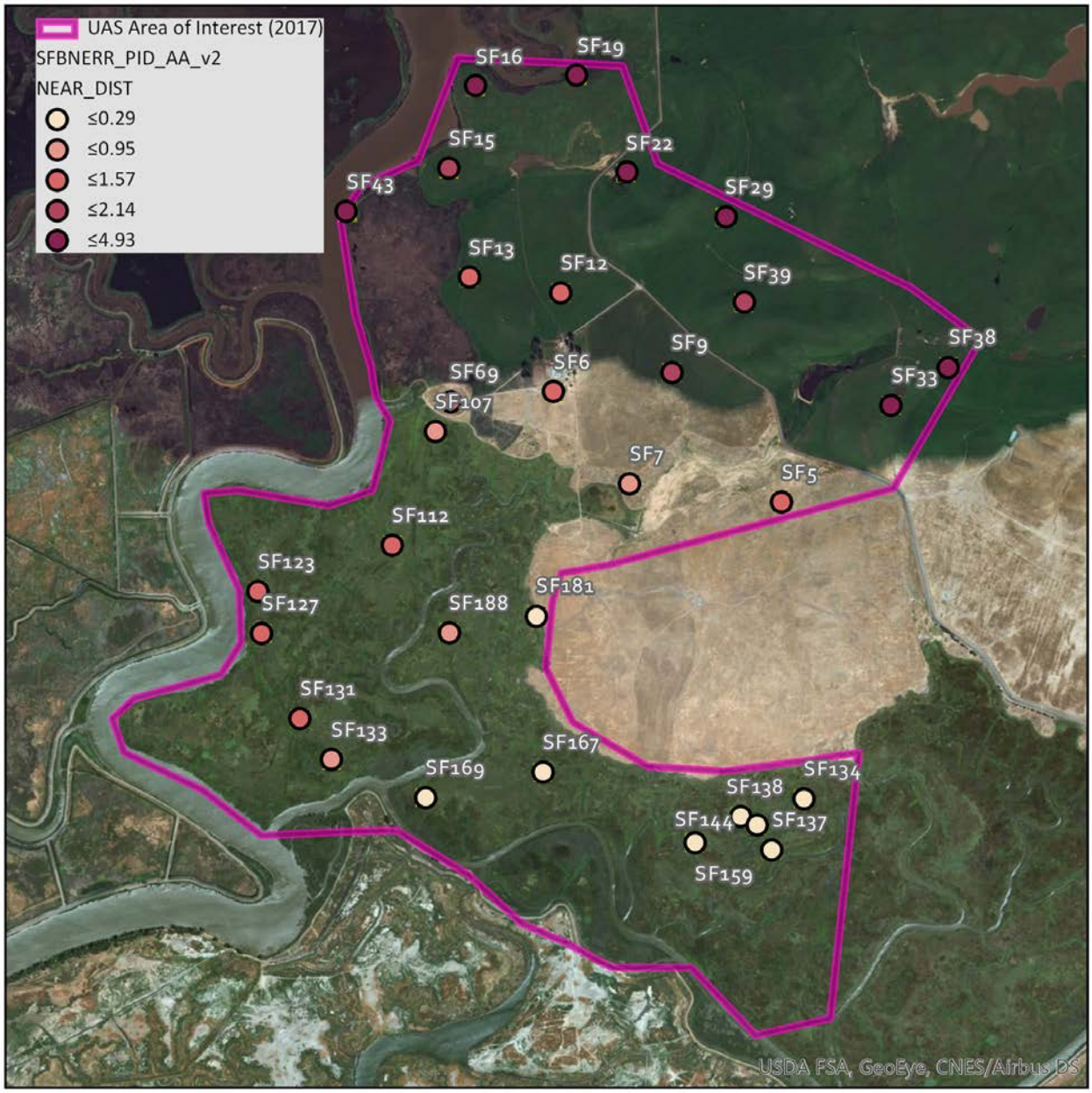


Figure 50. Imagery spatial distribution of positional errors for Rush Ranch

The analysis of the imagery resolving power indicated that the final mosaicked imagery is able to resolve features that are approximately the width of a pixel. The image in Figure 51 clearly shows the 4-centimeter wide fabric strips. Surprisingly, the largest bars on the USAF-1951 chart can also be resolved, although they should have been slightly smaller than a pixel.

Figure 51. USAF-1951 chart as seen in the mosaicked imagery. No stretch was applied.

4.3.2.2 Imagery Analysis

Plant spectral characteristics have been researched extensively since the early 1960s (Gates, Keegan, Schleter, & Weidner, 1965). Known relationships between spectra and spectral indices (e.g. leaf area index and normalized difference vegetation index) have found increasingly wide use in vegetation and plant population studies. Despite the tremendous body of research, applications for remote sensing in vegetation and ecological studies are constrained by accuracy, accessibility, and scalability challenges. UAS platforms, sensor technologies, and software innovations have significant potential to address these challenges by improving data quality, providing repeatable analytical methods, and streamlining data management and processing routines. More importantly, these emerging technologies have proven effective at reducing the level of effort and ecological impacts of data collection in sensitive habitats, while maximizing the information content of imagery data outputs. Our analysis of the Rush Ranch imagery aims to evaluate information gain as the spatial resolution of imagery increases. We compare UAS, aerial, and satellite imagery products along a gradient of spatial resolutions ranging from approximately 0.04-meter to 2-meter pixel resolution. We summarize the classification methodology and image-processing approach and present a quantitative and qualitative analysis of our results. Image classification data products and accuracy metrics are included at the end of the report.

4.3.2.2.1 Approach

We focused primarily on evaluating spatial and spectral characteristics that differentiate UAS, aerial, and satellite imagery. To quantify results, we compared the accuracy of classified imagery data across five unique imagery products (Table 10) within a supervised classification framework. To streamline data processing and analysis we constrained the study area to an approximately 4-acre area of interest. The area of interest provides a representative sample of the site's vegetation assemblage and reduces the possibility of introducing spectral and spatial variability from image processing artifacts, stitching, edge effects, environmental factors, and radiometric calibration. Data inputs were limited to three spectral bands (R, G, B) for consistency and to facilitate analysis across different imagery types. Classification was accomplished by applying multiple synthetic classification routines to a set of UAS, aerial, and satellite imagery data (Table 10). Intermediate classification products were evaluated for suitability, and a final classification approach was selected from a qualitative analysis of the intermediate data.

Table 10. Imagery resolution, acquisition dates, and product type (Rush Ranch)

Imagery Type	Resolution	Bands	Acquisition Date	Product
UAS	.04 meters	R, G, B	September 4-10, 2017	Precision Hawk, 5-band multispectral
Aerial	.12 meters	R, G, B	August 25, 2017	TerrAvion, 5-band multispectral
Aerial	.12 meters	R, G, B	September 21, 2017	TerrAvion, 5-band multispectral
Satellite	.4 meters	R, G, B	August 19, 2017	World View 2, Pan-sharpened NC
Satellite	2 meters	R, G, B	August 6, 2014	World View 2, 8-band multispectral

4.3.2.2.2 Imagery Analysis and Processing

Imagery analysis was conducted in ArcGIS Pro utilizing classification, segmentation, and image processing tools available commercially through ESRI's ArcGIS platform. Classification and segmentation routines were scripted and executed in a Python 3.6.4 programming environment to ensure repeatable results and analytical consistency. Scripts are archived in the project database along with imagery, intermediate data, and classification results. The final classification approach was based on results of the initial exploratory data analysis, which included the following:

- Principal Components Analysis
- Maximum Likelihood Algorithm
- Nearest Neighbor Classification
- Random Forrest Algorithms

4.3.2.2.3 Image Segmentation

Image segments were generated for each imagery product outlined in Table 10. Image segmentation recursively groups pixels until homogeneity criteria are met, producing relatively homogenous polygons with associated quantitative characteristics (e.g., length/width, spectral range and variance). These multi-dimensional "image objects" characterize the plant population based on spatial and spectral patterns. Three main parameters were adjusted to control image segmentation:

1. Spectral detail: relative measure of importance (weight) of the spectral variability of image segments that are generated via the segmentation algorithm
2. Spatial detail: relative measure of importance (weight) of the size of image segments that are generated via the segmentation algorithm
3. Minimum segment size: defines the minimum mapping unit

Image segmentation results (see Figure 54, Figure 57, and Figure 60) varied widely depending on input and scale parameters (see 1-3 above). Spatial and spectral detail influenced segmentation results to a lesser degree than segment size alone. However, we did not quantify the significance of individual input parameters, as this was beyond the scope of our investigation. Input parameters were iteratively adjusted to optimize segmentation and reduce the spatial variability of image segments across imagery products. Segmentation tended to obscure vegetation boundaries in the high-resolution UAS imagery segmentation. Segmented boundaries were consistently disassociated with vegetation boundaries that were clearly delineated at the pixel level. However, this effect was less pronounced in the satellite and aerial image segments due to larger pixel size.

4.3.2.2.4 Final Classification Approach

While manual optimization (parameterization) of image segments generally improved results, it had the undesirable effect of introducing bias from the analyst. Ultimately, manual optimization precluded the standardization of the classification approach across imagery products. More importantly, segmentation consistently failed to accurately delineate vegetation communities at smaller spatial scales (see image segmentation section). To minimize information loss and biases resulting from image segmentation, we opted for a pixel-based classification approach. This approach classifies each pixel independently, whereas segmentation aggregates pixels to produce larger discreet image objects from recursively grouped pixels. Pixel-based classification provides a more detailed classification at a finer spatial resolution but suffers from "noise" in the scene. This approach may be best suited for vegetation communities with relatively high diversity and spatial heterogeneity, such as Rush Ranch.

The final image classifications were generated with a supervised classification approach using identical algorithms and input parameters. Training polygons were digitized from vegetation communities readily identifiable in the UAS imagery and were then used to generate spectral training data for the UAS, aerial, and satellite imagery. A maximum likelihood classification algorithm was applied to assign individual pixels to one of five classes (Table 11) representing a dominant vegetation type. Image processing was scripted to ensure that the classification methodology was consistent for all classified outputs.

Table 11: Vegetation classification class names and classification codes

Class Name	Class Code
Salicornia pacifica	C_1
Typha spp.	C_2
Schoenoplectus americanus	C_3
Schoenoplectus acutus	C_4
Juncus balticus	C_5

4.3.2.2.5 Accuracy Assessment

An accuracy assessment was conducted for each of the classification results and a standard-error matrix was generated along with a statistical estimation of classification accuracy. Accuracy results were compared to quantify differences in classification accuracy between UAS, aerial, and satellite imagery and across vegetation types.

4.3.2.2.6 Imagery Classification Results

The accuracy assessment was conducted in ArcGIS Pro. Error matrices, producers and users accuracy, and kappa statistics were calculated for all classified data sets. The classification accuracy improved with increasing spatial resolution in all cases (Table 12). Previous research in this ecological community are consistent with these results. Interannual variation and temporal differences may account for some variation in classification accuracy, particularly with respect to the August 2014 World View-2 product. However, the magnitude of the accuracy difference was much higher than would be expected due to variation in the spatial extent (or shifts) of vegetation communities across acquisition dates (e.g. UAS imagery vs. WV-2 2014 imagery). Further investigation is recommended to qualify these results.

Table 12: Comparison of accuracy assessment results

Imagery Type	Resolution	Overall Accuracy	Kappa	Product
UAS	0.04 meters	.936	.92	Precision Hawk, 5-band multispectral
Aerial	0.12 meters	.724	.655	TerrAvion, 5-band multispectral
Satellite	0.4 meters	.704	.63	World View 2, Pan-sharpened NC
Satellite	2 meters	.64	.55	World View 2, 8-band multispectral

Our research demonstrates that UAS imagery provides higher classification accuracy and is more representative of the vegetation assemblage when compared to satellite and aerial imagery and their derivative classification products. Vegetation classes are clearly delineated in the UAS classification. Conversely, satellite classification often confounds classes, and lacks the well-defined boundaries that are apparent in the UAS classification. Aerial imagery classifications sufficiently represent vegetation community boundaries but with less precision than the UAS imagery. With the exception of the 2-meter World View-2 classification, all classified data outputs exceeded the overall accuracy of the existing 2015 habitat map (overall accuracy = 0.697) generated from the 2-meter World View-2 data. Our results demonstrate the value of high-resolution imagery data products for improved classification accuracy, particularly in estuarine systems with high ecological diversity and spatial heterogeneity. Likewise, this study highlights opportunities for future habitat mapping efforts within the research reserve system that leverage UAS technologies.

4.3.2.2.7 Imagery and Classification Outputs at Rush Ranch

Graphic and tables from the imagery and multiple landcover classification (MLC) analysis are shown on the next several pages.

Figure 52: WV-2 (2-meter) natural color imagery

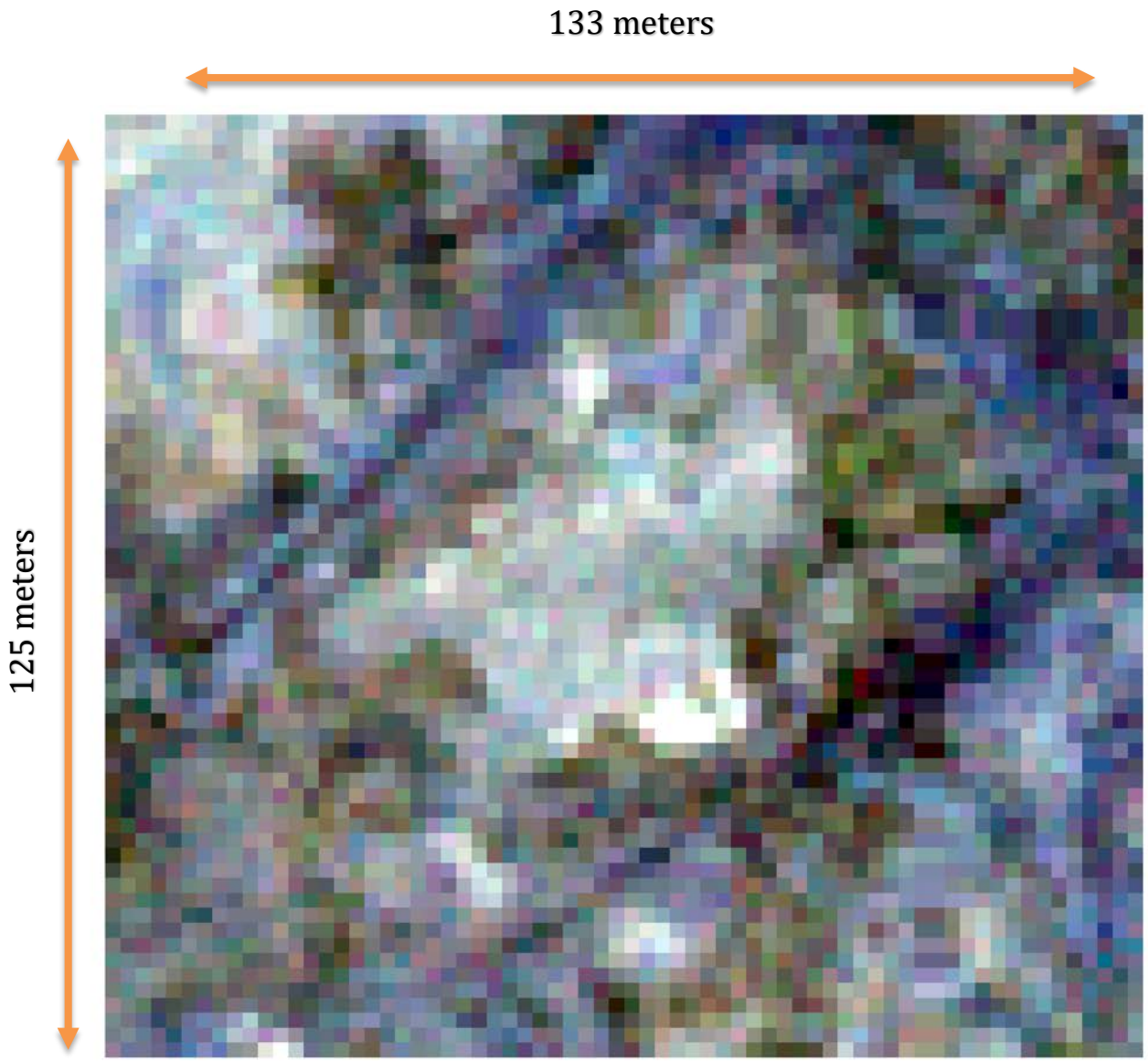
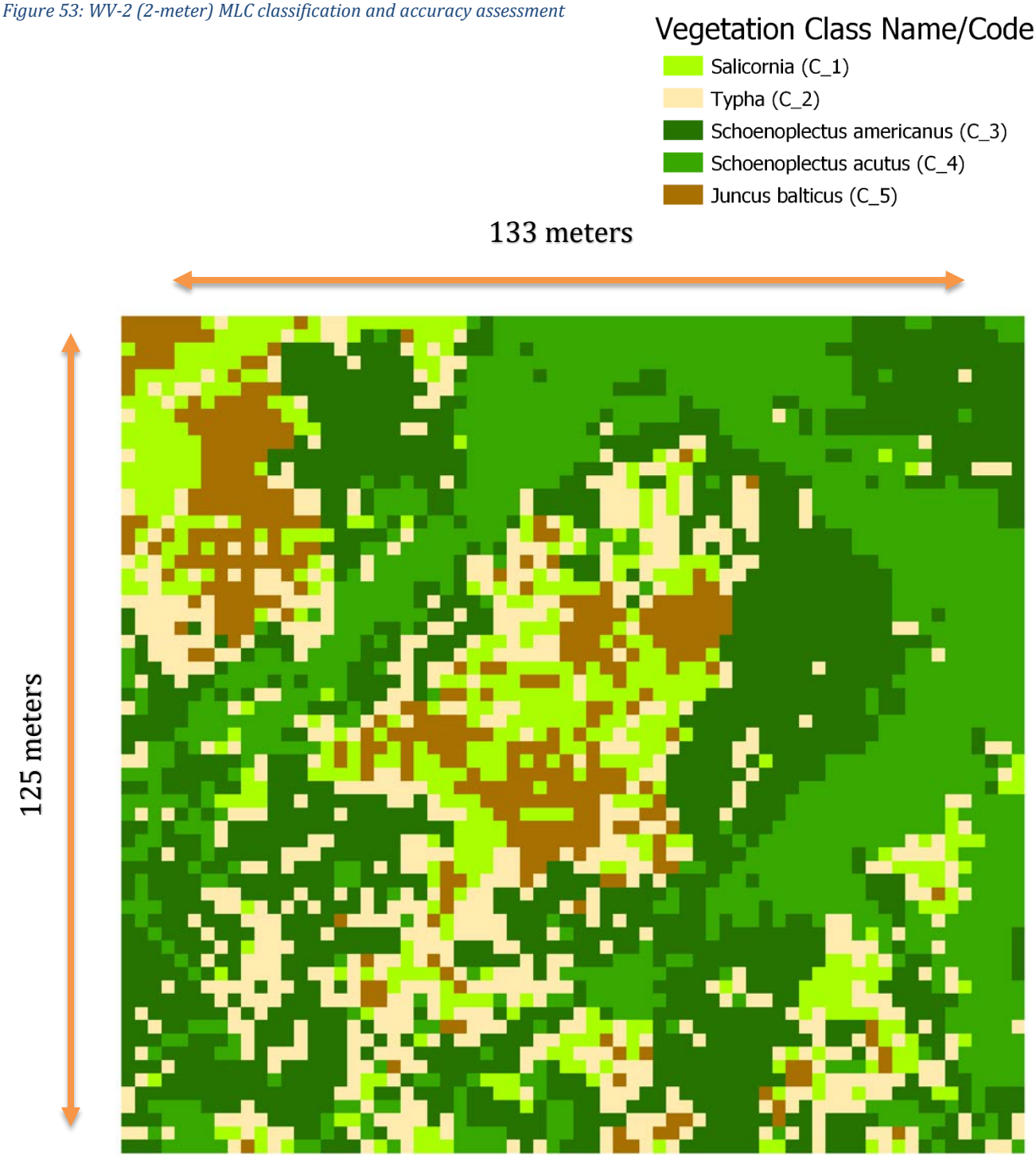


Figure 53: WV-2 (2-meter) MLC classification and accuracy assessment



ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	22	11	4	2	13	52	0.423076923	0
C_2	5	22	5	1	6	39	0.564102564	0
C_3	0	2	39	1	0	42	0.928571429	0
C_4	3	6	1	46	0	56	0.821428571	0
C_5	20	9	1	0	31	61	0.508196721	0
Total	50	50	50	50	50	250	0	0
P_Accuracy	0.44	0.44	0.78	0.92	0.62	0	0.64	0
Kappa	0	0	0	0	0	0	0	0.55

Figure 54: WV-2 (2-meter) close-up of image segmentation

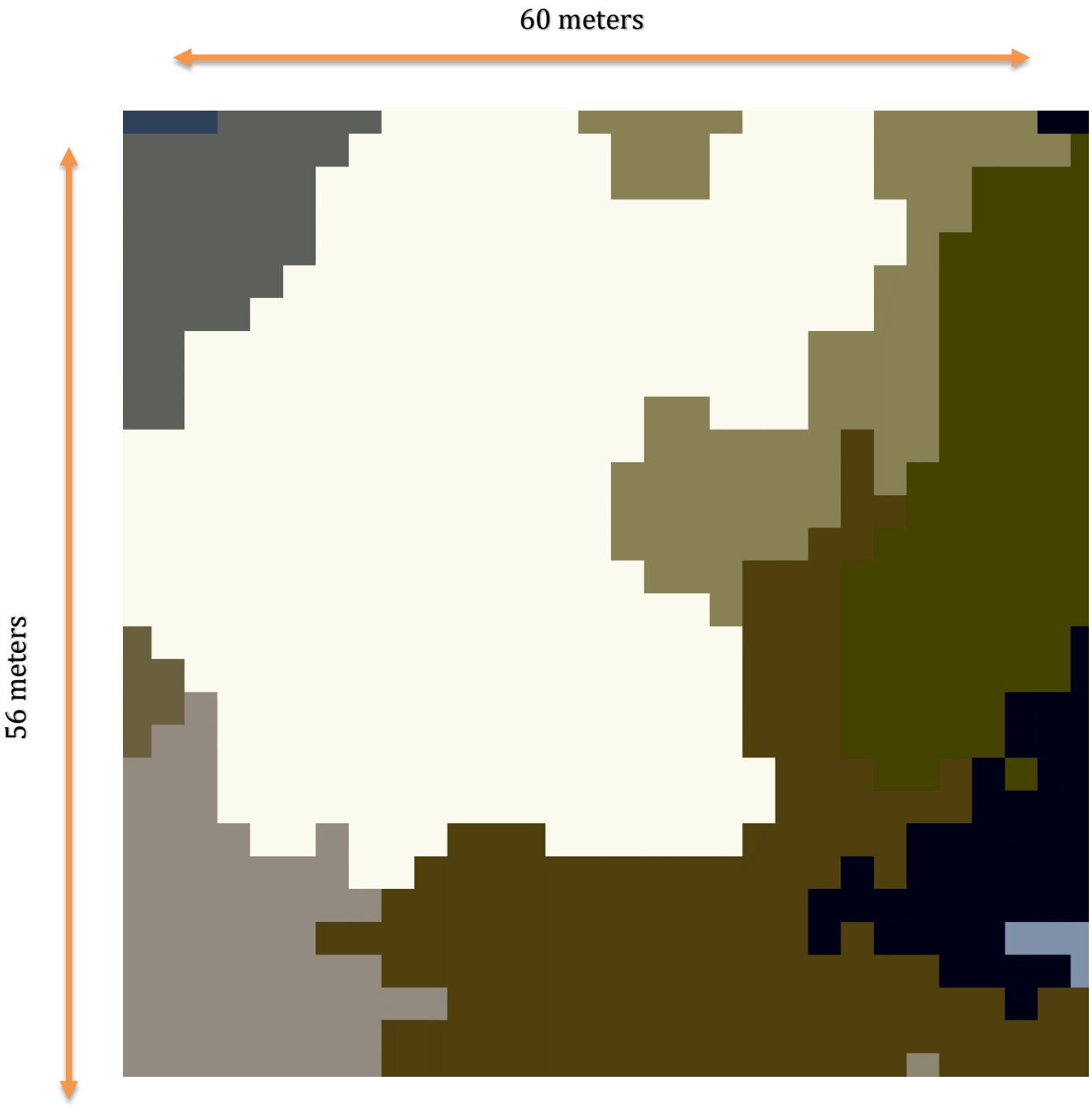


Figure 55: WV-2 (0.4 meter) natural color imagery

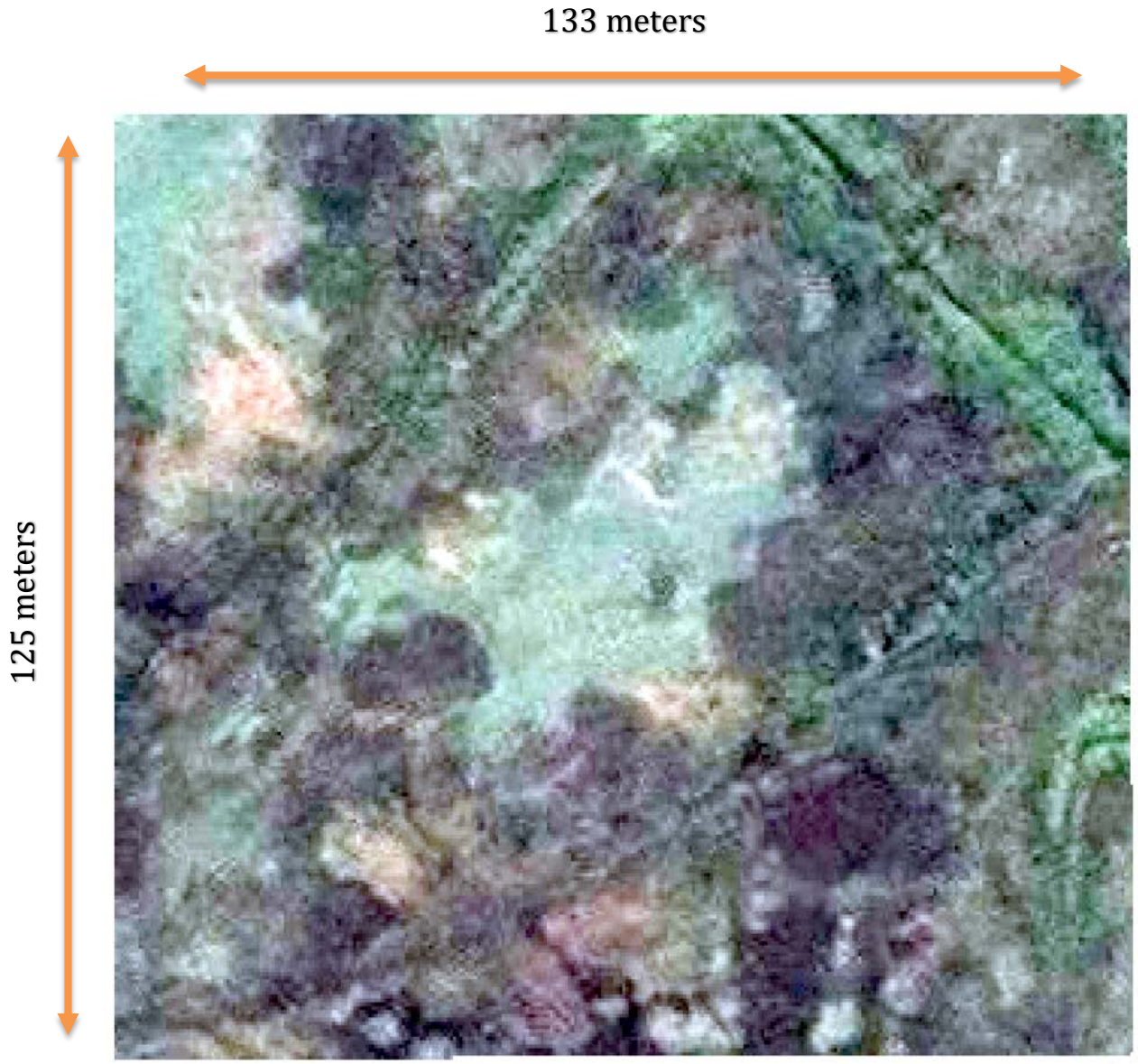
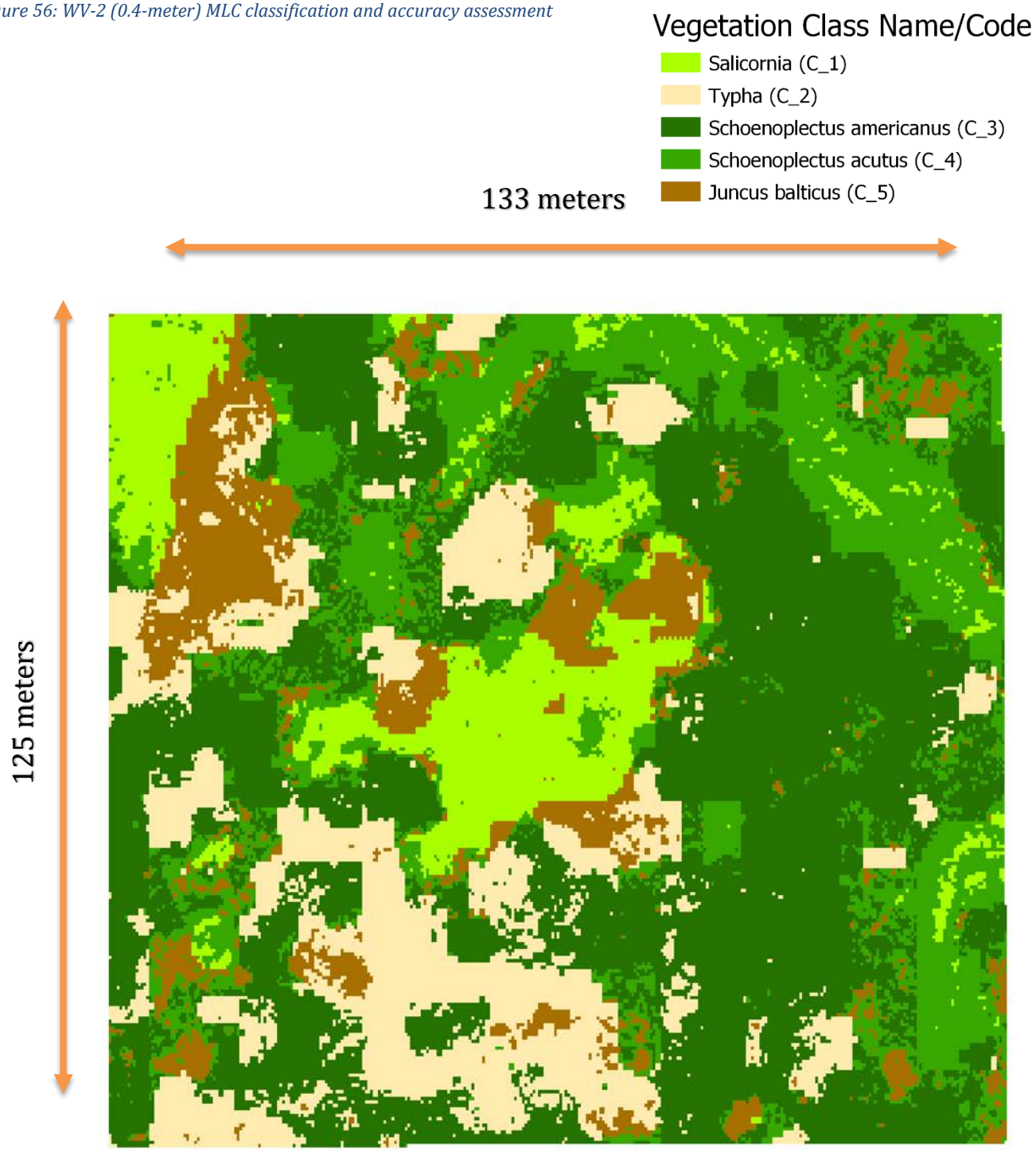


Figure 56: WV-2 (0.4-meter) MLC classification and accuracy assessment



ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	44	0	0	4	14	62	0.709677419	0
C_2	1	33	5	1	3	43	0.76744186	0
C_3	0	4	45	23	0	72	0.625	0
C_4	1	3	0	22	1	27	0.814814815	0
C_5	4	10	0	0	32	46	0.695652174	0
Total	50	50	50	50	50	250	0	0
P_Accuracy	0.88	0.66	0.9	0.44	0.64	0	0.704	0
Kappa	0	0	0	0	0	0	0	0.63

Figure 57: WV-2 (0.4-meter) close-up of image segmentation

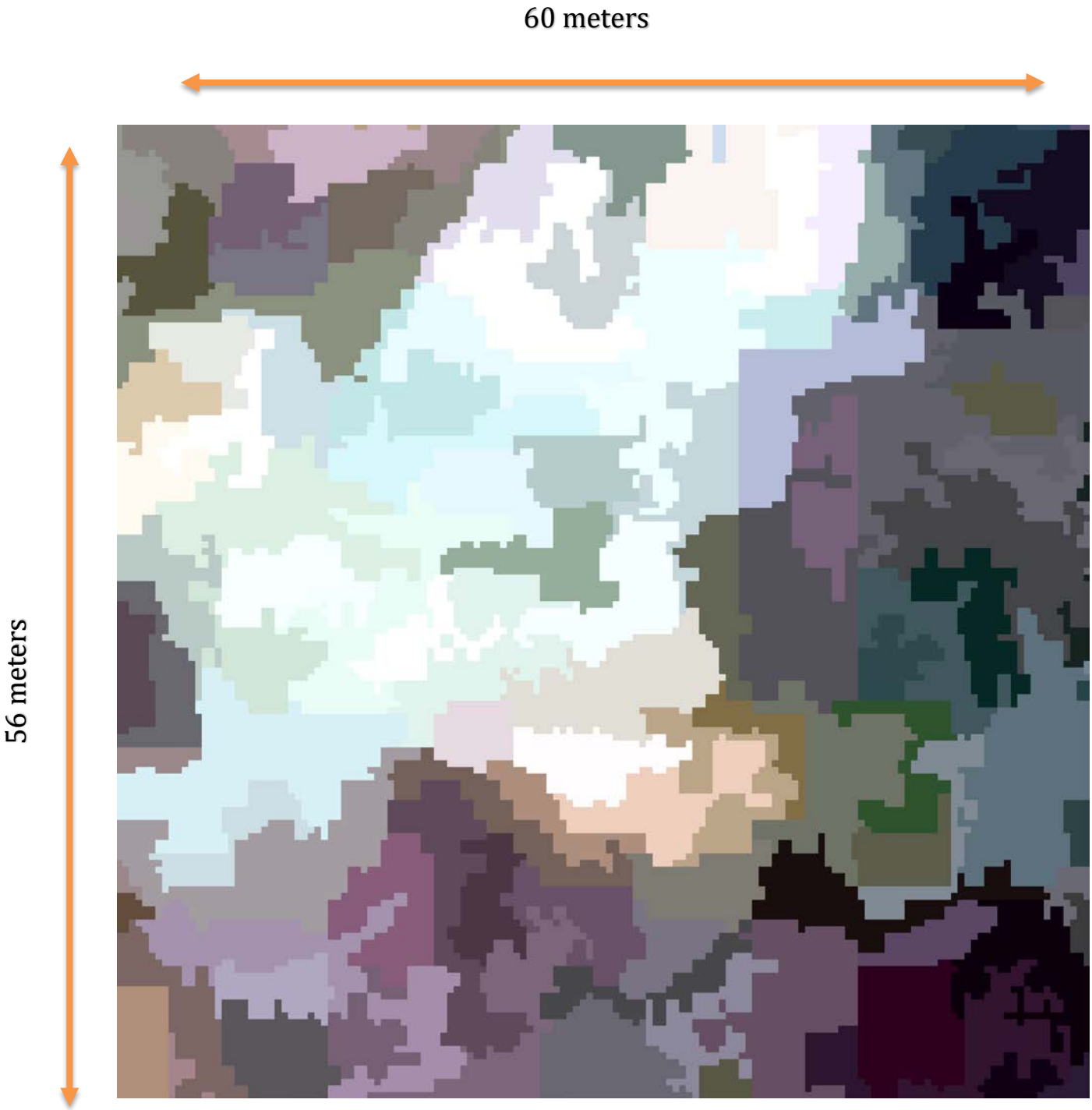


Figure 58: TerrAvion (0.12-meter) natural color aerial imagery

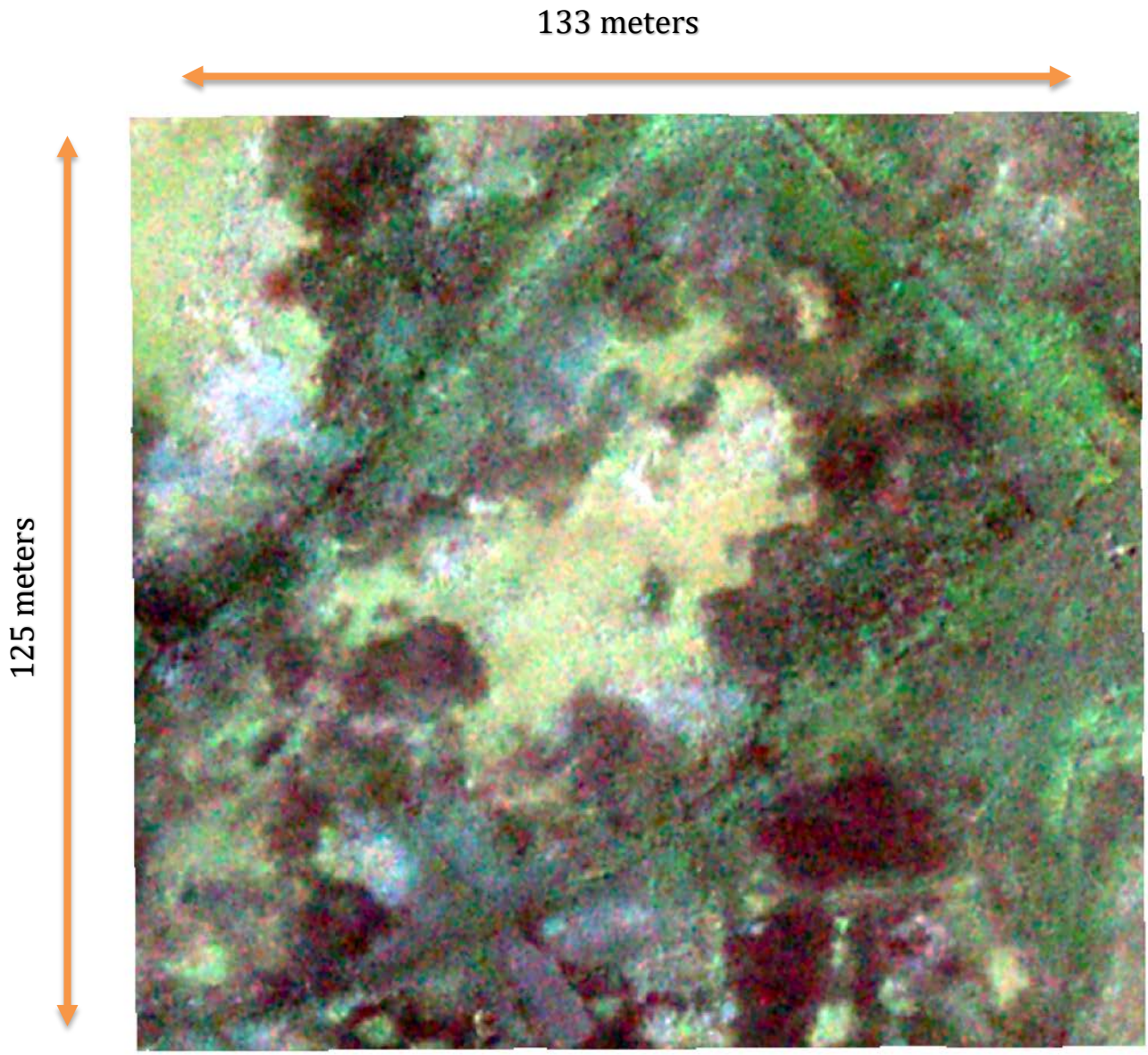
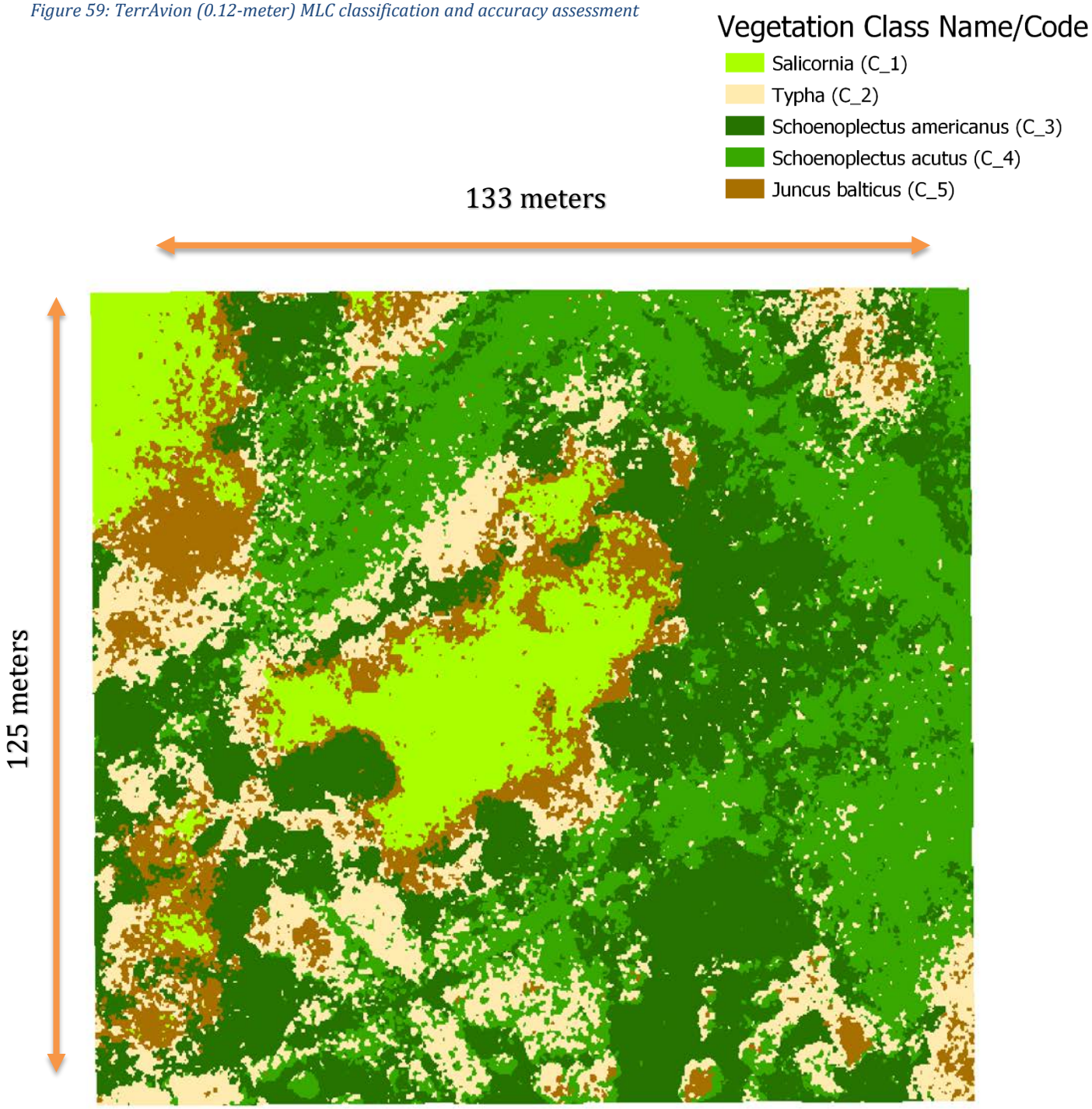


Figure 59: TerrAvion (0.12-meter) MLC classification and accuracy assessment



ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	44	3	0	0	25	72	0.611111111	0
C_2	0	33	4	3	1	41	0.804878049	0
C_3	0	0	41	6	2	49	0.836734694	0
C_4	0	8	5	41	0	54	0.759259259	0
C_5	6	6	0	0	22	34	0.647058824	0
Total	50	50	50	50	50	250	0	0
P_Accuracy	0.88	0.66	0.82	0.82	0.44	0	0.724	0
Kappa	0	0	0	0	0	0	0	0.655

Figure 60: TerrAvion (0.12-meter) close-up of image segmentation

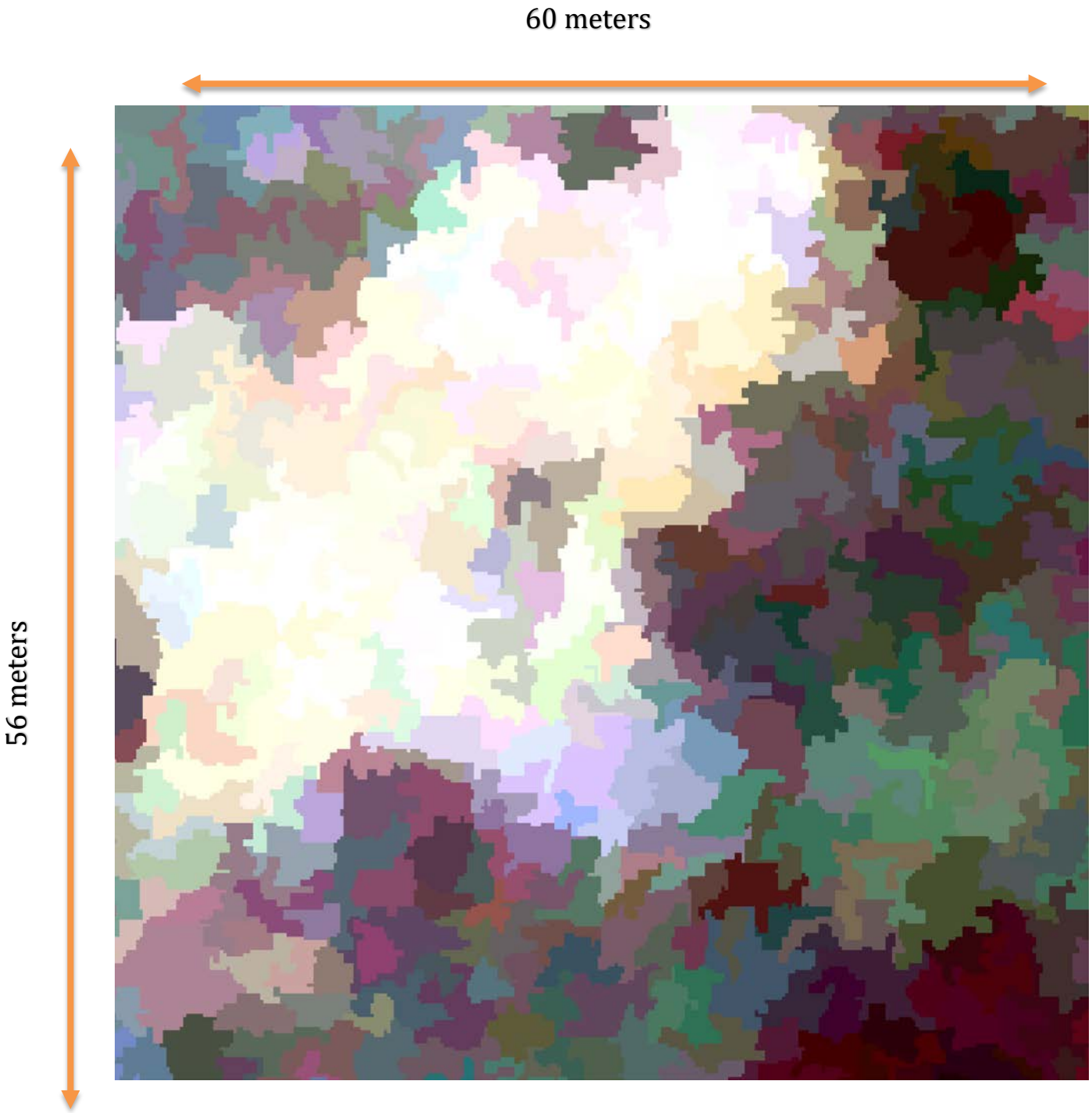


Figure 61: PrecisionHawk (.04-meter) UAS natural color imagery

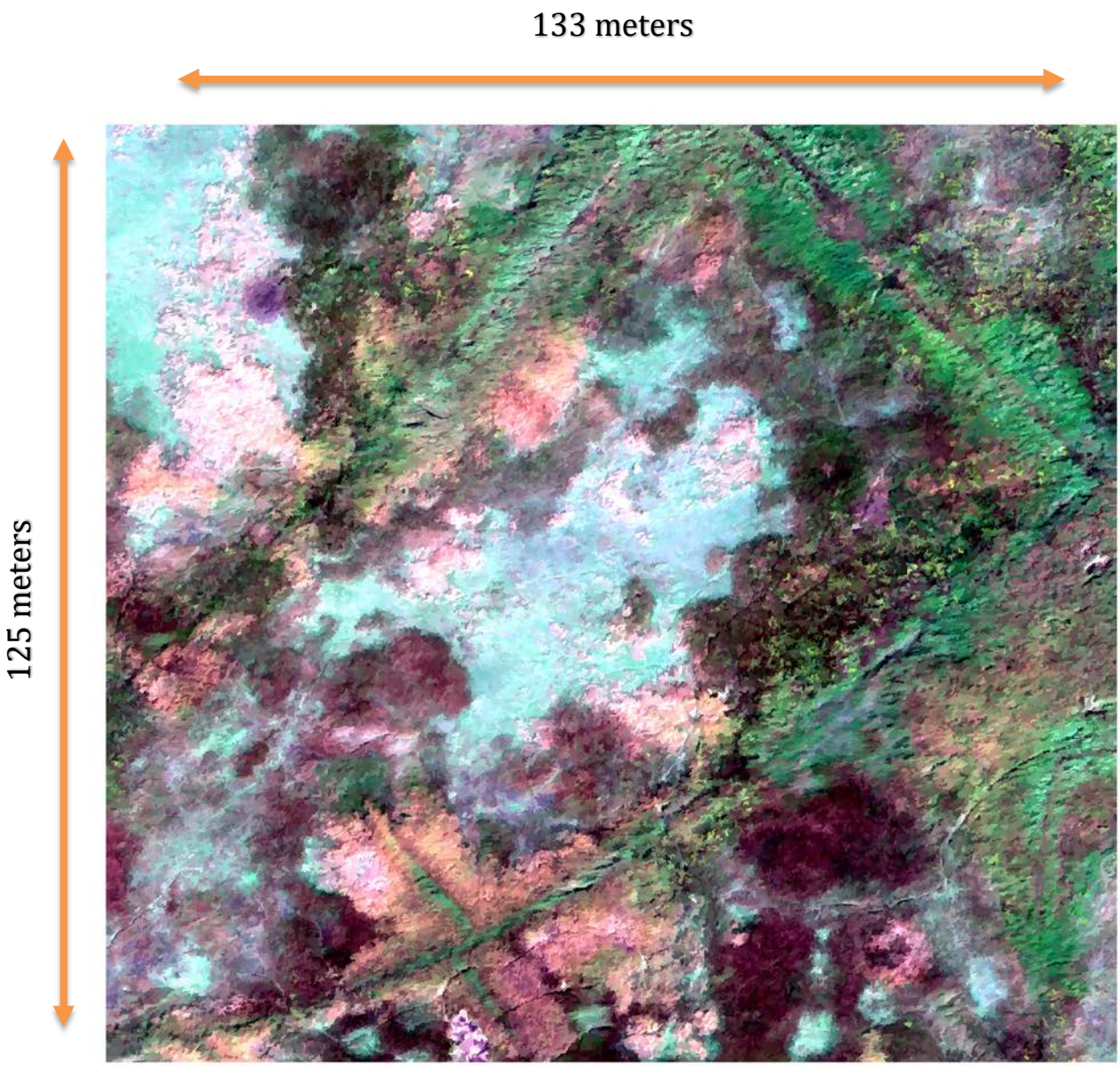
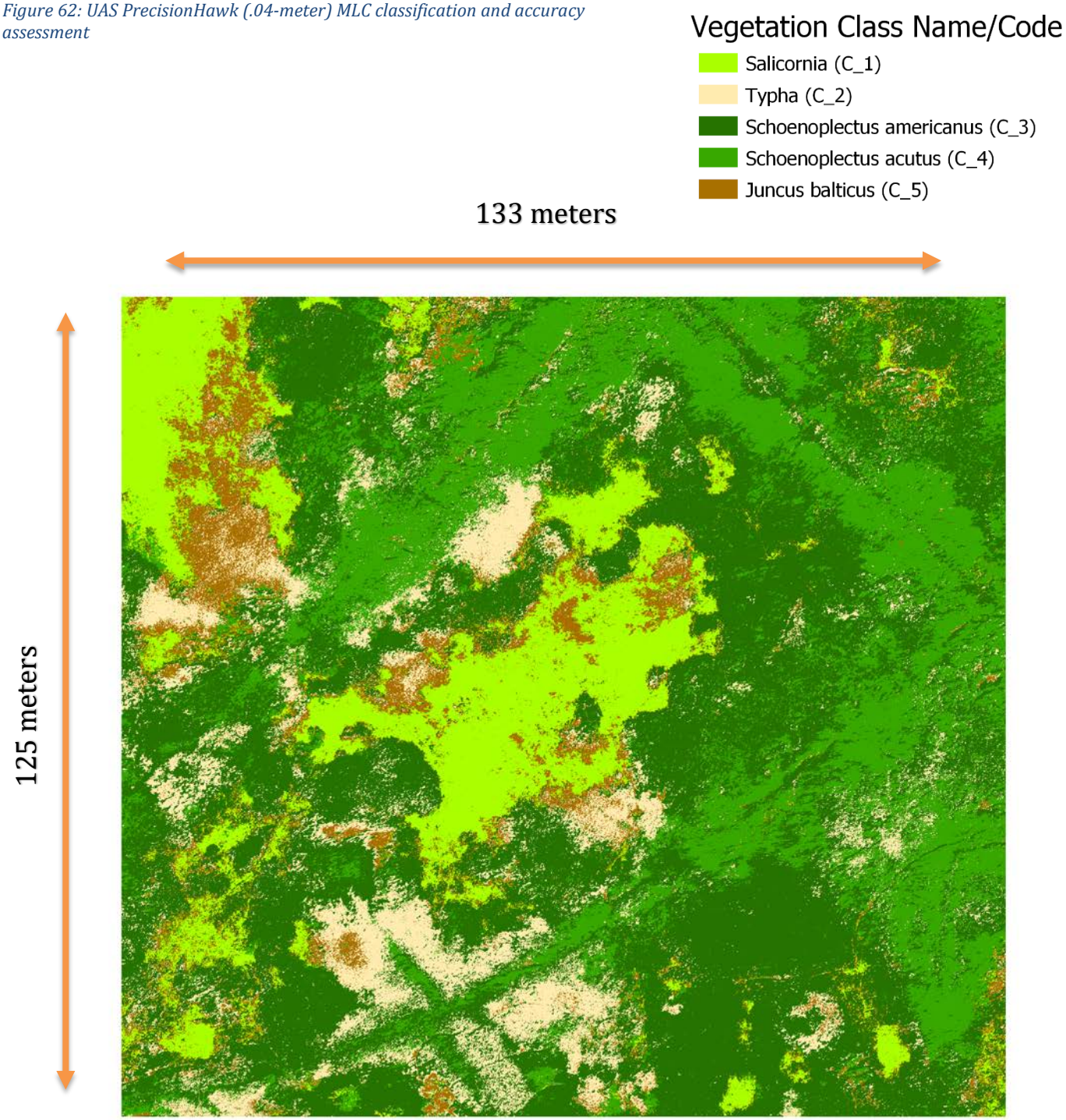
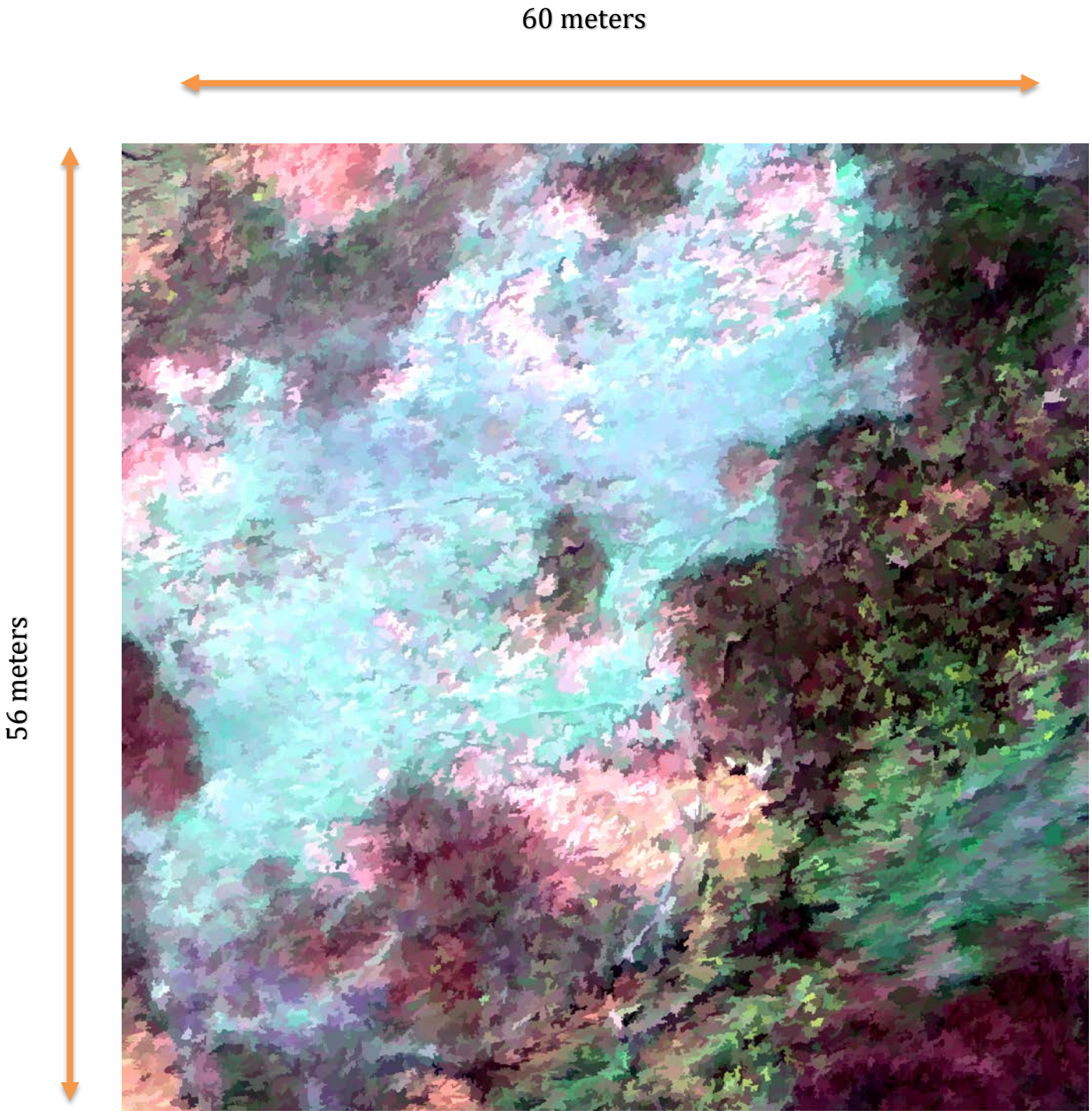


Figure 62: UAS PrecisionHawk (.04-meter) MLC classification and accuracy assessment



ClassValue	C_1	C_2	C_3	C_4	C_5	Total	U_Accuracy	Kappa
C_1	49	0	0	1	1	51	0.960784314	0
C_2	0	45	1	0	4	50	0.9	0
C_3	0	1	47	0	0	48	0.979166667	0
C_4	0	2	2	48	0	52	0.923076923	0
C_5	1	2	0	1	45	49	0.918367347	0
Total	50	50	50	50	50	250	0	0
P_Accuracy	0.98	0.9	0.94	0.96	0.9	0	0.936	0
Kappa	0	0	0	0	0	0	0	0.92

Figure 63: PrecisionHawk (.04-meter) close-up of image segmentation



4.3.3 Lidar

The overall pattern of the bare-earth lidar agrees with the expectations for the landscape, varying from near zero NAVD88 on the rivers to almost 70 meters in the hills (Figure 64).

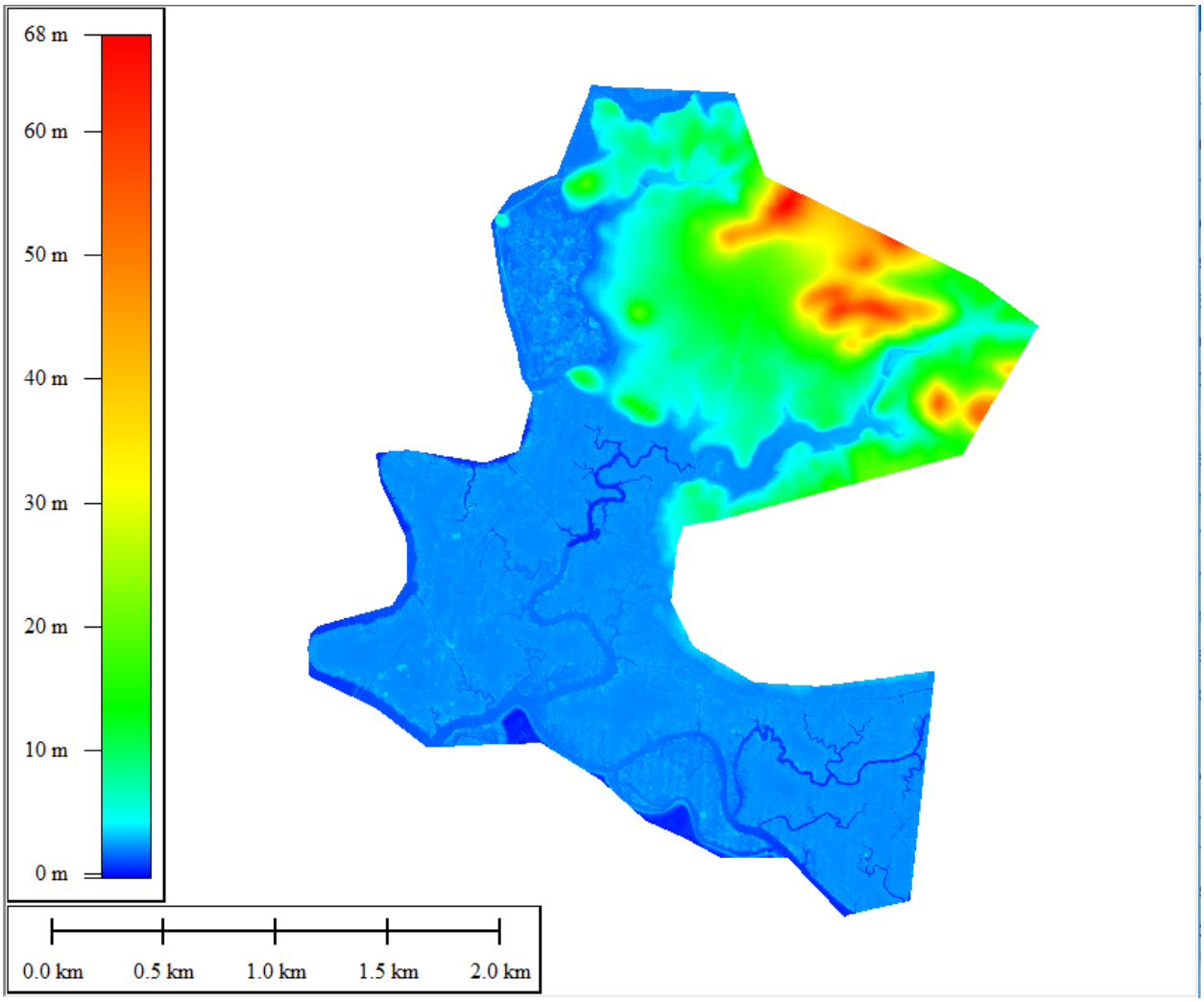


Figure 64. Bare-earth elevations from lidar at Rush Ranch

Subtracting the highest points per 1-meter cell from the bare-earth, we can examine what the lidar shows for canopy and structures (Figure 65). The hills have low vegetation of dry grass. The marsh shows that the higher grasses are near the water bodies. The highest vegetation is the trees surrounding the buildings.

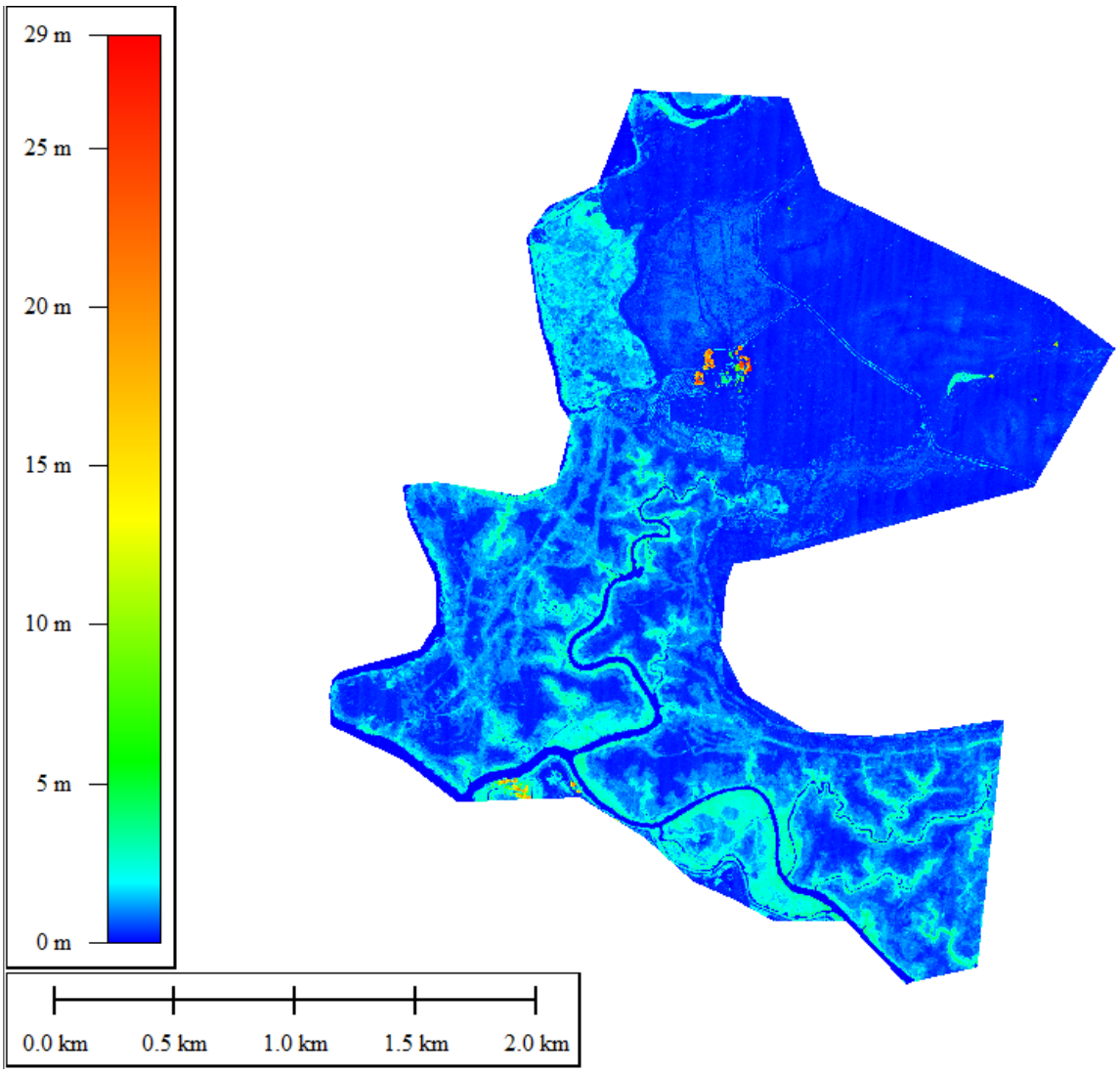


Figure 65. Canopy and structure heights derived from lidar at Rush Ranch

We used 62 bare-earth validation points to calculate the non-vegetated RMSEz accuracy as 0.047 meters. This easily meets the contract specifications of 0.10 meters. The RMSEz for the vegetated classes is shown in the graph (Figure 66) and table (Table 13) below. The graph only shows the classes with more than a few points.

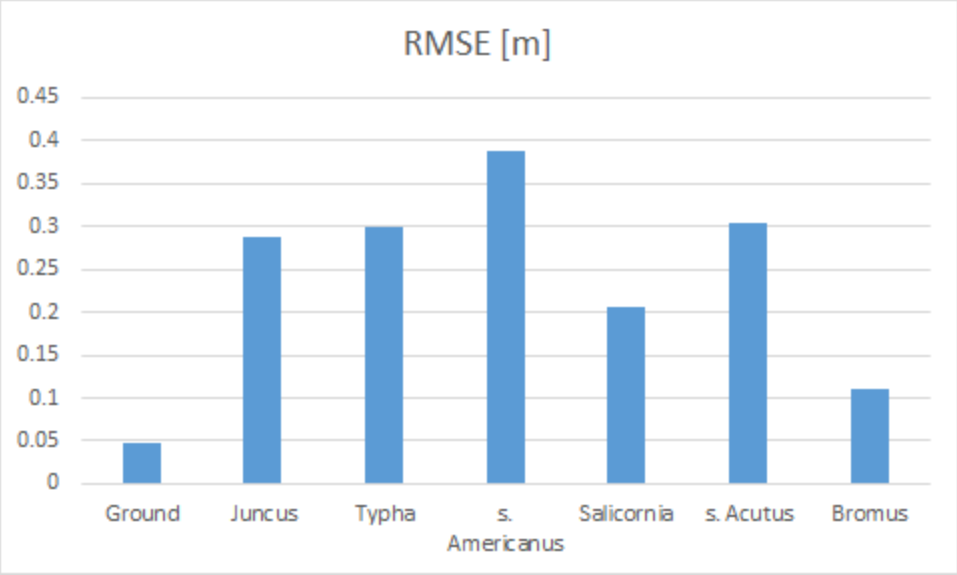


Figure 66. Lidar RMSEz for lidar by vegetation type at Rush Ranch

Table 13. Tabulation of lidar errors for Rush Ranch

Class	RMSE [m]	Mean lidar - truth [m]	N
Ground	0.047	-0.003	62
Juncus	0.287	0.264	25
Typha	0.300	0.279	17
S. americanus	0.387	0.369	16
Salicornia	0.207	0.192	15
S. acutus	0.304	0.276	12
Bromus	0.109	0.066	8
Grindelia	0.276	0.275	2
Birdsbeak	0.023	0.023	2
Thistle	0.154	0.119	2
Phragmites	0.464	0.464	1
Lepidium	0.324	0.324	1

We expected that the marsh vegetation in Rush Ranch would be extremely difficult to penetrate with lidar, and this appears to be the case. With few exceptions, the RMSE is very close to the mean error, with the lidar above the ground truth. The lidar is likely returning from the top of the vegetation or from some point within the vegetation where the system cannot differentiate between the vegetation and the ground return. In the case of short dense plants such as *Salicornia*, the mean error is approximately the same as the height of the plants. For the taller (chest to head high) plants, such as *S. acutus*, *S. americanus*, *Typha*, and *Phragmites*, the mean error is deep within the

vegetation structure and may be an issue with the pulse width and signal response such that a signal approximately 30 centimeters from the ground cannot be differentiated from a signal from the ground. There were not sufficient holes in the vegetation for the pulses to get to the ground without a reflection within the plant structure. From visual observation in the field, this was not surprising, as these vegetation types create dense thickets and overlay dark, damp soils.

4.3.4 Structure from Motion

The point cloud derived from the imagery via the SfM process had significant issues. While the lidar point clouds had a range of values between 0 and 70 meters NAVD88, the SfM point cloud had a range from -183 to 271 meters (Figure 67). It seems likely that using the SfM-derived digital surface model (DSM) contributed to the horizontal error found in the imagery data and might explain the majority of the error. The SfM point cloud was generated for each of the imagery processing blocks. The point clouds were merged to generate the raster DSM below, partially obscuring the differences at block edges, which could be significant. A particularly egregious block is block 11 in the northwest. It has the full range seen in the SfM point cloud and considerably greater range than the rest of the blocks.

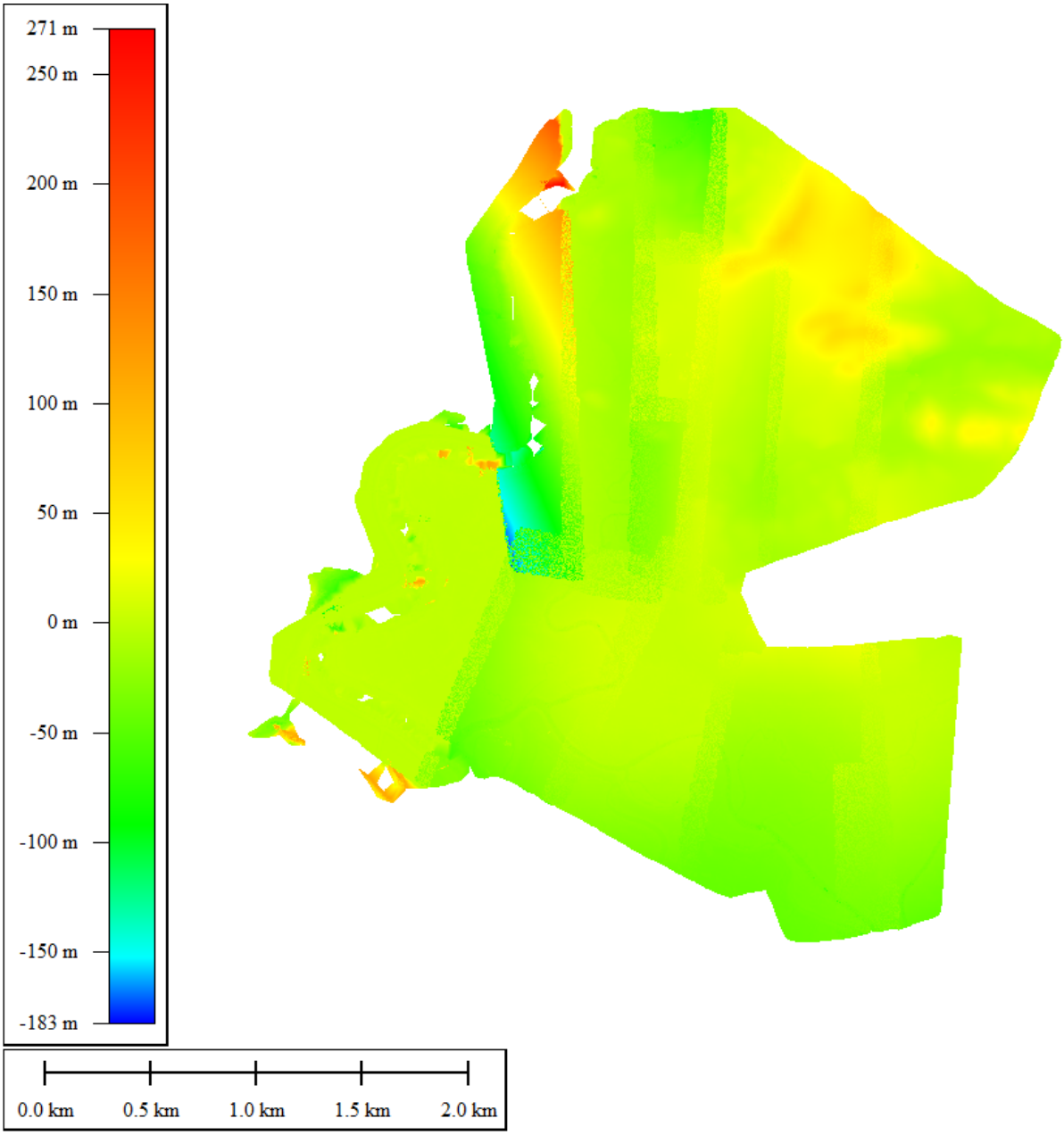


Figure 67. Digital surface model (DSM) derived from the imagery SfM point cloud

To illustrate some of the issues, we compared the SfM-derived DSM to the lidar-derived DSM (Figure 68). The default color ramp covered such a large range that it was very difficult to see where good agreement might be. A color ramp was applied such that differences greater than five meters were solid blue or red while smaller differences were in yellow and green. This is still a far greater range than we would expect between the data sources, but it is evident that most of the area exceeds a five-meter difference. A profile was drawn through the data (yellow line in the figure) showing some of the extreme changes between image blocks.

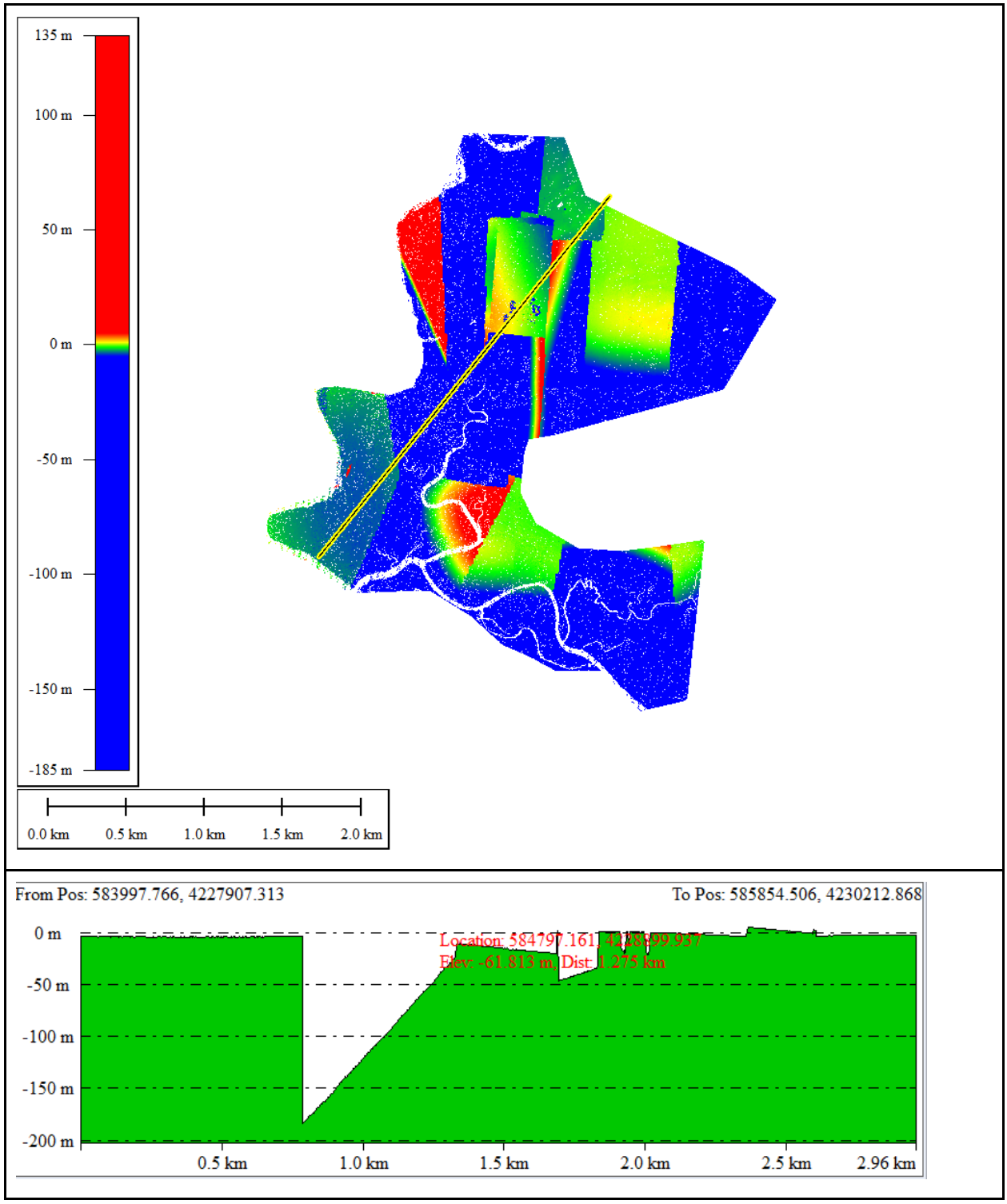


Figure 68. A profile is drawn on the difference between the SfM DSM and the lidar DSM in the upper panel and shown in the lower panel. The difference is colored such that red and blue are more than 5 meters from zero difference while greens and yellows are less than 5 meters apart.

We also drew a profile through the main ranch house area that included buildings and trees to compare the lidar and the imagery (see Figure 69). Here we can see the building peak height was in good agreement between the two at 18.18 meters in the lidar and 17.84 meters in the imagery-derived points. However, the trees do not agree. While we did not measure the tree heights during our fieldwork, the trees were significantly taller than the buildings. This is in agreement with the lidar, but the imagery-derived data show much lower trees.

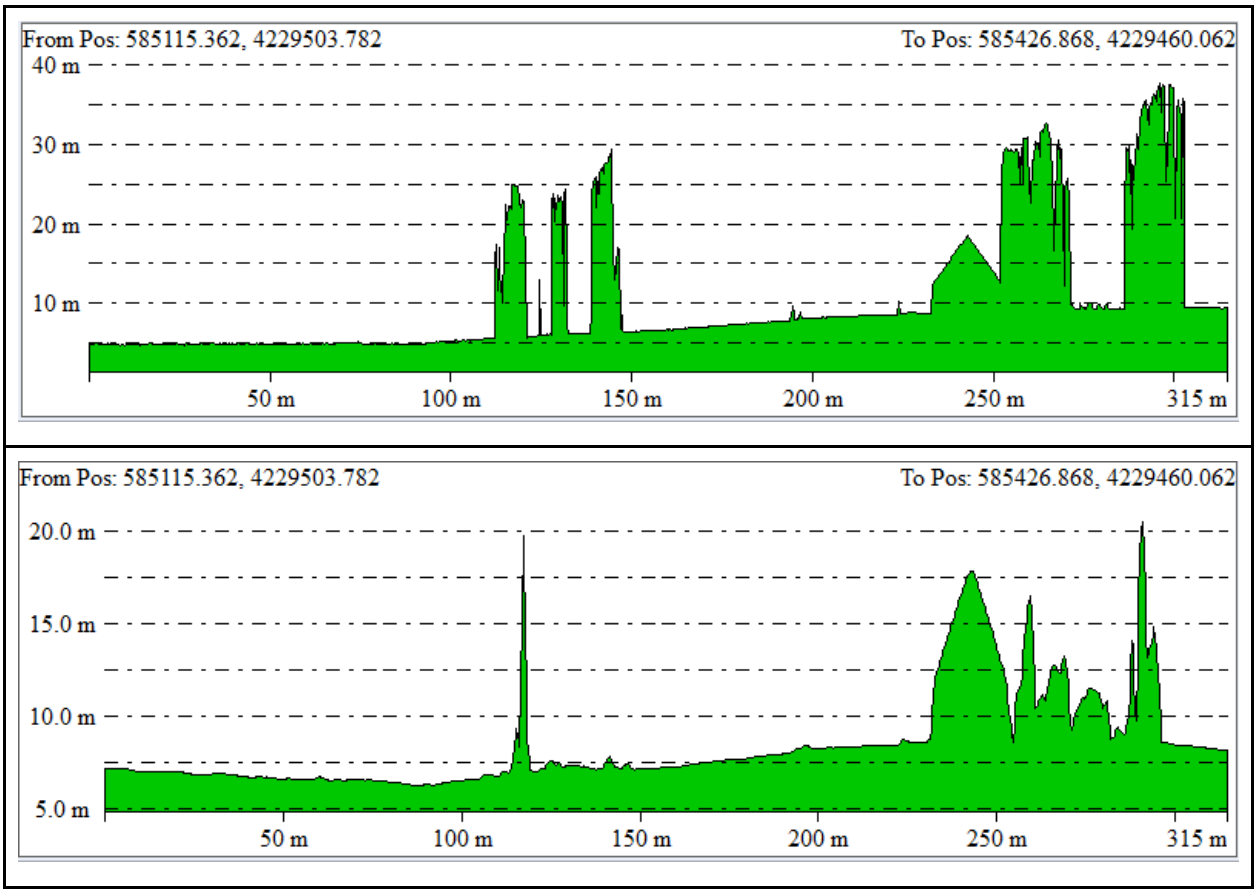


Figure 69. Comparison of a lidar (top) first surface profile through the Rush Ranch main compound and the same profile from the imagery-derived elevation model (bottom)

Overall, while the lidar data could be reasonably used to generate a canopy model, the SfM-derived elevation could not be combined with the lidar to provide anything meaningful, and generally was not a useful product. It is unclear to us why this is the case, as we have heard other reports of good success with SfM. Possibly this is a case of too much homogeneity combined with vegetation moving in the wind.

5 CONCLUSION

There were many changes during the course of the project that made some of the originally proposed research impossible. The intent and the explanations are as follows:

- We had proposed to fly each reserve twice to allow an analysis of repeatability of the UAS systems and provide two seasons of imagery. Cost constraints during contracting changed this to flying only one reserve, San Francisco Bay, twice. The higher complexity and expected greater seasonal signal were the primary criteria in selecting San Francisco Bay Research Reserve as the repeated reserve. Delays in obtaining a certificate of authorization and birding closures resulted in our first opportunity to fly the intended marsh with the correct vegetation state in September 2017, approximately 16 months from the start of the project. This left us with too little time to fly the marsh twice within the grant period, so we enlarged the area and flew it only once. Even if the grant period was not a constraint, we had no assurance that the marsh would not remain closed in the spring and early summer for bird nesting.

- We proposed to fly the same platform and sensor each time to evaluate the system in different ecosystems. However, the technology is moving so fast that it became apparent that we would be evaluating old technology by the time we flew the last research reserve. We opted to let the contractor select their most appropriate technology at the time of each flight. This resulted in increasingly better data as the ecosystems became more complex, simply because of the order of flights. While we lost the ability to compare the ecosystems based on the same technology, we gained insight into the speed of advancement and additional benefits of letting the private sector stay abreast of technology.
- We proposed to evaluate the imagery for use in habitat mapping with the intent to compare to the existing habitat maps. We knew we could only do that at Grand Bay and San Francisco Bay Research Reserves, since the Jacques Cousteau Research Reserve did not have a habitat map completed. However, the Grand Bay map was found to have enough questions about its quality that it was deemed unreliable for that comparison. Instead, we could only compare segmentation results with the same process used on manned or satellite imagery without something to supply a "truth," limiting the analysis.

We were able to test the spatial accuracies of the data. For both imagery and lidar, we found that ground control was required to attain the expected accuracies. For the imagery, we also found that the size of the area was important for attaining the required accuracy. If the area had to be broken into blocks for processing, whether for data volume or illumination consistency, each block needed to have enough control. In the San Francisco Bay Research Reserve, breaking into blocks resulted in approximately three to four control points per block, which was insufficient for attaining the required spatial accuracy. The lidar data did not appear to have this dependency.

Analysis of the elevation data derived from the imagery (using commercial structure from motion algorithms) indicated poor results in all cases and showed a lack of consistency. Possible explanations for these results include homogeneous landscape lacking clear features to tie images and movement of the vegetation in the wind between frames, particularly sidelap. It also indicates that structure from motion algorithms might have difficulty identifying coincident tie points in and on vegetation, and that vegetation density and camera look angles play variable roles in the relative success of this approach. Further research may be required to look at performance in an urban setting versus grasslands.

The poor results for imagery-derived elevation prevented an analysis of the possible benefits of combining lidar and imagery for vegetation height studies. The lidar appeared to do a better job than the imagery in capturing the tops of the vegetation, contrary to expectations. The lidar also performed very well when control was included in the acquisition, and exceeded our expectations for both accuracy and performance in dense marsh vegetation. Understanding the limitations of SfM and the environments where poor results are likely may be an important next step to avoid selecting the wrong technology in future work.

The analysis of imagery suitability and comparative advantage for habitat mapping had some shortcomings, but generally showed that the higher resolution imagery allowed for better and more detailed segmentation in the initial object creation step. The primary shortcoming was the lack of manned imagery that was both temporally similar to the UAS imagery and representative of the higher resolutions possible. This may have put the manned imagery at more of a disadvantage than would be realistic and essentially compared what you could get by paying for UAS flights versus what was available at no cost. However, it did show that we could use the UAS imagery for habitat mapping and that it did a good job in the initial automated step. As UAS imaging manufacturers continue to integrate ambient illumination sensors with downward-looking multispectral sensors, we expect the illumination issues to improve and more consistent radiometrically balanced imagery to become predictable.

We found that the mosaicked imagery from the MicaSense camera did very well in our resolving power test. Bars of approximately 2.5 centimeters (1 inch) width were resolved in imagery with 4 centimeter pixels, exceeding our expectations. This adds to our confidence in using this imagery where we need to resolve small targets. However, we also noticed the potential for blurring in the marsh grasses due to movement. Higher speed cameras would be required to solve that problem.

One of our objectives was to evaluate the costs for UAS collections compared to manned flights. Conducting this comparison was problematic for two reasons. First, what we paid for the UAS contract did not fully cover the costs incurred, since the private sector contributed an unknown amount. This would not be the expected mode of operation. Second, we did not pay for manned flights and we can only use previous experience regarding costs. On average, the UAS collections cost us over $20,000 per square mile collected for combined lidar and imagery, with

the lidar processing contributing a significant portion of that cost. For a manned flight, we might expect $15,000-$20,000 for mobilization costs. The per square mile costs for lidar (QL1) and imagery (4-band, 6-inch) are in the neighborhood of $300 and $100 respectively, making the mobilization the dominant cost for a small area. Those numbers are very rough and the products are not directly comparable, but they suggest that in areas somewhere around a square mile, the costs are comparable between manned and UAS. As the area shrinks, the UAS is favored.

5.1 Lessons Learned

We learned a number of valuable lessons during the course of this research. These were not part of the original research goals, but they became evident as we proceeded.

- Attaining the positional accuracy needed at UAS imagery resolutions that would support change analysis requires control with photo-id points. While standard procedure to mappers, the Jacques Cousteau data reinforced the point. Continued advancements in direct geopositioning may soon allow conducting UAS data acquisitions without control, though this will need to be rigorously tested before becoming standard practice.
- Trying to create a 4-band (red, green, blue, NIR) image by combining imagery from two cameras that were not flown simultaneously results in a product that is difficult to use for computational analysis. Anything that moved is displaced between the bands. This is particularly problematic with shadows and mobile objects. It is also problematic for objects that have considerable vertical structure relative to the sensor height, as the orthorectification process may not adequately adjust for differing view angles.
- The acquisition duration and number of images may put a limit on the reasonable area of coverage where UAS can make sense. Creating color-balanced mosaics may require capturing the amount of ambient light with each image. Setting stringent acquisition mission parameters such as sun angle, cloud cover, airspeed, flight AGL, frontlap, and sidelap may help improve product quality, but may also be too restrictive for efficiently covering large areas. Processing requiring seven months and lots of cloud processing after five days of collection for two square miles is neither cheap nor fast.
- With the possible exception of very small areas, UAS does not break the adage about "cheap, fast, and good; pick up to two." As the area grows, the small footprint of UAS becomes a serious problem, potentially making it more expensive, slower, and affecting quality. Our areas varied in size from the small 7-acre polygon at Grand Bay Research Reserve to the 2 square mile (1,280 acre) site at San Francisco Bay Research Reserve. Grand Bay included areas sized 7, 29, 50, and 91 acres and was the fastest delivery of any site. It also had the best spatial accuracy. It still required a couple months.
- The industry is in a rapidly evolving mode. New platforms, sensors, and processing approaches are becoming available all the time. Conclusions drawn from one experience, including this one, may have a limited shelf life.
- Private industry works hard to be a good partner to the public sector. This project was research for both sectors and pushed the bounds of what they had done previously. While not all the products met the hoped-for specifications and the turn-around times were longer than originally anticipated by anybody, the private sector put in significant resources of their own to make the best product they could under the conditions. They brought options to the table and discussed the pros and cons, involving the public sector in the decisions regarding platforms and sensors for a particular job. Had we taken the route of buying the platforms and sensors available at the start of this project and doing this project ourselves, we would have had an inferior product that required much more of our staff time. Although not every product met the specification, nor did we expect them too, it was a very successful public–private partnership.
- The combination of FAA regulations and the need to protect endangered species can lead to narrow or nonexistent operating windows. The allure of the UAS as a rapid deployment technology has to be tempered by the other factors that are not technology related.

6 Acknowledgments

The NOAA UAS Program supported this study through its grant program. We would like to thank the landowners within the research reserves: the State of New Jersey; the U.S. Fish and Wildlife Service; and Solano Land Trust for Jacques Cousteau, Grand Bay, and San Francisco Bay Research Reserve, respectively, for granting permission to fly

on their land. Staffs at the research reserves were invaluable to completing the work. Our point of contact within the UAS Program, John "JC" Coffey, provided a great deal of advice and guidance to help us navigate the complexities of NOAA UAS requirements. Patmarie Nedelka was instrumental in ensuring we complied with National Environmental Policy Act requirements. Management within the NOAA Office for Coastal Management, particularly Nicholas Schmidt, Mary Culver, and Erica Seiden, allowed staff members the opportunity to conduct this work. Perhaps most importantly, we would not have been able to do this work without the private sector. Quantum Spatial, Inc., and PrecisionHawk were excellent and trustworthy partners. They engaged in this project as a team effort to move the state of the art forward, contributing their own resources to make better products and push the envelope. While continually improving their technology and processes, they were very open about the issues they found collecting and processing the data, allowing us to better understand the data.

7 GLOSSARY

COA	certificate of authorization
DEM	digital elevation model
DSM	digital surface model
EASC	Eastern Administrative Support Center
FAA	Federal Aviation Administration
GBNERR	Grand Bay National Estuarine Research Reserve
GPS	Global Positioning System
JCNERR	Jacques Cousteau National Estuarine Research Reserve
MPiA	multiple pulses in air
NAVD88	North American Vertical Datum of 1988
NERRS	National Estuarine Research Reserve System
NGS	National Geodetic Survey
NIR	near infrared
NOAA	National Oceanic and Atmospheric Administration
NPS	nominal post spacing
NSSDA	National Standard for Spatial Data Accuracy
OCM	Office for Coastal Management
PH	PrecisionHawk
QSI	Quantum Spatial, Incorporated
RGB	red green blue
RMSE	root mean square error
RMSEz	vertical root mean square error
RTK	real time kinematic
SFBNERR	San Francisco Bay National Estuarine Research Reserve
SfM	structure from motion
TIN	triangulated irregular network
TRL	technology readiness level
UAS	unmanned aircraft systems
UAV	unmanned aerial vehicle
USAF	United States Air Force

8 BIBLIOGRAPHY

Gates, D. M., Keegan, H. J., Schleter, J. C., & Weidner, V. R. (1965). Spectral properties of plants. *Applied Optics, 4*(1), 11-20. Retrieved 5 8, 2018 from https://osapublishing.org/abstract.cfm?uri=ao-4-1-11

Wu, K.-L., & Yang, M.-S. (2007). Mean shift-based clustering. *Pattern Recognition, 40*(11), 3035-3052. Retrieved 3 31, 2018 from http://dblp.uni-trier.de/db/journals/pr/pr40.html