Theses and Dissertations   1. Thesis and Dissertation Collection, all items

2019-12

TERRAIN CATEGORIZATION CAPABILITIES OF LIDAR SYSTEMS OVER DENSELY VEGETATED AREA

Sonarch, Chedpong

Monterey, CA; Naval Postgraduate School

http://hdl.handle.net/10945/64073

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**TERRAIN CATEGORIZATION CAPABILITIES OF
LIDAR SYSTEMS OVER DENSELY VEGETATED AREA**

by

Chedpong Sonarch

December 2019

Thesis Advisor:	Richard C. Olsen
Co-Advisor:	Jeremy P. Metcalf

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		*Form Approved OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE December 2019	3. REPORT TYPE AND DATES COVERED Master's thesis	
4. TITLE AND SUBTITLE TERRAIN CATEGORIZATION CAPABILITIES OF LIDAR SYSTEMS OVER DENSELY VEGETATED AREA		5. FUNDING NUMBERS	
6. AUTHOR(S) Chedpong Sonarch			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		10. SPONSORING / MONITORING AGENCY REPORT NUMBER	

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	12b. DISTRIBUTION CODE A

13. ABSTRACT (maximum 200 words)

Light Detection and Ranging (LiDAR) technology has various useful applications, such as in surveying tasks. This study continues research previously conducted at the Naval Postgraduate School by Andrew S. Davis, and documented in his thesis, "Forestry Identification with LiDAR Waveform and Point Clouds." The present study, which aims to evaluate the classification capability of LiDAR systems over various tree species in a particular area of interest, collected sample data over Point Lobos State Park, California. The data set was separated into two categories, aerial platform and ground survey. The aerial platform consisted of an Optech Titan system and Airborne Hydrography AB Chiroptera system (AHAB). Analysis was performed by comparing the results from the ENVI software classifier and the actual location of tree species from ground surveying. The study also extracted the features of waveforms from each tree species and used these features to distinguish them from among other samples of tree species and their surrounding environment, such as roads and trails. The classifications were done by the classifier tools provided in ENVI (K-means, Spectral Angle Mapper, and Support Vector Machine). The results showed that waveform data can accurately distinguish class samples. The analysis also pointed out that the most common error occurred when classes had a narrow gap in data values and shared similar pulse shapes.

14. SUBJECT TERMS LiDAR, terrain classification, LiDAR system, LiDAR waveform			15. NUMBER OF PAGES 87
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

TERRAIN CATEGORIZATION CAPABILITIES OF LIDAR SYSTEMS OVER DENSELY VEGETATED AREA

Chedpong Sonarch
Lieutenant Junior Grade, Royal Thai Navy
BEE, Royal Thai Naval Academy, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED PHYSICS

from the

NAVAL POSTGRADUATE SCHOOL
December 2019

Approved by: Richard C. Olsen
Advisor

Jeremy P. Metcalf
Co-Advisor

Kevin B. Smith
Chair, Department of Physics

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Light Detection and Ranging (LiDAR) technology has various useful applications, such as in surveying tasks. This study continues research previously conducted at the Naval Postgraduate School by Andrew S. Davis, and documented in his thesis, "Forestry Identification with LiDAR Waveform and Point Clouds." The present study, which aims to evaluate the classification capability of LiDAR systems over various tree species in a particular area of interest, collected sample data over Point Lobos State Park, California. The data set was separated into two categories, aerial platform and ground survey. The aerial platform consisted of an Optech Titan system and Airborne Hydrography AB Chiroptera system (AHAB). Analysis was performed by comparing the results from the ENVI software classifier and the actual location of tree species from ground surveying. The study also extracted the features of waveforms from each tree species and used these features to distinguish them from among other samples of tree species and their surrounding environment, such as roads and trails. The classifications were done by the classifier tools provided in ENVI (K-means, Spectral Angle Mapper, and Support Vector Machine). The results showed that waveform data can accurately distinguish class samples. The analysis also pointed out that the most common error occurred when classes had a narrow gap in data values and shared similar pulse shapes.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

x

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

2D	Two-dimensions
3D	Three-dimensions
AHAB	Airborne Hydrography AB
ALS	Airborne Laser Scanning
CHM	Canopy Height Model
DBH	Diameter at Breast Height
DEM	Digital Elevation Model
DSM	Digital Surface Model
DTM	Digital Terrain Model
EM	Electromagnetic
ENVI	Environment for Visualizing Images
GLAS	Geoscience Laser Altimeter System
GPS	Global Positioning System
ICESat	Ice, Cloud and land Elevation Satellite
IDL	Interactive Data Language
IHO	International Hydrographic Organization
IMU	Inertial Measurement Unit
INS	Inertial Navigation System
LASER	Light Amplification by Stimulated Emission of Radiation
LiDAR	Light Detection and Ranging
LM	Local Maxima
LVIS	Laser Vegetation Imaging Sensor
MR	Multi-resolution
NASA	National Aeronautics and Space Administration
NCALM	National Center for Airborne Laser Mapping
NIR	Near-Infrared
nm	Nano meter
PRF	Pulse Repetition Frequencies
RADAR	Radio Detection and Ranging
RG	Region Growing

ROI	Region of Interest
RTK	Real-Time Kinematic
SAM	Spectral Angle Mapper
SVM	Support Vector Machine
SWIR	Shortwave Infrared
TERCAT	Terrain Categorization

ACKNOWLEDGMENTS

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PURPOSE OF RESEARCH

Light Detection and Ranging (LiDAR) provides an essential new tool for mapping and surveying both urban and forested environments. Among aircraft-mounted LiDAR systems, Airborne Laser Scanning (ALS) has become operationally viable for commercial surveying. In particular, the Optech ALTM 3100, released circa 2005, marked the advent of an ever-expanding commercial industry (Shan and Toth 2018).

The laser profiling and scanning system, or LiDAR system, has capabilities for measuring the distance from the source to the target, collecting data as a composite of actual locations and elevation data, and creating a model of the artificial structure of the collected target (ArcGIS Desktop n.d.). One primary area of interest for the application of LiDAR capabilities is forestry, as forests are significant to both organisms living in them and humans (Yilmaz et al. 2017). In terms of forestry, the LiDAR system is helpful for, among other things, studying the increase in tree structures and the decrease of forest areas (Leeuwen and Disney 2017). Despite the popularity of this technology, the capability of LiDAR in various cases still needs to be evaluated. In particular, the so-called 'waveform' data utility is a question still of interest.

This research investigates the utility and use of LiDAR waveform data, following a previous thesis effort at the Naval Postgraduate School (Olsen, Davis, and Metcalf 2019). At the end of the prior research, it was suggested a future researcher further study the characteristics of waveform data obtained from other LiDAR systems (Davis 2018). Following Davis's suggestion, this research focuses on and evaluates LiDAR data's ability to facilitate accurate identification of vegetation in an area of interest.

B. OBJECTIVE

The objective of this research is to evaluate the Terrain Categorization (TERCAT) capabilities of LiDAR systems by performing both ground observation and analysis of LiDAR's data of low-growing plant species found within Point Lobos State Park, located in Carmel-by-the-Sea, California. The study evaluates the capability of a LiDAR system

to distinguish various tree species in the area of interest. Furthermore, the study compares the characteristics of low vegetation's spectral profile to the characteristics of its surrounding environment. The comparison is to validate the unique pattern of the spectrum of each object in the area of interest.

C. THESIS OVERVIEW

There are five chapters included in this research paper. The current chapter (Chapter I) presents the purpose of this research, its scope, and related topics reviewed. Chapter II provides readers with a history of the LiDAR system, the fundamentals of LiDAR, and case studies of data analyzing techniques in terms of forest study. Chapter III presents the data collection and preparation methods. Chapter IV presents the methodology for analyzing and evaluating the data, as well as the results of this research. Finally, Chapter V concludes this research and offers recommendations and considerations for future work.

II.   BACKGROUND

This chapter provides essential background information about LiDAR, including its history, and explains how the system is used. Also, the fundamentals of LiDAR are covered to explain how it works. Finally, this chapter also provides examples of LiDAR data analysis techniques in terms of forest study.

A.   HISTORY OF LIDAR

LiDAR is a technique that measures the distance to an object by firing a laser pulse, hitting that object, and reflecting a detector of the system (Wasser n.d.). According to the Encyclopedia Britannica, "LiDAR Scientific Technique" written by Erik Gregersen, description of the beginnings of the LiDAR system, the technology emerged in the 1930s when the first attempt to measure distance by using light beams was made. The study used searchlights rather than a laser to examine the atmosphere's structure. A few years later, in 1938, light pulses were used to determine the height of clouds (Gregersen 2016). With the advent of the laser (Schawlow and Townes 1958), short, intense pulses of light became practical. The q-switch allowed for pulses extending over fractions of micro-seconds (McClung and Hellwarth 1962). By 1965, experiments in altitude ranging were being conducted from aircraft (Miller 1965).

In the 1970s, attempts to observe the properties of the ocean water and atmosphere, ice sheets, and the forest canopy for topographic mapping were made by National Aeronautics and Space Administration (NASA). In the early years, the limitations of a reliable altitude measurement limited the utility of such systems. With the advent of the Global Positioning System (GPS) in the early 1990s, and more reliable inertial measurement units (IMU), the technology came into its own as the sensor was of concern (Wikipedia 2019).

By the mid-1990s, the manufacturers of laser scanners were introducing LiDAR sensors that could transmit 10,000 to 100,000 pulses per second for topographic mapping purposes (BCC Research 2018). The previously noted Optech ALTM 3100, for instance, could be operated at pulse repetition frequencies (PRF) of 33–100 kHz, depending on

altitude (Olsen 2007). Since instruments and techniques were introduced that made LiDAR systems more reliable and increased users interested in this field, the applications of LiDAR have gradually been included in everyday life, such as in navigation systems. The growth of the study and the use of LiDAR systems will undoubtedly lead to more development in the future.

B.   FUNDAMENTALS OF LIDAR

Since LiDAR is one of the active remote sensing techniques, its fundamentals are related to those of the laser. All laser ranging, profiling, and scanning applications involve the use of some type of laser-based ranging. A laser rangefinder, for example, is an instrument that measures the distance to an object, and that measurement can be used for a variety of purposes (Shan and Toth 2018). One such purpose is to analyze the content of the Earth's surface. Furthermore, depending on the type of sensor and the user's purposes, the data under consideration can be generated as either 2D or 3D form.

C.   COMPONENTS OF LIDAR SYSTEMS

By definition, the phrase " active remote sensing system" means the system can emit energy by itself, such as electromagnetic wave or light, to determine objects or any type of structure on the ground without actually physically measuring (i.e., using a sensor instead of measuring by hand) (Wasser n.d.). The source, a light wave in this case, transmits to the target and reflects a sensor detector (see Figure 1). As indicated in Figure 2, a LiDAR system typically comprises four main components: laser, GPS, an inertial navigation system (INS), and computer. A laser system is used to obtain a measurement by firing energy or pulses to the target and recording the return pulses, which provide target information such as the properties of targets. The second component, the GPS, is used to determine the location of each point of data we obtain, and this will combine with the output from the laser system as raw data. INS is used to control roll, pitch, yaw, and heading of the platform. A computer system is used to store data (Figure 3) (ArcGIS Desktop n.d.).
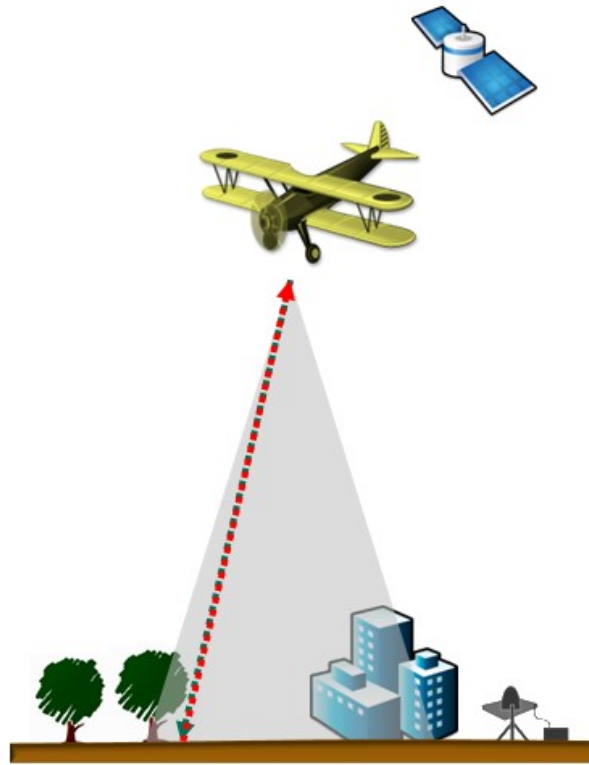
Figure 1. Showing how a LiDAR system operates. Source: ArcGIS Desktop (n.d.).
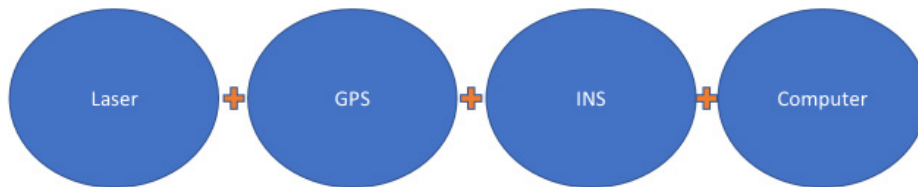
Figure 2. LiDAR system components.
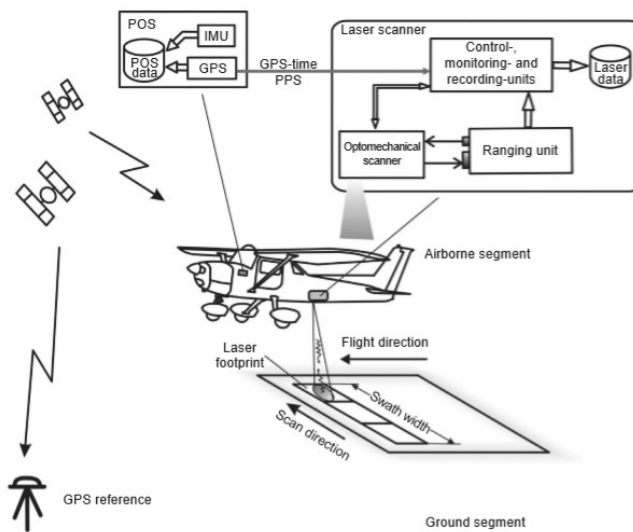
Figure 3.Overview of LiDAR system components' detailed operation.
Source: Shan and Toth (2018).

D.   TYPES OF LIDAR

The textbook *Topographic Laser Ranging and Scanning* explains that the two main types of LiDAR applications are based on the type of carrier. The first type is airborne and spaceborne applications. The second is terrestrial applications (Shan and Toth 2018).

Based on the description from the website ArcGIS Desktop, airborne and spaceborne applications include topographic LiDAR and bathymetric LiDAR (Figure 4). A topographic LiDAR application is the study of the contents of areas on land. On the other hand, a bathymetric LiDAR involves the study of water and sea surface.
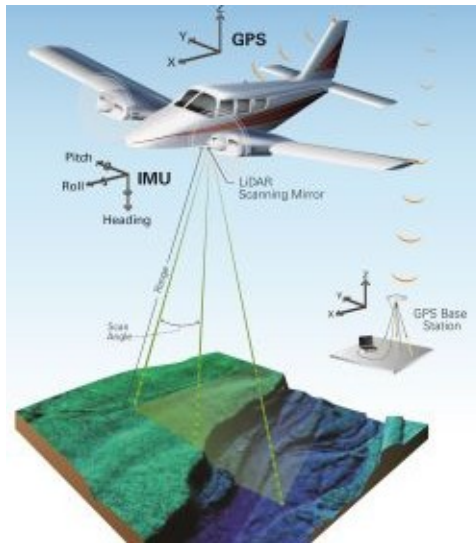
Figure 4. Airborne and spaceborne applications. Source: LiDAR News (2016).

Terrestrial applications include mobile and static techniques. A mobile LiDAR consists of a moving vehicle on which a LiDAR sensor is mounted (Figure 5). This type of application is mostly used to observe roads, light poles, and infrastructure.
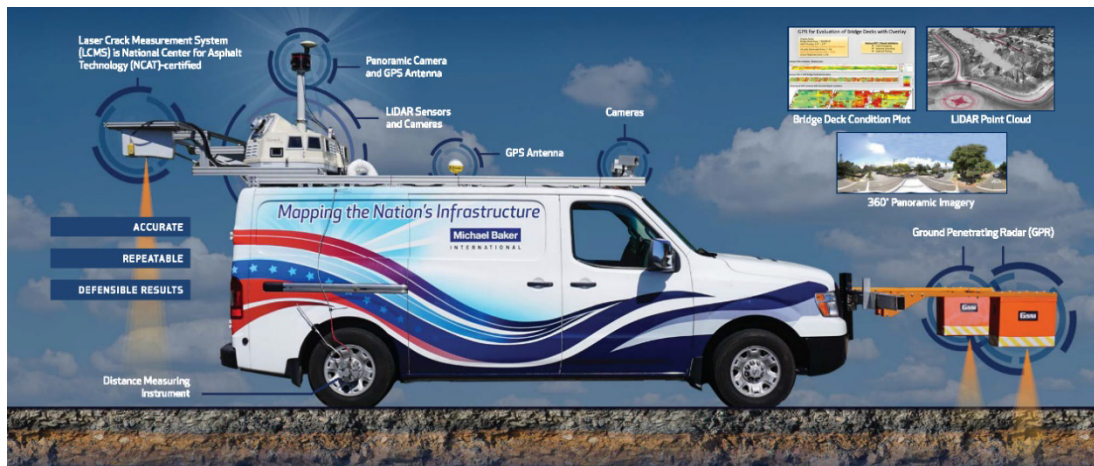
Figure 5. Mobile LiDAR. Source: Baker Mobile LiDAR (2016).

A static LiDAR consists of data collected from a static location (Figure 6). The LiDAR sensor is mounted on a tripod. Mostly, this technique is used for mining, surveying, and archaeology (ArcGIS Desktop n.d.).

Figure 6. Example of static LIDAR, Leica Scan Station P50. Source:
HEXAGON GEOSYSTEMS (n.d.).

E. DATA TYPES – DISCRETE RETURNS AND WAVEFORMS

Commonly, a LiDAR system is operated in visible light (532 nm) and near-infrared (NIR, 1,064 nm, and 1,555 nm) ranges of the electromagnetic spectrum. The selection of operating wavelength depends on the purpose of use (Olsen 2016). A wavelength of 532 nm is proper for use in bathymetric LiDAR applications because of its ability to penetrate through the water with less attenuation when compared to other options. On the other side, the NIR beam is better when measuring land or sea surface (Crutchley and Crow 2010).

Since LiDAR data measures reflections of energy or returned pulse, we can group the types of data that are recorded in two different ways. The first is a discrete return. Discrete data are extracted from a point at each peak of the laser pulse. The second is waveform data. Waveform data provides a distribution of entirely returned pulse (Wasser n.d.). Following Figure 7, we see that waveform data gives us more detail of a target than

discrete data. Nonetheless, this does not mean that discrete data is not worth using. In contrast to waveform data, discrete data is less complex, lower in cost, and requires less processing time.

"The Light Fantastic Using Airborne Lidar in an Archaeological Survey" by Simon Crutchley and Peter Crow (2010) gives descriptions of how to interpret LiDAR's return pluses. The peaks of waveform indicate the order of the data number. The first return typically represents a top of the target, such as the top of a tree or the top of a building. The last return usually represents the ground or the base of an object. The first and the last returns are considered the most significant data because these data are involved in calculating the Digital Surface Model (DSM) and Digital Terrain Model (DTM) (Crutchley and Crow 2010). The DTM and DSM are necessary for evaluating height data. The other returns between the first and the last return may represent shapes, characteristics, and detail of an object.
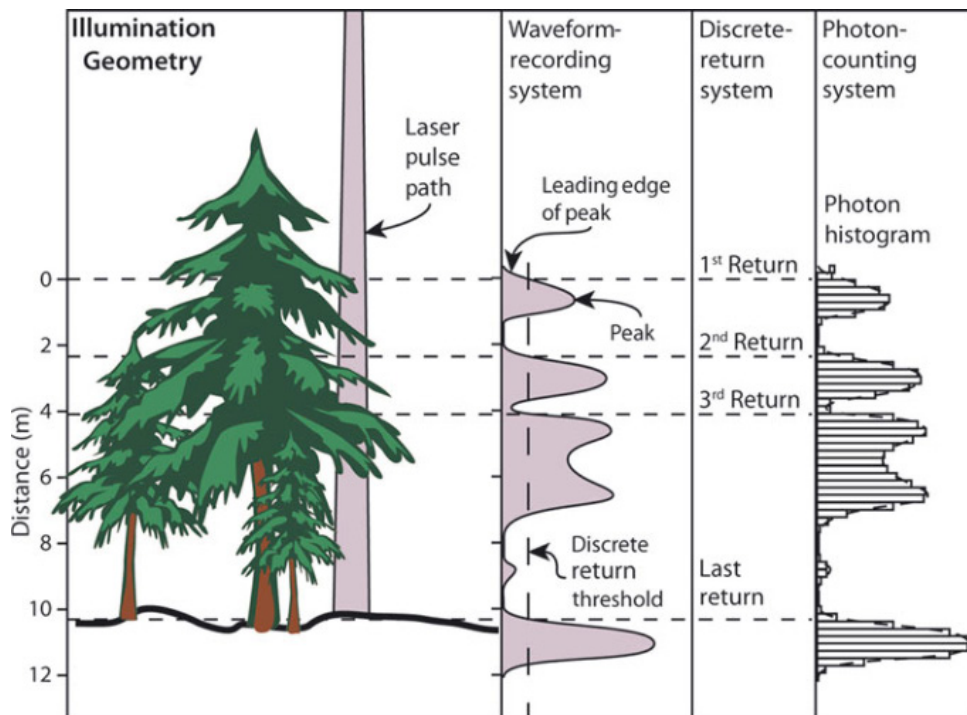
Figure 7. Full-waveform data vs. discrete data. Source: Deems, Painter, and Finnegan (2013).

9

## F.	WAVEFORM LIDAR

LiDAR waveform data can be considered in a few general forms: terrestrial systems that have a generally large footprint (approximately 100 m, ICESat/GLAS) or small footprint (Airborne RIEGL systems, typically 50 cm to 1 m); and bathymetric LiDAR (intrinsically "small footprint" systems in general).

### 1.	Bathymetric LiDAR

Bathymetric studies are a primary application of airborne laser scanners, and waveform data are intrinsic to such measurements. There is a large body of work on such systems, largely out of the scope and topics analyzed here. Guenther et al. (2000) illustrate the complexity of the problem. Guenther points out two main problems that occur while dealing with bathymetry studies: the accuracy and reliable determination of the location. Based on the International Hydrographic Organization (IHO) standard, the vertical accuracy accepted for shallow water hydrography is about 25 cm from all sources. The factors are water depth, and the interface between air and water. Due to the factors mentioned previously, the use of green laser range only is not acceptable. It requires at least two wide ranges of laser sources, such as red and IR, to maintain acceptable accuracy. The second dilemma is at the interface between vigorous strong water and unsteady bottom returns, which cause fluctuating results. This limitation must be considered and fixed by configuring the ability of a detector to avoid irregular effects on the results. Furthermore, the measurement result has to be formed into the acceptable digitizer ranges (Guenther et al. 2000).

Collin et al. (2011) illustrate many features in a typical LiDAR return for a bathymetric system. The integration of environmental factors performing at the interface between air and water (Figure 8), as well as underwater, and instrument factors lead to a denoising method. The shape and the returned signal power depend on the depth of the water because of scattering and absorption (Figure 9). Therefore, the regression method can be applied to correct the statistical variable at different depth levels, and the only variance analysis was established for the residuals (Collin, Long, and Archambault 2011).
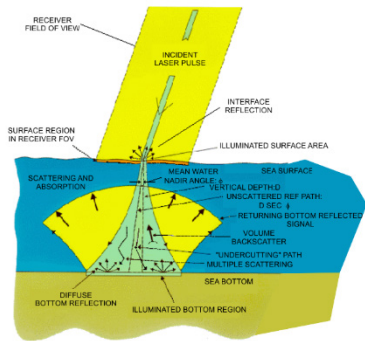
Figure 8.Schematic diagram of the effects of scattering on the green LiDAR beam. Source: Guenther et al. (2000).
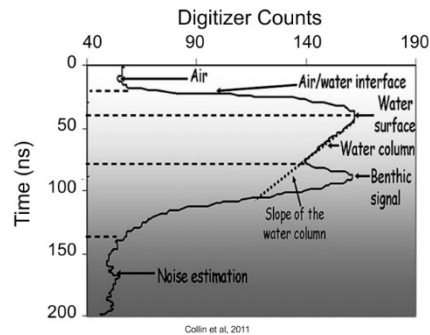
Figure 9.Segmentation of a generic bathymetric LiDAR backscatter. Source: Collin, Long, and Archambault (2011).

2.   Large Footprint Systems

The Earth researchers have used 'large footprint' systems for several years, particularly the ICESat satellite sensor and the airborne Laser Vegetation Imaging Sensor (LVIS) for mapping terrestrial properties and the evaluation of digital elevation models. The priority to obtain a large scale of observation in these uses, however, limits the spatial resolution of these systems.

LVIS has been flown routinely since 1998 by the NASA Goddard Space Flight Center group (Blair, Rabine, and Hofton 1999). Similar technology was flown on the space shuttle (Bufton, Harding, and Garvin 1999). Eventually, this design evolved into the Geoscience Laser Altimeter System (GLAS) launched as the primary payload for the Ice, Cloud, and land Elevation Satellite (ICESat) mission in 2003 ("ICESat" 2019). The ability to collect data about vegetation structure (height and density) and the spatial distribution of vegetation within the Geoscience Laser Altimeter System (GLAS) footprint were limited but could be addressed by comparing ground measurements and modeling of the LiDAR returns (L. Neuenschwander et al. 2006). Hofton et al. (2000),show a variety of Gaussian decomposition techniques appropriate to systems with

50–100 m spot sizes. The waveform can be a simple shape, as shown in Figure 10a, or complex and multimodal shape, as shown in Figure 10b. Typically, a simple waveform represents the ocean and bare-ground regions. The complex waveform represents vegetated areas (Hofton, Minster, and Blair 2000).

In general, the surface type and the return pulse shape are not considered by current waveform processing methods. These factors lead to the problems in data processing because of inconsistent sample ranges to a reflecting area. Thus, the techniques locate the crown of the amplitude of the waveform or point out the centroid of the return waveform. In the case of a simple waveform performed by a single mode, the methods previously mentioned are reliable for locating the reflecting surface of interest. On the other hand, for a complex waveform, the area of interest may have attenuation and absorption, which can cause attenuation in the return pulse amplitude. Consequently, the significant information, i.e., the centroid, may not be delivered with the greatest accuracy. Therefore, by decomposing and turning those return pulses into several fragments can provide a better way to estimate the structure of the waveform, its location, and its geolocation (Hofton, Minster, and Blair 2000).
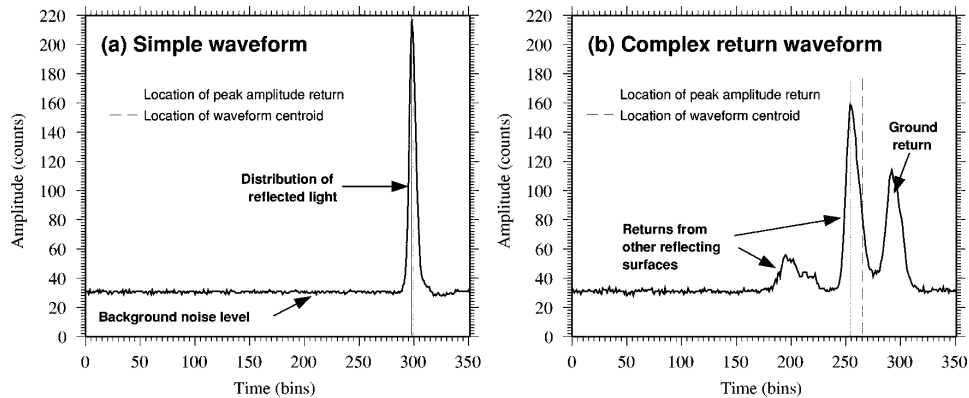
"The waveform represents the distribution of light reflected from all intercepted surfaces within the footprint. The positions of the peak amplitude and centroid of the waveform are shown by solid and dashed lines, respectively. (a) Simple and (b) complex laser altimeter return waveforms, collected by the LVIS airborne altimeter system" (Hofton, Minster, and Blair 2000).

Figure 10.   Return pulse waveform. Source: Hofton, Minster, and Blair (2000).

Despite some of the indications just noted, and also noting the multi-year period required to cover the full earth, there have been several successful applications of ICESat/ GLAS data to forestry (Harding et al. 2002). The study of David J. Harding et al. (2002), for instance, focused on evaluating the accuracy of measuring topography, vegetation height, and attributes of canopy structure and function. The result showed that LiDAR can produce accurate results. Those authors also predicted that a variety of applications relating to the use of LiDAR would increase (Harding et al. 2002). In the same year, Hudak et al. (2002) conducted a study of the integration of LiDAR and Landsat ETM+ data for estimating and mapping forest canopy height. The result of their research showed that the combination of LiDAR and Landsat data can improve the accuracy of estimating and mapping forest canopy height (Hudak et al. 2002).

Further studies focused on evaluating the ability of ICESat waveforms to estimate forest canopy height. Results from such work showed that the use of ICESat's waveforms and the collection of data from topographic relief can successfully estimate forest height and aboveground biomass (Figure 11) (Lefsky et al. 2005). The preceding studies are only a few examples related to the study of the forest by using LiDAR. There are still several examples that show how LiDAR became used in forest study; some of them are discussed in the next section.
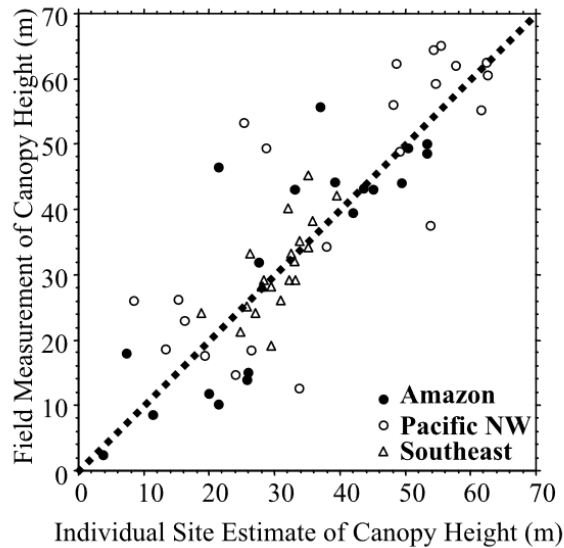
Figure 11.　　　Observed forest maximum canopy height vs. ICESat
estimates of same, for three study areas. Source: Lefsky et al. (2005).

3.　　　Airborne LiDAR – Small Footprint Systems

LiDAR technologies give us worthwhile information about forestry systems. LiDAR data provide us high-resolution data and a three-dimensional (3D) structure of targets. Because of these useful features, we can extract tree height, tree structure, and tree density, and even track changing of forest areas (Leeuwen and Disney 2017). Small-footprint airborne LiDAR has been used on a wide scale of applications, both commercial and research. Small-footprint airborne LiDAR, also referred to as airborne laser scanning or ALS, gives us the best accuracy when we perform terrain elevation and vegetation measurement (Popescu et al. 2011). Small-footprint LiDAR generally ranges from 10 to 30 cm in scale in the case of the discrete return system. Thus, it increases the probability of a beam to penetrate through the ground in a densely vegetated area. In the case of waveform return, the return pulse in a densely area possibly can associate with a single object. Yet, this technique also has drawbacks, such as the small beam may be completely absorbed by the tree before it hits the ground, and due to its small scale, the beam may not cover the peak of the tree, or the beam may double count the tops (Lewis and Hancock 2007).

14
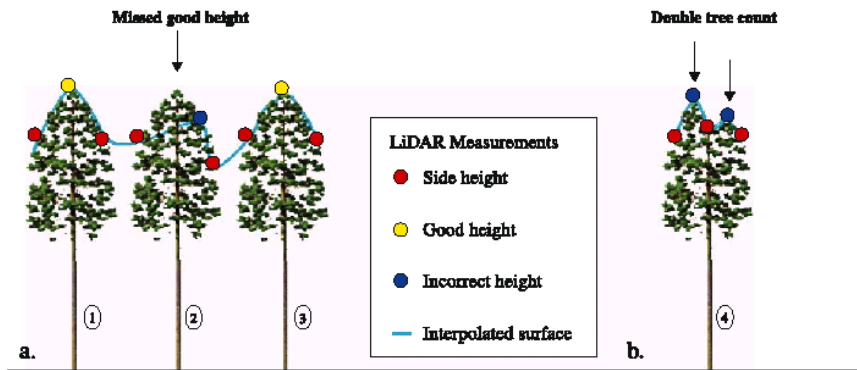
Fig. 5. Tree height variance can be inflated due to misperceived tree heights (from large post spacing) within the tree height-finding model. (a) Height of trees 1 and 3 is measured correctly because LiDAR returns intercept tree peaks (yellow). Height of tree 2 is incorrectly measured because the LiDAR return is from the side of the crown (blue). (b) Tree 4 is counted as two stems (and heights) due to a forked or irregular tree crown.

Figure 12.   Small-footprint issues. Source: Lewis and Hancock (2007).

There are several practical examples of using LiDAR in forestry work, such as in California. In a fire-prone state like California, the change in forest volume and abundance is significant to all creatures in those surroundings. Based on the study of Maggi Kelly and Stefania Di Tommaso (2015), it is evident that since 2000, several researchers have been interested in using LiDAR systems to map California's forests. LiDAR systems have been used to capture tree structures, map individual trees, predict quantity and biomass of forests, develop plans to deal with fire in forests, and map topography and infrastructure of forests (Kelly and Di Tommaso 2015).

Small-footprint LiDAR studies for tree species identification were done with the TopEye MK II sensor. Josef Reitberger, Peter Kyzystek, and Uwe Stiller (2006) conducted this research by using three main concepts. The first idea was to describe tree crowns by a segmentation of the canopy height model. The second was to extract characteristic features of the individual trees. Finally, an appropriate classifier was used to classify tree species, based on the extracted features of tree species (Reitberger, Krzystek, and Stilla 2006).

Specifically, Reitberger et al. (2006) focused on extracting a unique geometry of tree structures and intensity of waveform decomposition. The parameters of this study were salient features that include outer tree geometry (S_g), the internal geometrical tree

15

structure (S_i), and the intensity-related tree structure (S_I). The external geometry consisted of two variables, S_g^1 and S_g^2. The S_g^1 is evaluated by the crown shape surface. The S_g^2 is calculated by the mean radii, which are determined as the mean distances of all layer points to the tree trunk (Figure 13a). The internal geometrical S_i is separated into two groups, S_i^h and S_i^d. The S_i^h is considered the percentiles of LiDAR point height distribution of the tree (Figure 13b). The S_i^d is the normalized number of the LiDAR points distributed on each layer divided by the total number of LiDAR points. The S_I^1 is the mean intensity value of each layer. The S_I^2 is the mean value of the entire tree segment.
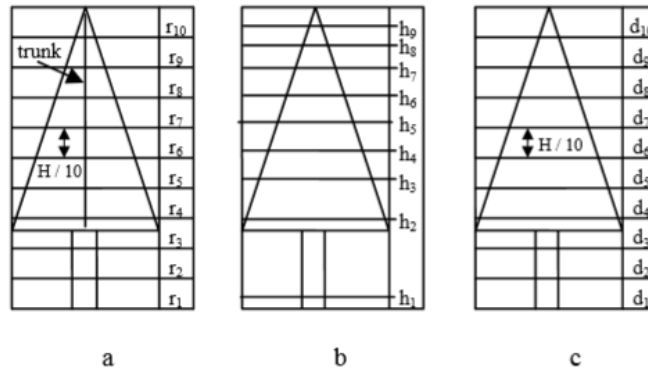
Figure 13.   Tree layer distribution. Source: Reitberger, Krzystek, and Stilla (2006).

The results in Table 1 show that waveform decomposition was generating more points than the original data collected by the TopEye system. The results in Table 2 show that the outer tree geometry S_g is proper for evaluating coniferous trees and the intensity-related tree structure S_I. is proper for deciduous trees (Reitberger, Krzystek, and Stilla 2006). From these results, one can see the performance and accuracy of the segmentation of Canopy Height Model (CHM) techniques in terms of forest study.

Table 1. Comparing points presented by TopEye and points presented by waveform decomposition. Source: Reitberger, Krzystek, and Stilla (2006).

Area		1	2	3	4	5
Tree specie / object type		Deci-duous (leaf-on)	Deci-duous (leaf-off)	Coni-ferous	Deci-duous (leaf-on) and coni-ferous	Mea-dow
Size [m²]		21.9	72.2	22.2	86.7	28.3
Points from Top-Eye	Total	768	5594	1109	1602	362
	First	503	4168	882	1191	362
	Last	265	1426	227	411	0
Points derived from wave-forms	Total	943	7436	2555	3261	456
	First (%)	553 (59)	4648 (62)	1483 (58)	1678 (51)	456 (100)
	Last (%)	280 (30)	1548 (21)	727 (28)	969 (30)	0 (0)
	Middle (%)	110 (11)	1240 (27)	345 (14)	614 (19)	0 (0)

Table 2. Evaluating accuracy of classification. Source: Reitberger, Krzystek, and Stilla (2006).

Saliency	S_g^1	S_g^2	S_i^h	S_i^d	S_I^1	S_I^2	$S_g^1 + S_I^2$	$S_g^2 + S_I^2$
Conif. [%]	86	94	54	69	79	80	80	80
Decid. [%]	65	65	74	61	65	83	83	82
Total [%]	81	88	58	67	77	81	81	80

G. CANOPY HEIGHT MODEL

A canopy height model, or CHM, is one of the essential products needed in forest studies. The CHM provides the height data of the trees and relevant forest information (L. Wasser 2018). The basic approach to generating the CHM is to subtract the digital elevation model (DEM) or digital terrain model (DTM) from the digital surface model

(DSM) (L. Wasser 2018; Marcu, Stătescu, and Iurist 2017). Figure 14 illustrates the terms DSM, DTM, and shows how the CHM is the difference between these terms (Marcu, Stătescu, and Iurist 2017); an outline of the process is depicted in Figure 15. The DSM is extracted from the first returns of the LiDAR data. The DTM, which is the elevation of the earth's surface, is extracted from last return (Figure 16). So, the difference between the DSM and the DTM indicates the height of objects, which in this case is CHM (as shown in an equation below).

$$DSM - DTM = CHM$$
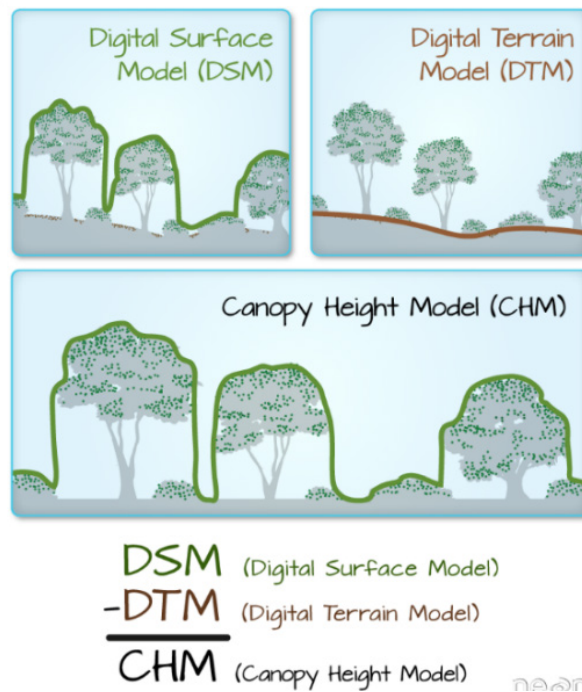
Figure 14.   Showing how to derive CHM. Source: L. Wasser (2018).
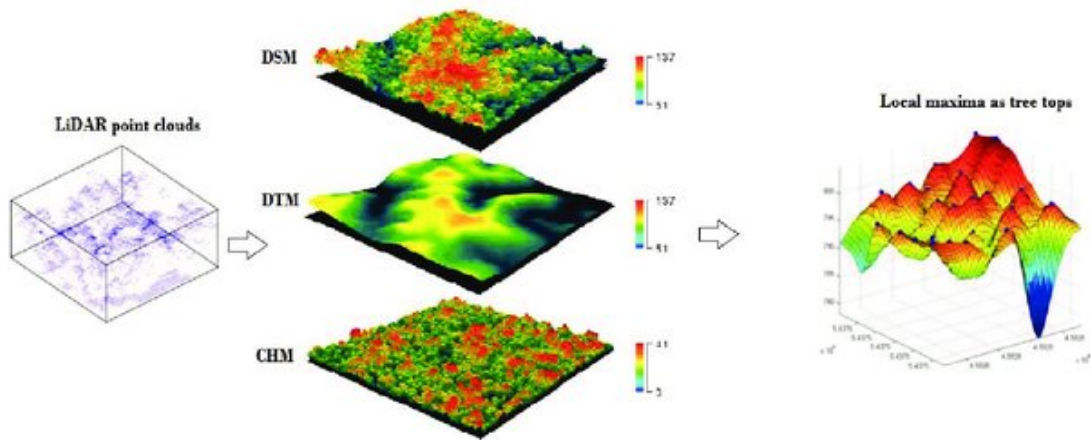
Figure 15.   Outline of the process to obtain the Canopy Height Model (CHM) and extract tree height from local maxima. Source: Dang (2012).
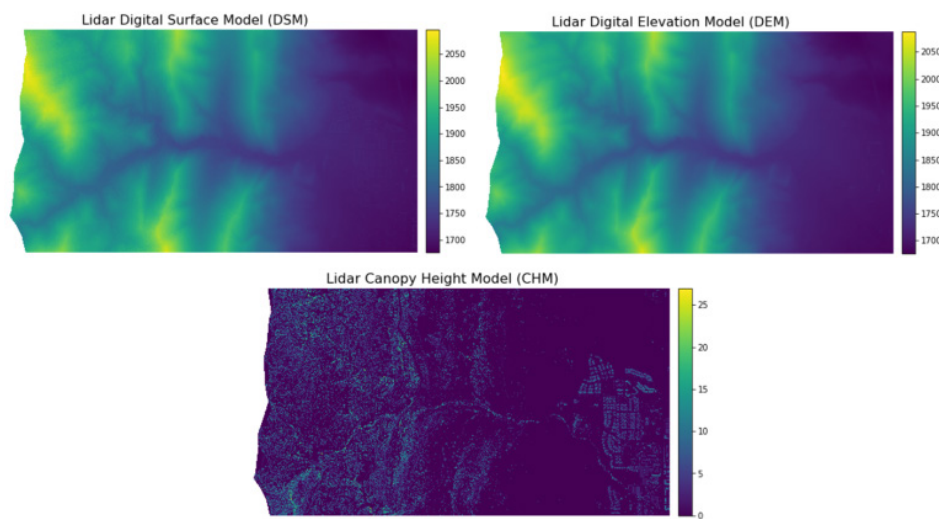
Figure 16.   Examples of DSM, DEM, and CHM extracted from LiDAR data. Source: L. Wasser (2018).

In terms of analyzing CHM, there are several image processing methods that are used. Mainly, CHM is used as input for the segmentation technique of tree crowns and calculation of forest parameters, such as biomass (Rizaev, Pogorelov, and Krivova 2016).

H.   SEGMENTATION METHODS

Segmentation methods are used as a threshold to eliminate data that are not expected to be used. Segmentation allows for a boundary to group the types, categories, and expected region of data. Volcan Yilmaz et al. (2017) provide an excellent example of the multi-resolution (MR) and region growing (RG) segmentation techniques. The MR segmentation technique is a technique related to a homogeneity criterion. This homogeneity criterion considers the color and shape criteria related to three parameters: scale, shape, and compactness. The size of an object is determined by the scale parameter. The ratio of the boundary of an object and the square root of the number of pixels in the object is defined by the compactness parameter. Smoothness is the ratio of the boundary of an object to the boundary of the bounding box of the object. Yilmaz et al.'s (2017) study used the eCognition software to perform MR segmentation. The study set the scale, shape, and compactness parameters as 15, 0.4, and 0.5, respectively. The resulting illustration after performing MR segmentation is shown in Figure 17.
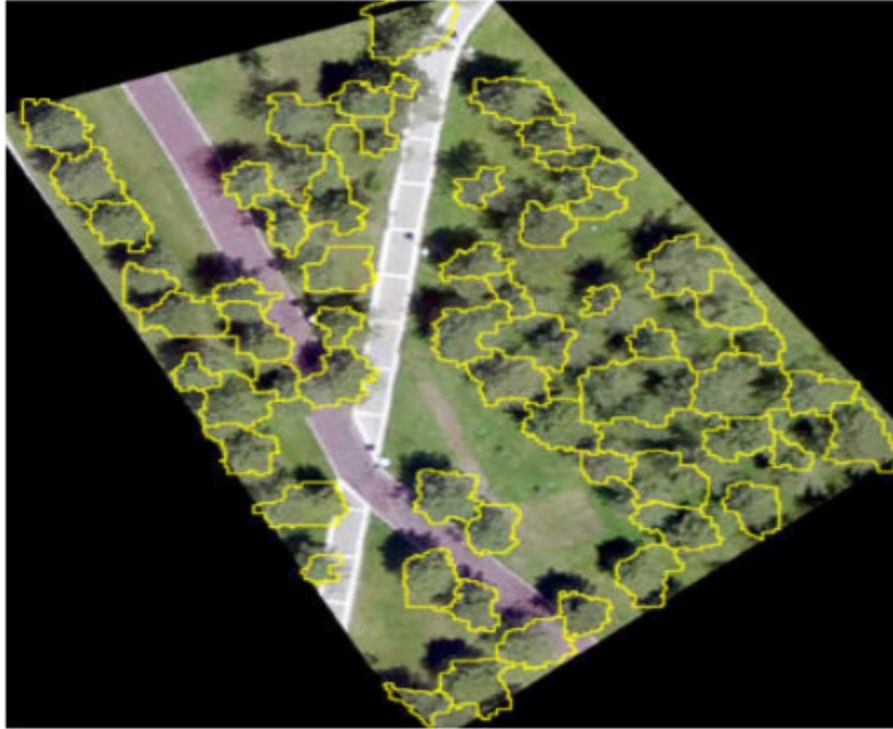
Figure 17.   Illustration of result from performing MR segmentation.
Source: Yilmaz et al. (2017).

RG segmentation uses some treetops (seed points) to start expanding. Seed points are typically obtained from doing local maxima (LM) filtering. The LM filter will extract the highest point within a window operator. The key parameter of this technique is tolerance values to grow regions. If a neighboring pixel is within the tolerance value, that point is defined as the same region as the seed point. Yilmaz et al. (2017) used QGIS software to perform RG segmentation. The study set a window size like 26 x 26, and the threshold parameter as 0.005. Figure 18 shows the resulting illustration after RG segmentation is performed.

Figure 18.   Illustration of result from performing RG segmentation.
Source: Yilmaz et al. (2017).

Even though segmentation techniques are deemed to be effective, over-segmentation problem is a problem users must deal with. Figure 19 shows errors caused by over-segmentation.
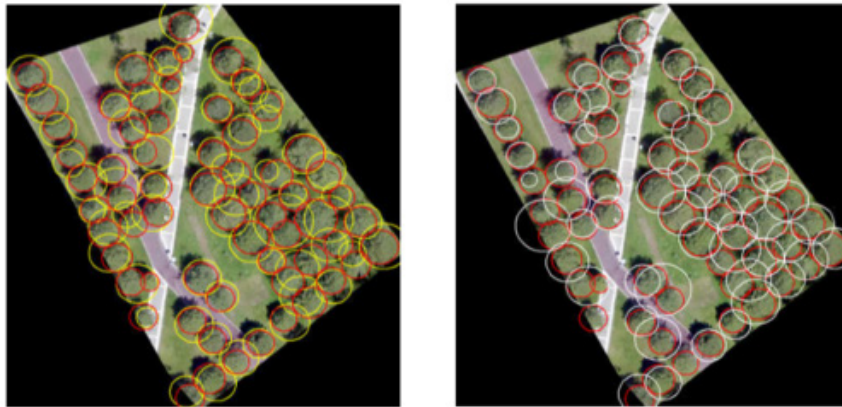
Figure 19.   The circles enclosing the tree crowns (yellow is MR, white is RG, and red is reference circle). Source: Yilmaz et al. (2017).

The result from Figure 19 shows that there were some missing tree crowns. According to Yilmaz et al. (2017), the results from MR and RG techniques are about 90 percent accurate (Yilmaz et al. 2017). From those examples of analysis techniques used in forest studies, it is evident LiDAR plays an integral role.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA SET AND ANALYTICAL METHODOLOGY

The previous chapters provided background on LiDAR and some studies related to LiDAR technology. This chapter provides the instruments, software, and methods we use to collect and enhance the quality of data. The first section discusses the data platforms. The second section discusses data collection and preparation.

A. DATA COLLECTION PLATFORMS

This section discusses the platforms used to collect the data set we used in this thesis: aerial platform and ground observation platform. Furthermore, this section explains how each data set differs.

1. Aerial Platforms

This thesis used two aerial collection systems. The first data set is collected from the Optech Titan multispectral LiDAR. Optech Titan is the same system used by Davis (Davis 2018). The second data set comes from the Airborne Hydrography AB LiDAR (AHAB).

a. Optech Titan Multispectral LiDAR

The Teledyne Optech Titan sensor includes three different wavelength channels. This sensor is well-known as the world's first multispectral airborne LiDAR sensor. The three active beams emit independent wavelengths of 532 nm (green), 1064 nm (NIR), 1550 nm (Intermediate shortwave infrared, SWIR). These three different channels covered for a day or night mapping and different environments, which are included in one sensor (Teledyne Optech n.d.). Figure 20 shows an electromagnetic (EM) spectrum curve with indications for where each channel is located. Figure 21 shows examples of products derived from multi-wavelength LiDAR. The three lasers are tilted at different angles, so they do not illuminate the same points on the ground at a given moment. Table 3 and Figure 22 show the pointing offsets.

To capture waveform data, digitizers can be attached to the Titan LiDAR during operation. For the work completed here, digitizers were attached to the green (532 nm) and NIR (1064 nm) channels. They worked independently, but not always at the same time.
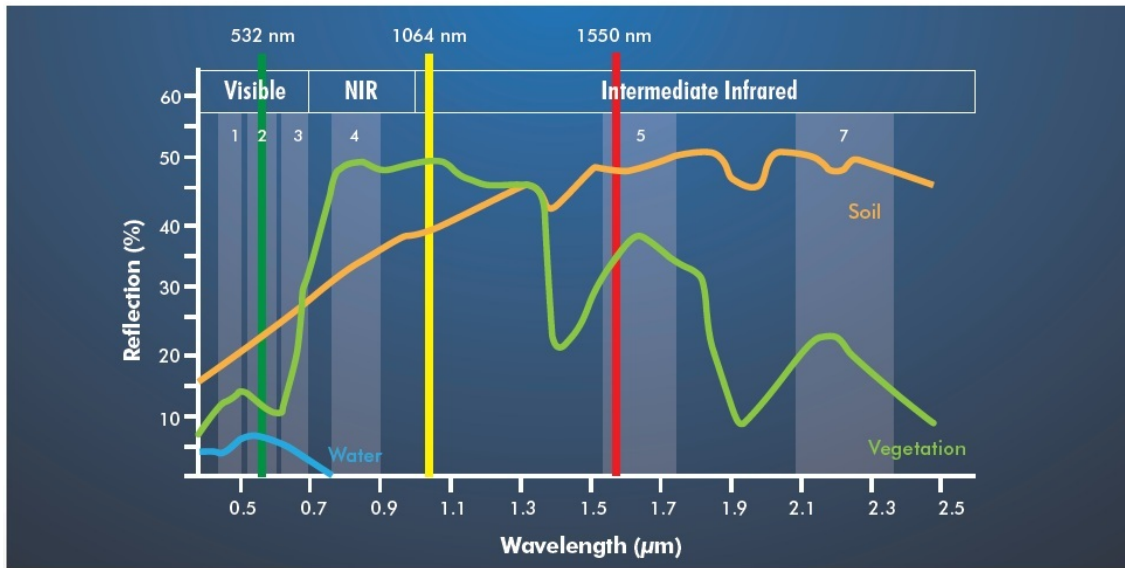
Figure 20. Optech Titan's wavelengths emitted. Source: Teledyne Optech (n.d.).
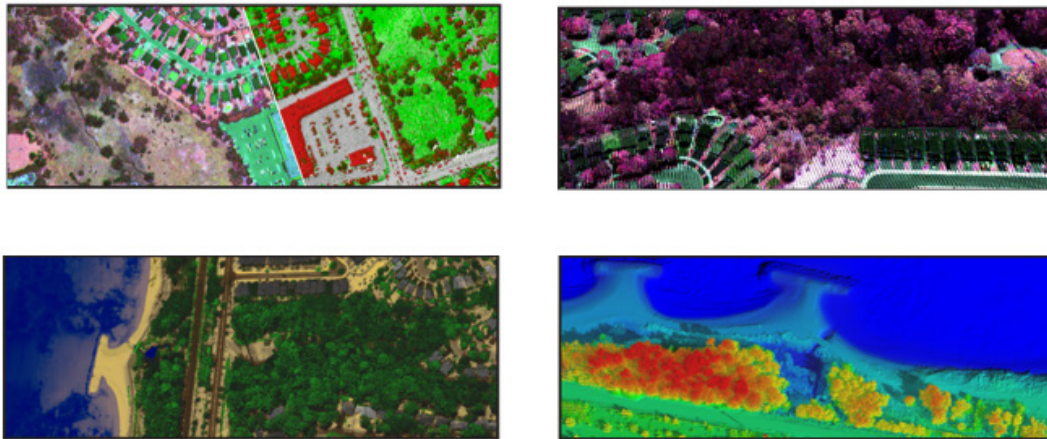
Figure 21. Top-left: 3D land cover classification; Top-right: Vegetation mapping; Bottom left: Shallow-water bathymetry; and Bottom right: Dense topography. Source: Teledyne Optech (n.d.).
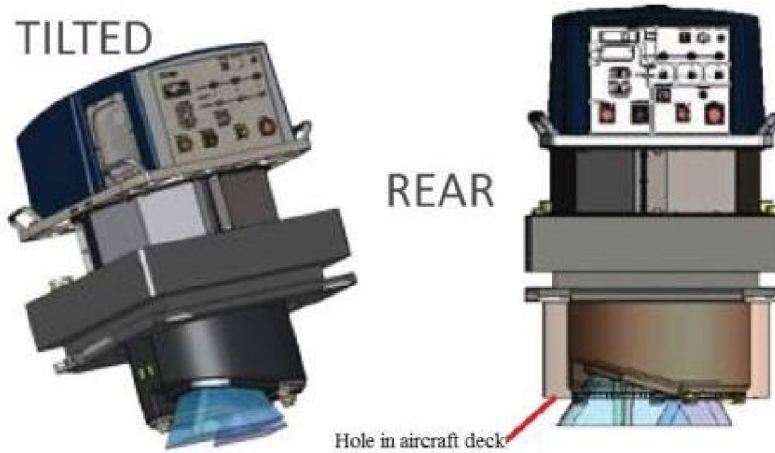
Figure 22.   Drawing of Titan laser. Source: McIver (2017).

Table 3.   Titan attributes. Source: McIver (2017).

	Wavelength	**Offset**	**Beam Divergence**
Channel 1	1550 nm (SWIR)	3.5° Forward	0.35 mrad
Channel 2	1064 nm (NIR)	0° (Nadir)	0.35 mrad
Channel 3	532 nm (Green)	7° Forward	0.7 mrad

b.   *Airborne Hydrography AB LiDAR*

The second aerial platform is the Airborne Hydrography AB (AHAB) Chiroptera II LiDAR system. The AHAB Chiroptera II scanner includes two individual channels, a 532 nm hydrometric scanner operating at 35 kHz, and a 1064 nm topographic scanner operating at up to 500 kHz (Geomatics Data Solutions n.d.). The bathymetric scanner can nominally penetrate to about 1.5 Secchi depths, or approximately 15 m in clear water (Stigermark 2012). "The beam divergence of the topo laser is 0.5 milliradians (mrad) and is 3 mrad for the bathy laser" (Webster et al. 2016, 33).

The AHAB LiDAR uses a Palmer Scanner design, which rotates in an elliptical pattern, as illustrated in Figure 23. The design is used to maintain a reasonably constant

incidence angle between the laser and the surface (water surface). In particular, the design is used to avoid a nadir view and the specular reflection that would result from the surface of the water (Figure 24). The ellipse shape is forward/aft ±14 degrees from nadir, with a transverse width of ±20 degrees. The sensor obtains two views (forward/aft) of most terrestrial targets in a single flight line. During processing, the forward/aft data were separated, as the waveform paths are quite different (Quadros 2013).

Figure 23.   AHAB's scanning pattern. Source: Stigermark (2012).
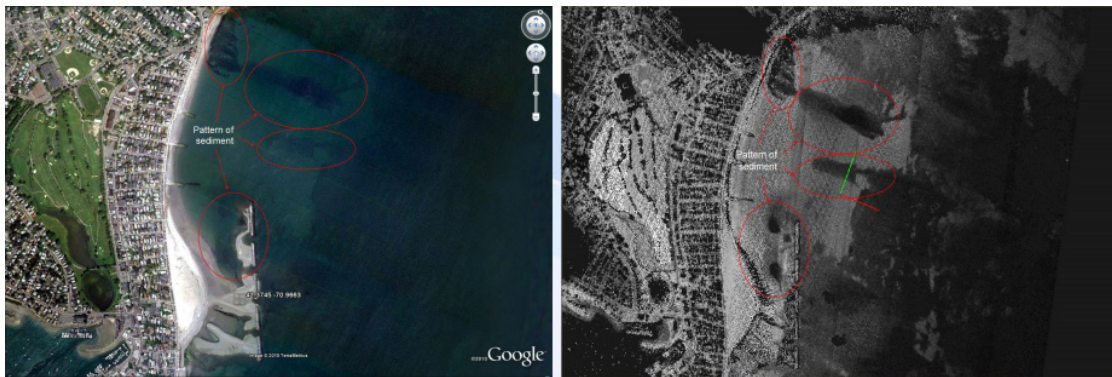
Figure 24.   Seabed reflectance, Orthophoto from Google Earth (left);
LiDAR processed relative reflectance intensity image (right). Source:
Stigermark (2012).

2. Ground Surveying Platform

The ground survey data is used to validate the aerial LiDAR data. The surveying covered the vegetation and terrain. This data set helps to locate and identify specific tree species within the area of interest. The metrics collected were tree species type, height, diameter at breast height (DBH), and a real-time kinematic (RTK) precision GPS location. This thesis mainly focuses on the ground surveyed area (Davis 2018). More detail is provided in the next chapter when we set up a region of interest.

B. DATA COLLECTION AND PREPARATION

Point Lobos State Park, which is located south of the Monterey peninsula in California, served as the study area. Figure 25 illustrates the study area. The park encompasses over 550 acres of natural and marine area which has over 300 species of plants. For this study, the ground survey specifically collected data on the Monterey Cypress, Monterey Pine, and Coast Live Oak trees. These highlighted species are densely distributed throughout the Monterey peninsula and are chosen as a group study for this research (Davis 2018).

The Optech data were collected over an area of interest (October 10, 2016) at Point Lobos, starting from flightline 1 to 30. The data were delivered by the National Center for Airborne Laser Mapping (NCALM) at the University of Houston in both the native Optech waveform data format and in the PulseWaves format. The PulseWaves format was developed by LAStools' Martin Isenburg to be a sensor-independent data format for data interchange. Figure 26 illustrates the contents of a typical data frame and helps illustrate the broader concept of the data analyzed for this thesis.

The AHAB flights were conducted on May 20, 2014. The flights were conducted by the University of Texas at Austin. "The digitization rate in full-waveform mode is 1.8 GHz, equivalent to ~8.33 cm of two-way range resolution. The nominal flight altitude was 400 m above ground level, and the concept of operations resulted in a discrete mode first-return point density of ~10 ppm^2 for both channels" (Leigh and Magruder 2016, 3). The green and NIR channels have a pulse width of 4 ± 1 ns and 2.5 ± 1 ns, respectively. These pulse widths correspond to distances of 1.2 and 0.75 meters in physical distance.

Nominal resolution for point data is then half these respective values. The waveform data were delivered in the native AHAB binary format, with MATLAB code provided to unpack the data (Leigh and Magruder 2016).

Figure 25.   Point Lobos State Park, the shaded area represents the area
of interest. Source: Davis (2018).
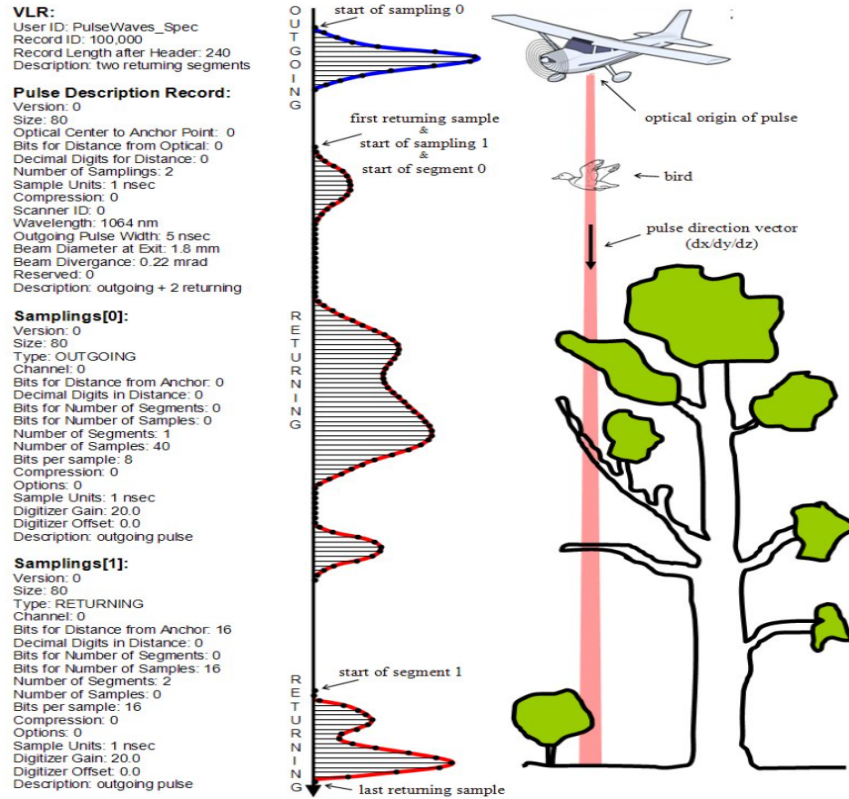
Figure 26.　　PulseWaves data description. Source: https://github.com/ PulseWaves/Specification/blob/master/pulsewaves.jpg; credit Martin Isenburg.

Raw data from each type of platform needs to be converted to an appropriate file format for software analyzing tools. The general process was similar to the two data sets in that data were extracted from the delivered formats, then structured into flat binary files that could more easily be analyzed using Interactive Data Language (IDL). Figure 27 shows one of the intermediate data products developed to view the data, to verify that the data were being read correctly, and to obtain insight into the structure of the waveforms. IDL code was written to grid the data into a standard "hypercube," following the structure generally used for the analysis of hyperspectral imagery (Tyo et al. 2003).
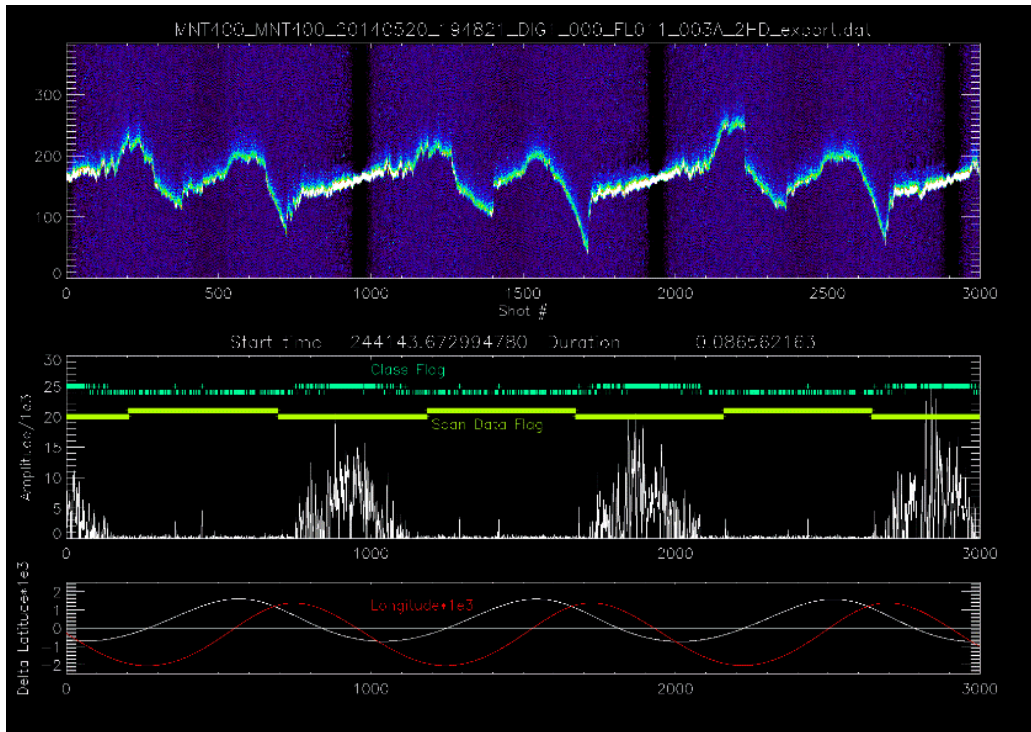
The top panel is a "waterfall plot" or spectrogram showing the waveforms as a function of time (0:3000). Various metadata elements are plotted below. The data shown here capture the measurements for three rotations of the scanner.

Figure 27.   Example of reviewing and converting raw data to a usable format by using IDL.

Standard IDL gridding routines were used with a grid size set to match data density, i.e., one return per a grid cell. Figure 28 illustrates the application of the technique to one short AHAB flightline. It shows the impact of the scanner pattern on the data (the elliptical scan).

FL054_006A_FWF_0pt5mPixels.bil
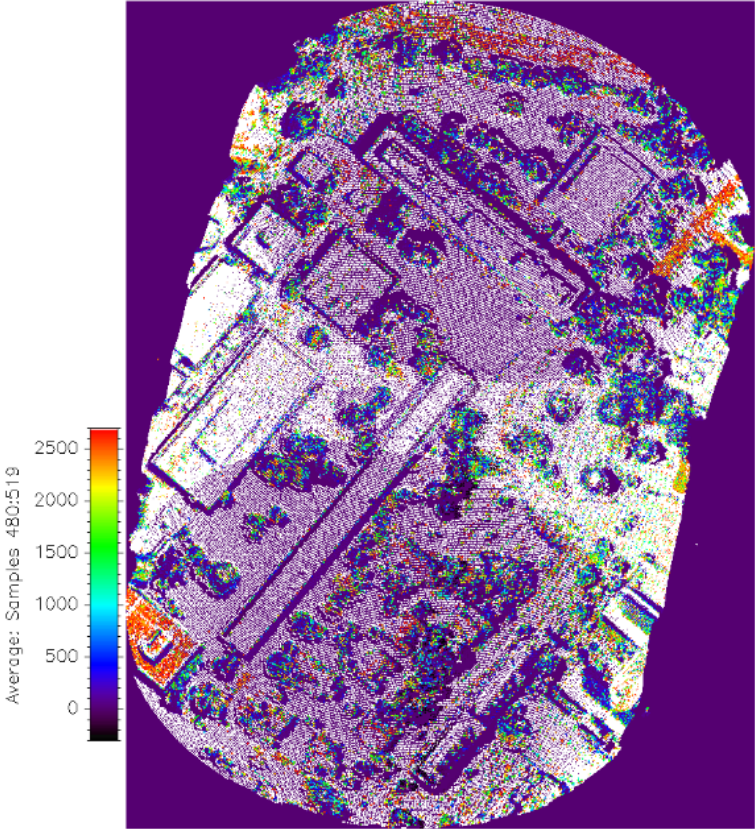
Figure 28.   Example of gridding the AHAB data.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. DATA ANALYSIS AND RESULTS

This chapter discusses the analysis process and evaluates the results. The first section provides an overview of the workflow and analysis processes. The second section describes the classifiers and results of each type of classification algorithm provided by Environment for Visualizing Images (ENVI) software tools. The last section compares and evaluates the results of the classification.

A. OVERVIEW OF WORKFLOW AND ANALYTICAL PROCESSES
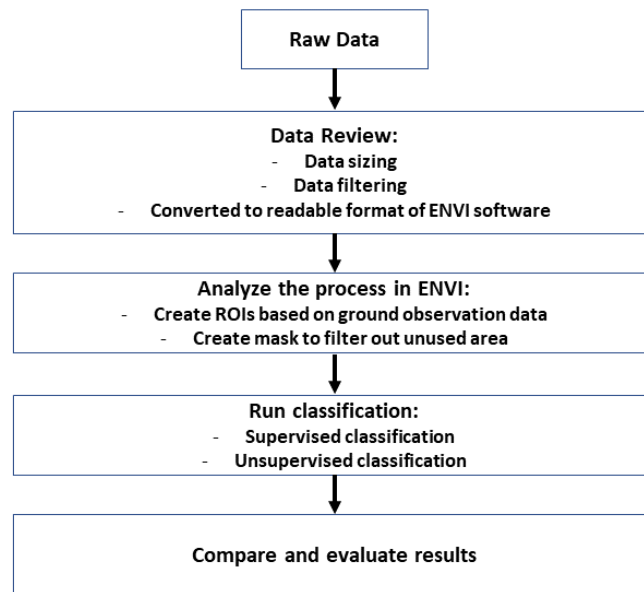
Figure 29.   Workflow diagram of the analysis process.

The workflow diagram (Figure 29) shows an overview of the processes from the initial step through the evaluation step. The first two steps are explained in the previous chapter. The next step, "Analyze the process in ENVI" is explained here. After we got the readable data format for the ENVI file, we initially reviewed the overall area of data samples. Figures 30 and 31 show the estimated covered area of the AHAB and Optech data. These estimations can be done by the function "Chip View to Google Earth" in the

ENVI software, and then using the ruler tool on Google Earth to measure distance and covered area. As shown in Figure 32, there are some dark areas for which we have to create a mask to filter out unused areas, primarily the ocean and coastal rocks.
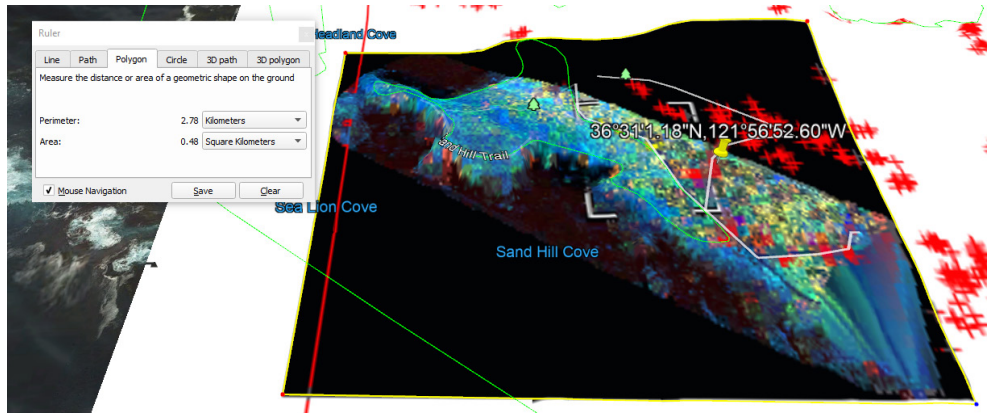
Figure 30.   Estimated covered area of AHAB data. Source: Google Earth (n.d.).
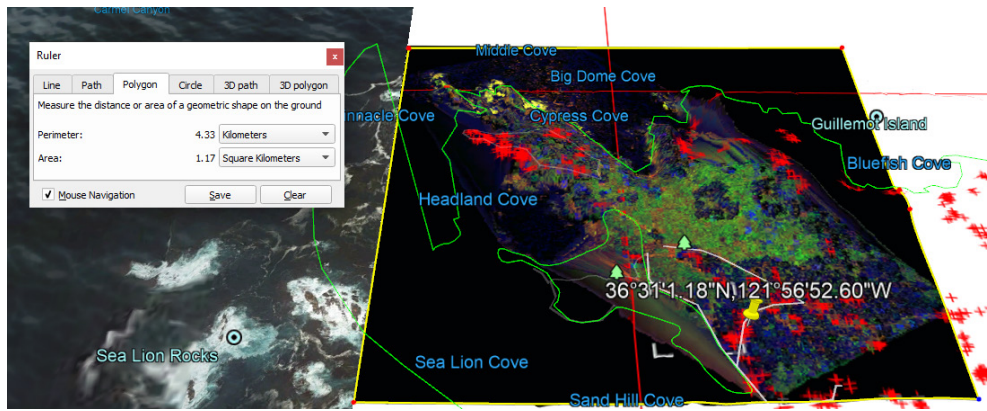
The Optech Titan data flight line 029 is approximately covered 1.17 square kilometers on the actual location.

Figure 31.   The estimated covered area of Optech Titan data. Source: Google Earth (n.d.).
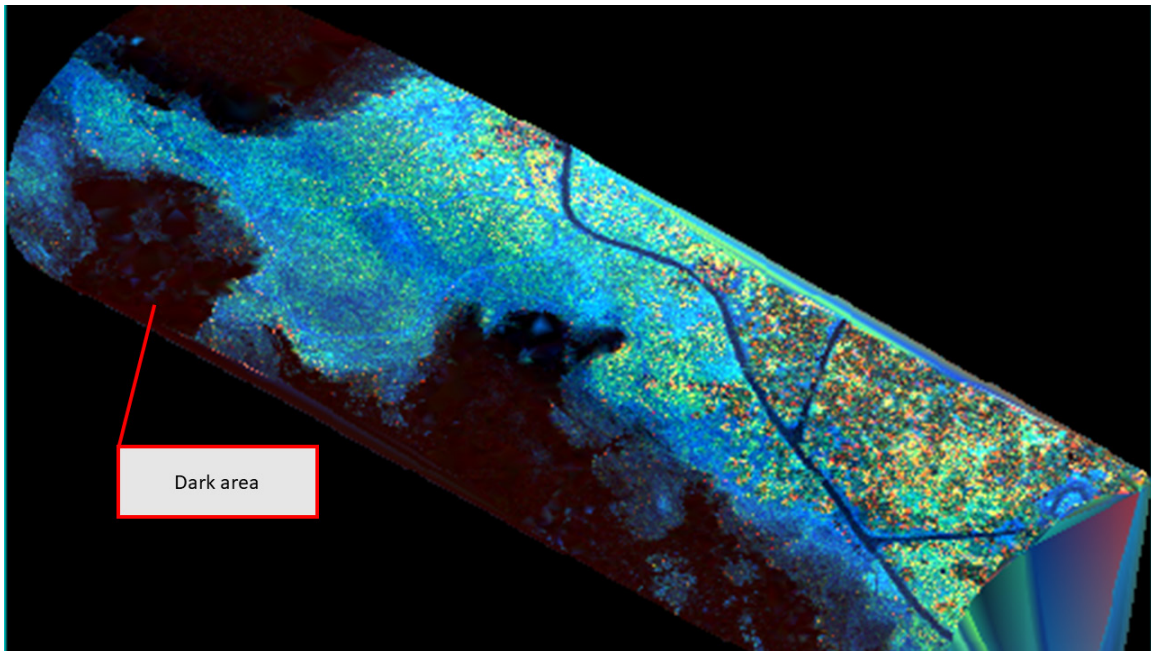
Figure 32.　　　Dark areas shown here are the areas that were not analyzed.

1.　　　Mask Creation

The dark areas in Figure 32 (i.e., water or ocean areas) were unused, so part of the analysis process is to create a mask to avoid problems when we run classifications. To create the mask, we use the masking tool in ENVI. First, we draw a region of interest (ROI) surrounding the usable area. Next, we run the ROI through the "Building mask tool" to define the unused area. Figure 33 shows the usable area. Figure 34 shows the mask for classification purposes.
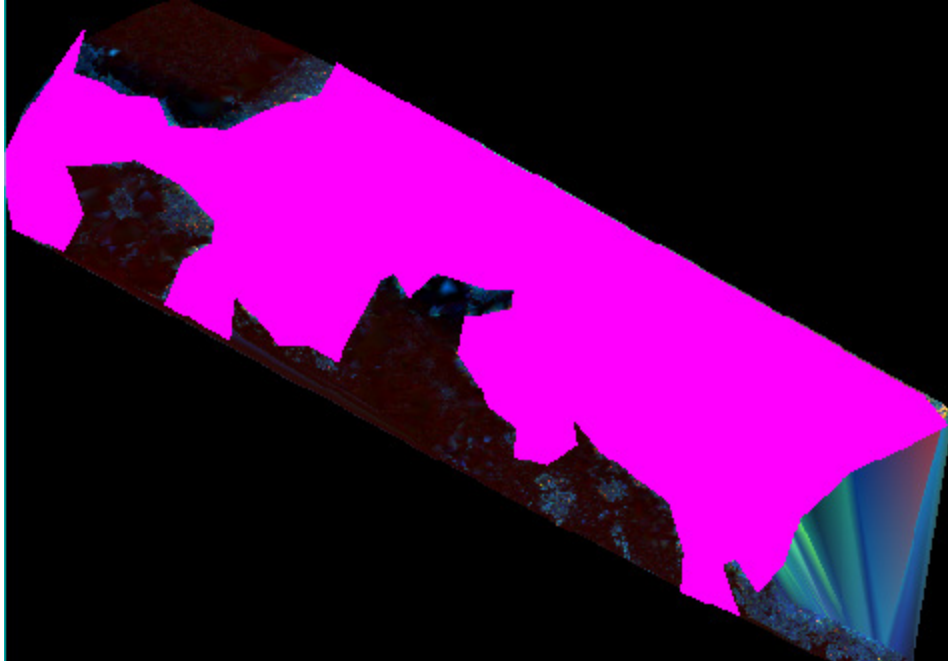
Figure 33.   AHAB usable area.

The white area is the focus area.

Figure 34.   AHAB mask for classification purposes.

2.   Region of Interest

The ROI is a selected sample area of each sample category for classification. Since we are focusing on the tree species, we located each type of tree, pine tree (Figure 35) and cypress tree (Figure 36), based on ground data. The ground data provided geolocation of each tree and the dimension of the tree structure, i.e., height and diameter at breast height. Figure 37 shows how to measure DBH. These features of the tree structure are important because they affect the shape, size, and quality of the return pulse, which depends on the complexity of the trees and how densely they are distributed.

From the Point Lobos Foundation (n.d.): "Monterey pine grows to 80−110 feet; it is symmetrical when young, with an irregular rounded crown at maturity (60−80 years). Its shape can also be distorted if it is crowded by other pines. In favorable conditions, i.e., near the coast, it can live to 100 years, but has a much shorter life (less than 50 years) if planted inland. Needles grow in a bundle of three, 3−6 inches long."

Figure 35.   Monterey pine tree at Point Lobos. Source: Point Lobos Foundation (n.d.).

From the Point Lobos Foundation (n.d.): "Visitors to Point Lobos are familiar with the Monterey cypress (formerly Hesperocyparis macrocarpa), perhaps the terrestrial symbol of the reserve. We see the cypress all around the Monterey peninsula and forget how unique it is. There are only two native stands: one at Allan Memorial Grove in Point Lobos and the other at Crocker Grove in Pebble Beach. California lists it as a Category 1 rare and endangered species, but since it is widely planted as a landscape tree, it is not on the federal list of endangered species – the federal list does not distinguish between natural populations and planted specimens."

Figure 36.   Monterey cypress tree at Point Lobos. Source: Point Lobos Foundation (n.d.).
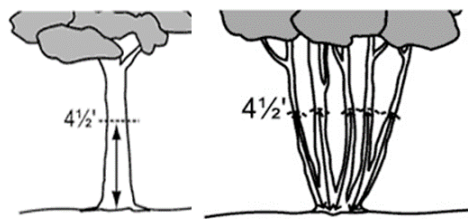
DBH is measured at 4.5 feet above the ground.

Figure 37.   How to measure DBH. Source: City of Portland, OR (n.d.).

After we collected the ground data, we generated the shapefile based on the data we collected. Then, we used this shapefile to plot the ROI. Figure 38 shows the plot of the shapefile and ROI of the Optech data. Figure 39 illustrates the plot of the shapefile and ROI of the AHAB data. After we obtained the ROI samples, we could quickly compute statistics of the samples. Figure 40 shows the mean value of each sample for Optech data, where black circles are roads/trails, green circles are low vegetation, blue circles are pine tree, and red circles are cypress trees. Figure 41 shows the same plot as Optech data, but based on the shapefile, there is no cypress sample over the data area.
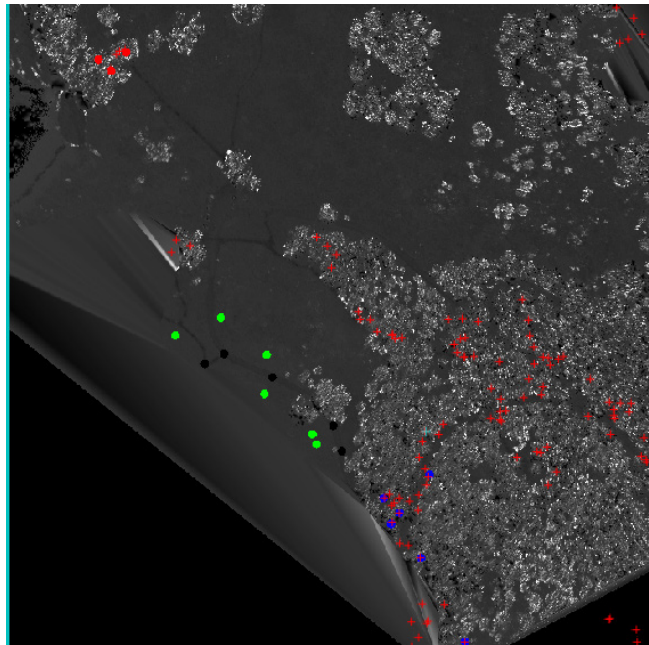
Grayscale image of Optech data is plotted by ground data as red plus-sign, roads/trails as black dots, low vegetation as green dots, pine trees as blue dots, and cypress trees as red dots.

Figure 38.   ROIs and shapefile over the Optech data area.
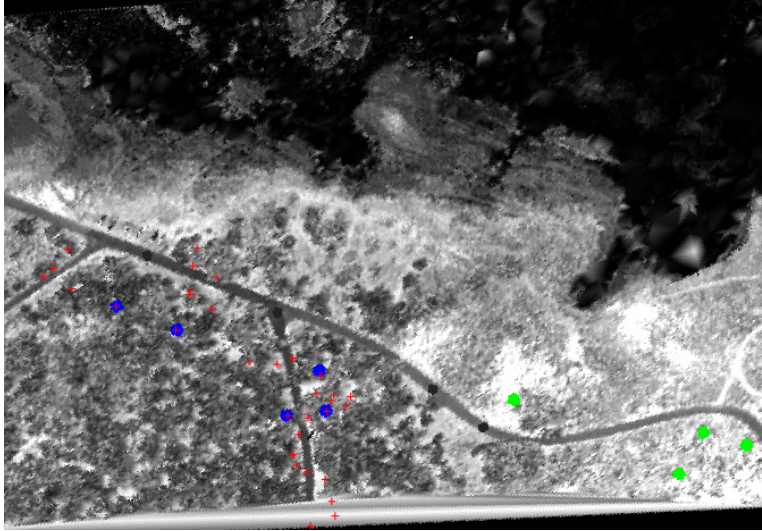
Grayscale image of AHAB data is plotted by ground data as red plus-sign, roads/trails as black dots, low vegetation as green dots, and pine trees as blue dots. There are no cypress trees.

Figure 39.   ROIs and shapefile over the AHAB data area.
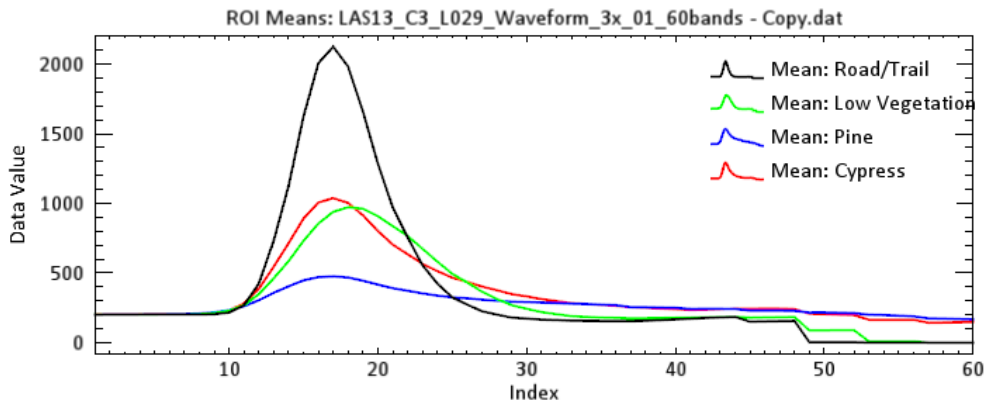
As shown in the figure, the mean for low vegetation and the mean for cypress trees are very close to each other. The x-axis indicates the band number and y-axis represents return number for each band.

Figure 40.   ROIs plot of Optech data as a function of waveform sample number or index.
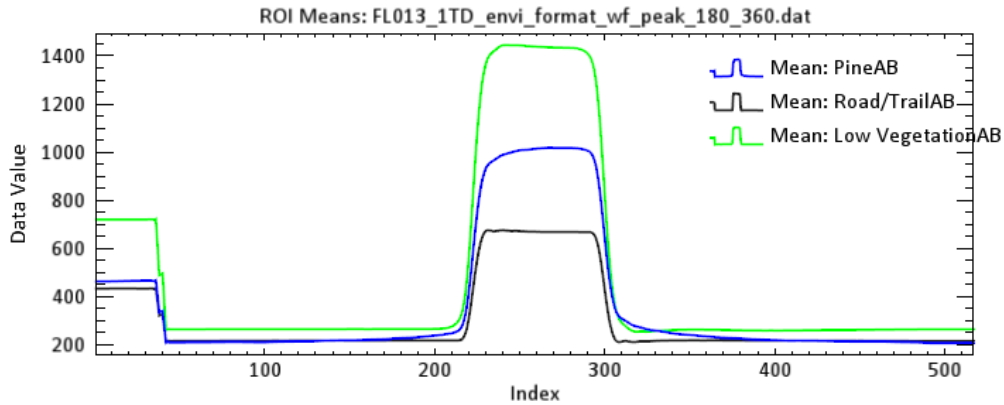
The figure shows the difference among three (3) samples. These curves are quite close to each other, but we did a smooth plot to clearly distinguish the differences.

Figure 41.   ROIs plot of AHAB data as a function of waveform sample.

As mentioned earlier, ROI samples are very important for classification. Therefore, the selection of sample data is very important. Tables 4 and 5 show selected ROIs that considered the features of the tree, i.e., height and DBH. The other ROI samples, low vegetation and roads/trails, were set up by visual inspection.

Table 4.   Summary of Optech ROIs.

Tree type	Tree height (m.)	Tree DBH (m.)	Assigned color
Pine_01	10	n/a	Blue
Pine_02	16	n/a	Blue
Pine_03	12.4	n/a	Blue
Pine_04	3	n/a	Blue
Pine_05	20	n/a	Blue
Pine_06	17	n/a	Blue
Pine_07	20	n/a	Blue
Cypress_01	7.8	1.8	Red
Cypress_02	5.6	0.81	Red
Cypress_03	n/a	1.51	Red
Cypress_04	11.6	1.11	Red

Tree type	Tree height (m.)	Tree DBH (m.)	Assigned color
Cypress_05	12	n/a	Red
Cypress_06	11	n/a	Red

Table 5.   Summary of AHAB ROIs.

Tree type	Tree height (m.)	Tree DBH (m.)	Assigned color
PineAB_01	19	n/a	Blue
PineAB _02	11.6	n/a	Blue
PineAB _03	9.6	n/a	Blue
PineAB _04	7.2	n/a	Blue
PineAB _05	12.4	n/a	Blue
PineAB _06	10.8	n/a	Blue
PineAB _07	6.8	n/a	Blue
PineAB _08	10	n/a	Blue
PineAB _09	17.4	n/a	Blue
PineAB _10	16	n/a	Blue
PineAB _11	17	n/a	Blue
PineAB _12	15.6	n/a	Blue

B.   WAVEFORM CLASSIFICATION

Since we have two data systems, we can obtain different results when we apply classifiers. The difference in sensor systems, date-time collection, and other variables make the return pulse dissimilar (Figures 42 and Figure 43). The tools to classify the samples we created are provided in ENVI software. There are two main types of classification tool, supervised and unsupervised classification.
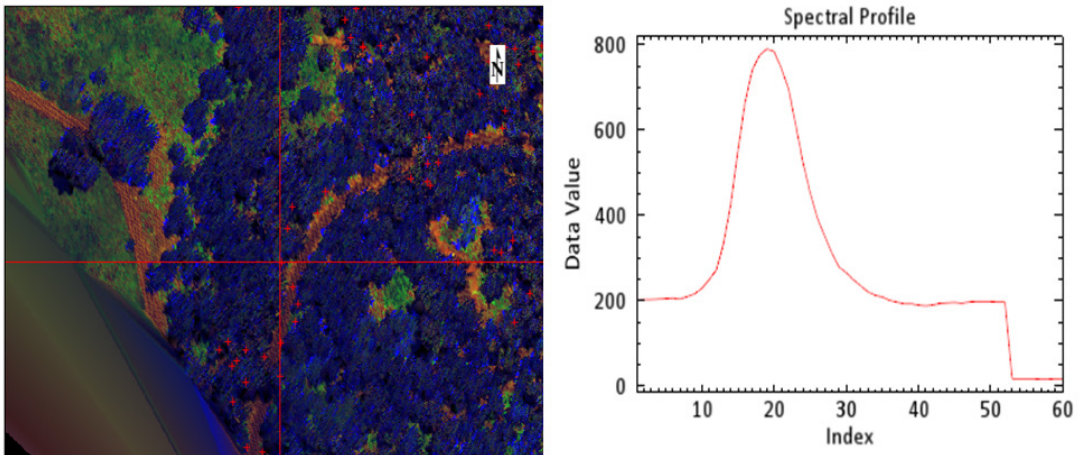
Figure 42.   A pine tree (left) and its spectral profile (right) in the
Optech data.
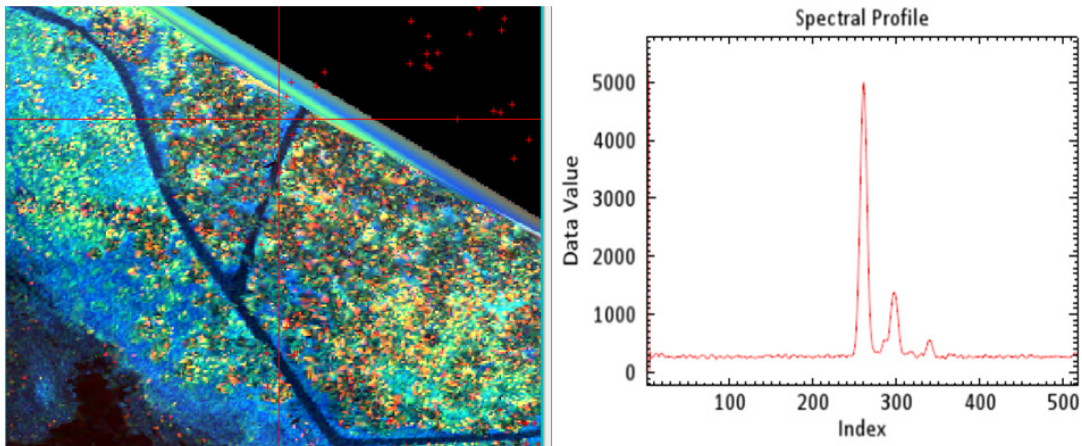
Figure 43.   A pine tree (left) and its spectral profile (right) in the
AHAB data.

1.   Unsupervised Classification Tool

There are two unsupervised tools provided in ENVI's toolbox. One is "IsoData Classification." The other, which we used in this research, is "K-means Classification." According to the explanation of K-means method provided with the ENVI software,

K-Means unsupervised classification calculates initial class means evenly distributed in the data space then iteratively clusters the pixels into the nearest class using a minimum distance technique. Each iteration recalculates class means and reclassifies pixels with respect to the new means. All pixels are classified to the nearest class unless a standard deviation or distance threshold is specified, in which case some pixels may be unclassified if they do not meet the selected criteria. This process continues until the number of pixels in each class changes by less than the selected pixel change threshold or the maximum number of iterations is reached (Harris Geospatial n.d.).

The K-means method can distinguish the classes of data, but it cannot indicate and distinguish the species of the tree. Figure 44 shows the product of the K-means method applied to the Optech data.
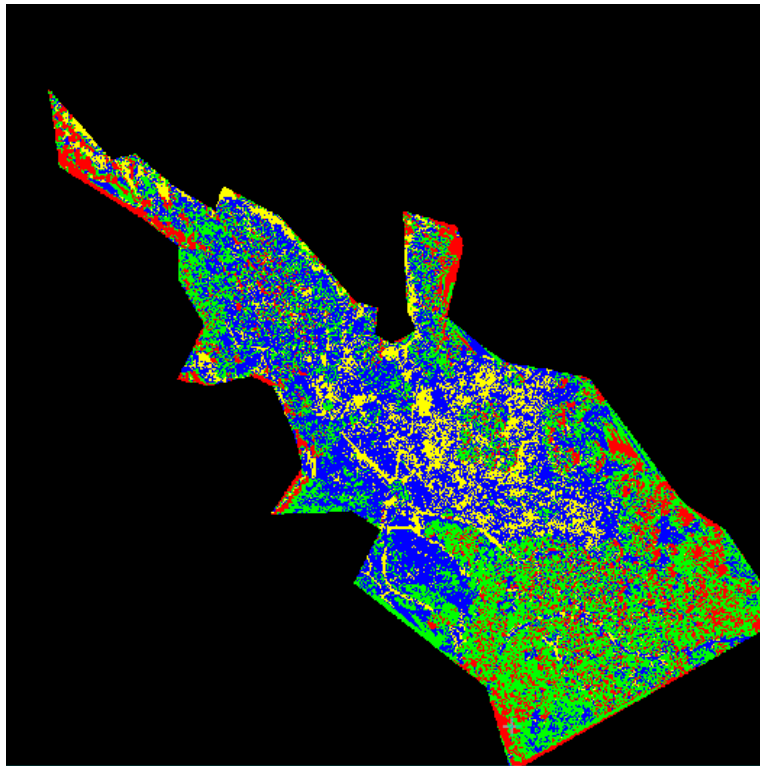
Figure 44.   K-means product of Optech data.

2.   Supervised Classification Tool

The supervised classification tools are the primary tool that we used to evaluate our results. Specifically, the two supervised methods we used in this research were "Spectral Angle Mapper (SAM) Classification" and "Support Vector Machine (SVM) Classification."

a.   Spectral Angle Mapper

SAM is the first method we applied for classification. SAM is used widely in optical imagery analysis (Davis 2018). The key for SAM is to use the n-D angle by calculating the spectral angle between each band. Then, the classifier groups the best-matched pixels to the reference spectral signature selected from the ROI regions (Harris Geospatial n.d.). Figure 45 shows the result of SAM applied to the Optech data. The first attempt of the SAM method set a radian value to 0.5, which ran through 60 bands of Optech data and 517 bands for AHAB data. The result shows that at this radian value, SAM can separate the classes through the entirety of the data, and there are only a few unclassified points. Table 6 shows the percentage of points that were classified and assigned to each class. Figure 46 and Table 7 illustrate the result of SAM applied to the AHAB data, which follows the same processes as those for the Optech data. For the AHAB data, however, there are some errors between the low vegetation class and the roads/trails class. The reason for these errors is when we created the ROI for the low vegetation class, it covered both low vegetation and roads/trails (soil and asphalt in this case).
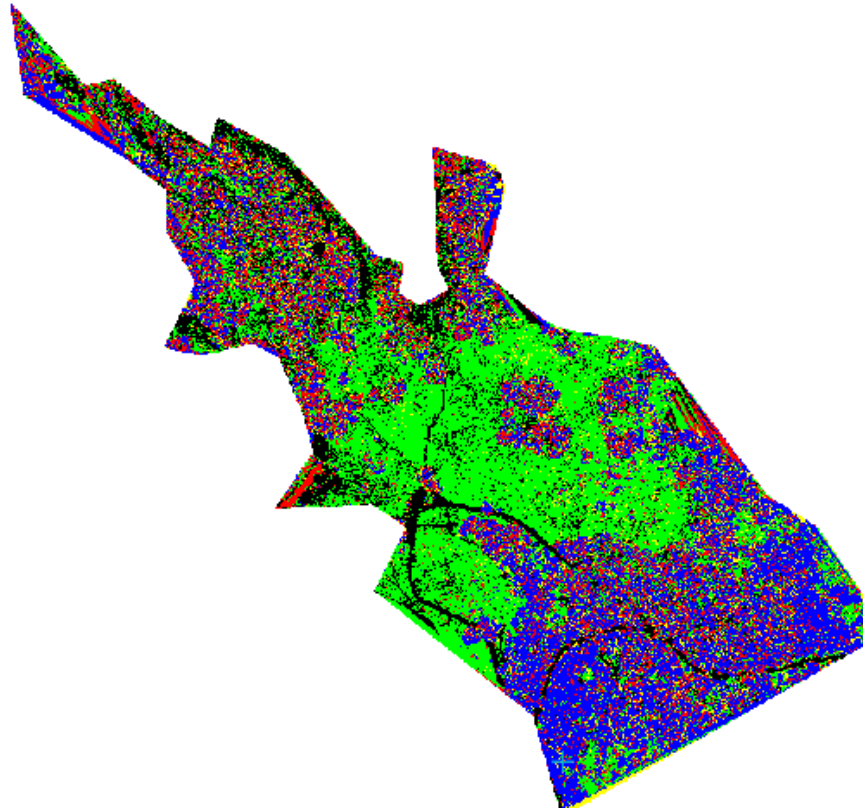
The result shows that roads/trails, low vegetation, and trees can be clearly separated.

Figure 45.　　　SAM classification at 0.5 radian for Optech data.

Table 6.　　　SAM classification summary report for Optech data.

Class summary	Percentage of counted pixels on the usable area
Unclassified	5.95%
Roads/Trails	16.20%
Low vegetation	34.33%
Pine	27.29%
Cypress	16.24%
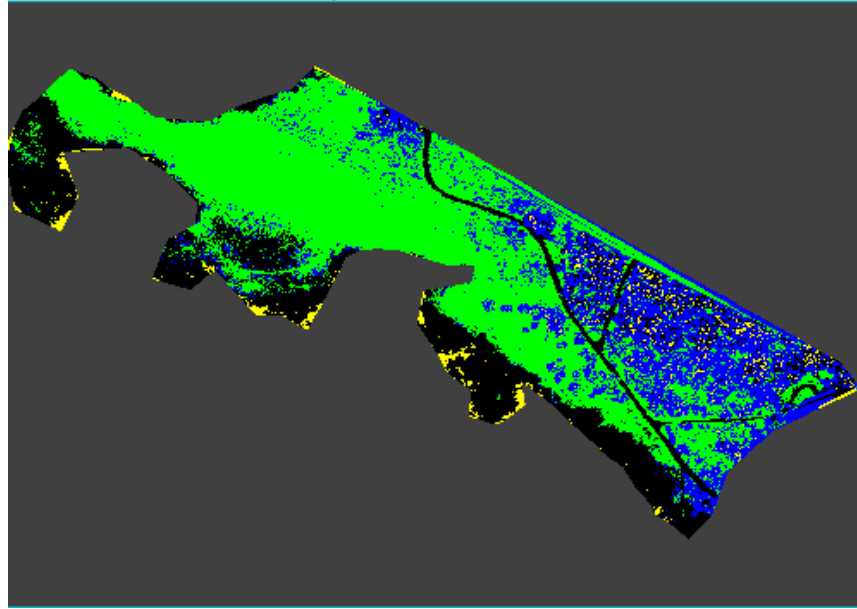
Figure 46.   SAM classification at 0.5 radian for AHAB data.

Table 7.   SAM classification summary report for AHAB data.

Class summary	Percentage of counted pixels on the usable area
Unclassified	2.89%
Roads/Trails	25.57%
Low vegetation	50.76%
Pine	20.78%

After that, we evaluated the accuracy of classification by using the ground data location as the reference point. Figure 47 shows classification results at the expected location. Table 8 represents the evaluation of the accuracy of Optech data. Table 9 represents the evaluation of the accuracy of AHAB data. The results show that some errors occurred. That is about 40 percent accurate.
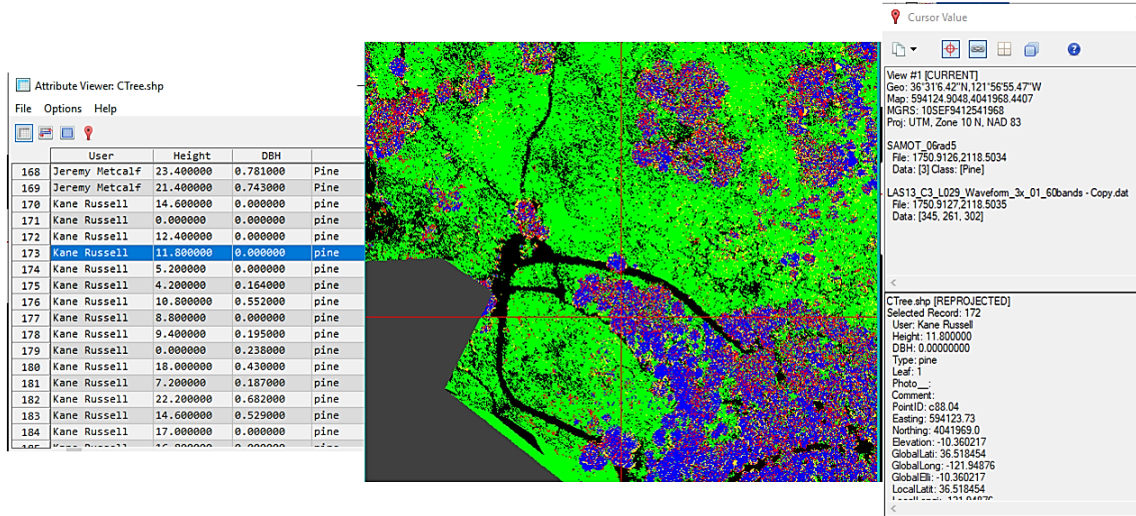
Figure 47.   Evaluation of the accuracy of SAM method at 0.5
maximum radian.

Table 8.   Evaluation of accuracy of SAM method (Optech data).

Class	Height/DBH (m)	Result	Comment
Pine01	12.4/n.a.	correct	
Pine02	11.0/n.a.	correct	
Pine03	7.2/0.18	correct	
Pine04	24.6/0.63	error	Low vegetation
Pine05	10/n.a.	error	Cypress
Cypress01	11/0.78	correct	
Cypress02	6.0/0.16	error	Pine
Cypress03	10.0/0.66	correct	
Cypress04	15.8/0.52	error	
Cypress05	11.8/0.49	correct	Low vegetation

Table 9.   Evaluation of SAM classification accuracy (AHAB data).

Class	Height/DBH (m)	Result	Comment
Cypress01	9.6/n.a.	correct	
Cypress02	9.8/n.a.	error	Unclassified
Cypress03	10.0/n.a.	error	Low vegetation
Cypress04	12.4/n.a.	correct	
Cypress05	16.0/n.a.	correct	

In the next approach, we try to adjust the maximum radian value to see which maximum radian value can get rid of errors. After we applied the SAM method with different values of maximum radian angle, we concluded that for Optech data, the smallest radian angle we can apply is 0.3. This value can maintain the same result as the radian at a threshold of 0.5. As we increased the radian angle, errors were about the same, but there was a change in the number of tree species counted. For AHAB, the best radian angle threshold we can apply is the range between 0.5−0.7.

b.   *Support Vector Machine*

The second classification approach is the SVM. The SVM method is a widely used classification technique for various applications (Bell 2015). The key to SVM is to separate the classes we created by the "data point closest" to the margin called the "support vector," which is based on the distance to the margin, "often called [the] optimal hyperplane" (Harris Geospatial n.d.). Following the results from the SAM analysis, we adjusted the location and covered area for better decision making of the classifier tool. We started SVM at 0.7 as a threshold to determine the class samples. For AHAB data, the results showed that SVM has the capability to distinguish the class samples (Figure 48b). Especially for the difference between low vegetation and roads/trails, SVM can locate even the small trail, which was surrounded by low vegetation. On the other hand, tree species still have some errors. After that, we varied the threshold to 65 percent and 85 percent. At 65 percent of the threshold, the capability to distinguish the class appears better when we inspect visually (i.e., the unclassified points decreased) (Figure 48a). For

85 percent of the threshold, the unclassified points are increased, and it is difficult to locate each class by visual inspection (Figure 48c). Table 10 shows the results of evaluating the accuracy of AHAB data classified by SVM. For Optech data at a 70 percent threshold, SVM can distinguish low vegetation and roads/trails from the tree species (Figure 49b). Even though there are some unclassified points, they do not disturb the outline structure of the main considered classes. At a 65 percent threshold, the separation of each class is more apparent than the previous attempt (Figure 49a). However, when we determined the accuracy based on ground data, we determined that the error in this attempt is higher than the previous attempt (70% threshold) (Table 11). In the last experiment, at 85 percent of the threshold, it increased in the level of unclassified points, but the class type of low vegetation and road/trail still dominate (Figure 49c).

(a)   (b)   (c)
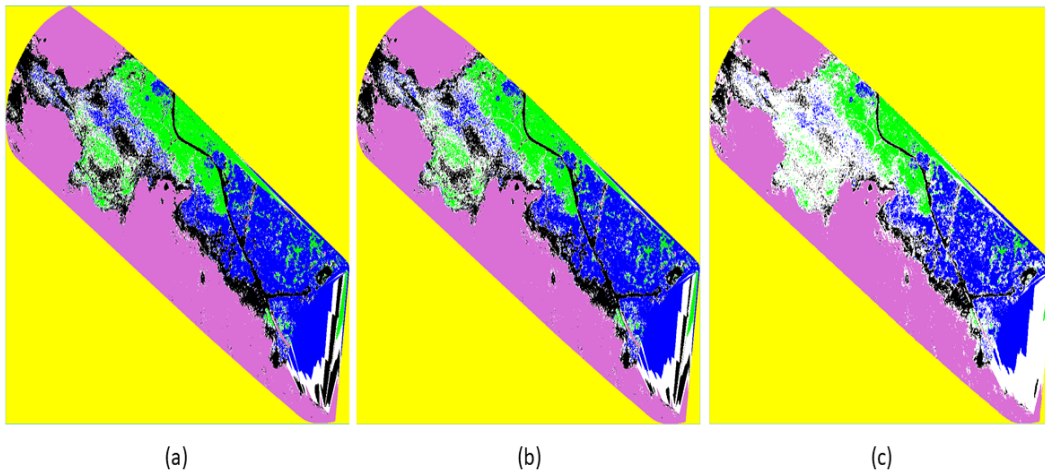
a) At threshold 60%; b) at threshold 70%; c) at threshold 85%. Blue is pine class; Black is road/trail; Green is low vegetation class, and other colors are ignored.

Figure 48.   The SVM results of AHAB data at various threshold values.

Table 10.   Evaluation of SVM classification accuracy at various threshold values (AHAB data).

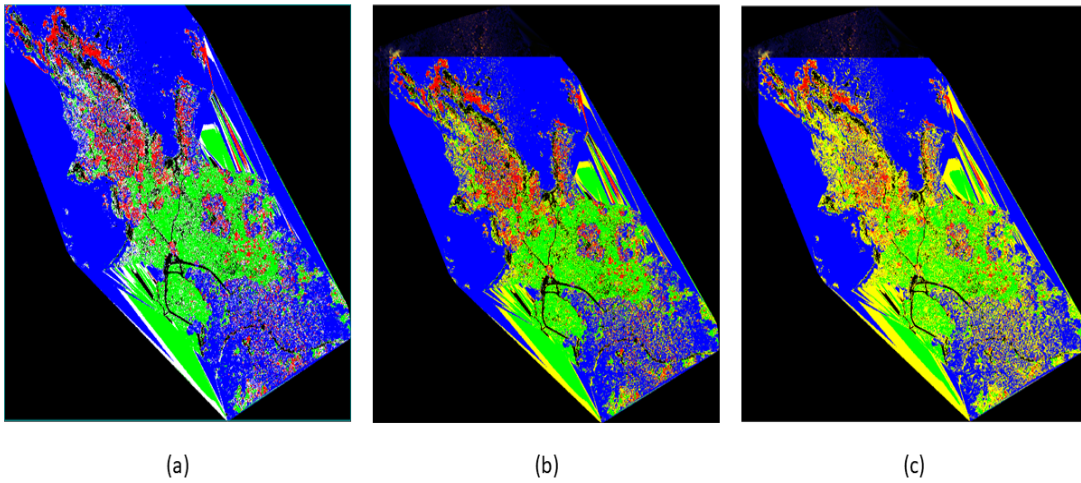|	At threshold (65%)	At threshold (70%)	At threshold (85%)
Unclassified	22.94%	29.02%	49.55%
Pine	33.85%	32.00%	25.98%
Roads/trails	24.91%	21.94%	11.69%
Low vegetation	18.29%	17.03%	12.77%
Verified pine tree with ground data (%)	100%	100%	100%

(a)   (b)   (c)

a) At threshold 60%; b) at threshold 70%; c) at threshold 85%. Blue is pine class; Black is roads/trails; Green is low vegetation class; Red is cypress; and large blue/black colors outside the area of interest are ignored.

Figure 49.   The SVM results of Optech data at various threshold values.

Table 11. The evaluation of SVM classification accuracy at various threshold values (Optech data).

	At threshold (65%)	At threshold (70%)	At threshold (85%)
Unclassified	23.46%	26.08%	45.52%
Pine	22.4%	28.03%	17.06%
Roads/trails	13.23%	12.14%	9.61%
Low vegetation	29.6%	26.94%	20.00%
Cypress	14.17%	12.36%	7.81
Verified Pine and Cypress tree with ground data (%)	60%	70%	50%
Comment	One error on Pine, three errors on Cypress	One error on Pine, two errors on Cypress	One error on Pine, four errors on Cypress

c.   *Evaluate Overlap Area*

Next we evaluated the overlap area of the Optech and AHAB data sets. First, we created a new ROI, which covered both the Optech data and the AHAB data. We expected to evaluate the accuracy of the classification of tree species (in this case, pine tree), and the capability to distinguish the low vegetation area from its surroundings. Consequently, we set up new ROIs to cover the known area of the pine trees. This new ROI was on the overlap area of those two data sets. After we created the new ROI, we ran the classification via the SAM tool. In SAM processing, we set the maximum radian angle to 0.5. Figure 50 illustrates the results of classification from the overlap area, using the SAM method. The accuracy of classification over the two data systems was evaluated by ground data samples. The Optech result proved 100 percent correct, but for the AHAB data, the accuracy of classification was only 40 percent.
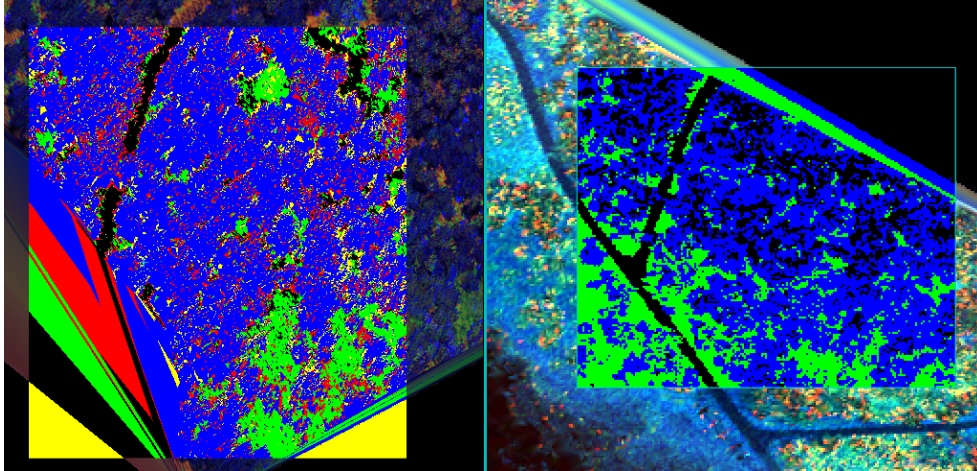
On the left is Optech data; on the right is AHAB. Blue color indicates pine tree, green is low vegetation, and black is roads/trails. Ignore the colored bars on the bottom.

Figure 50.   Evaluation of known tree species in the overlap area.

In another experiment, we created a new ROI over the known low vegetation area. After we created the ROI, we ran it through the SAM tool for classification. The result showed that on the Optech data it distinguished the low vegetation area from its surroundings well. In addition, the results of Optech data indicated that it can also locate the road and ground (soil) effectively. On the other side, the AHAB data result had an error located in the ground area known as pine tree (Figure 51).
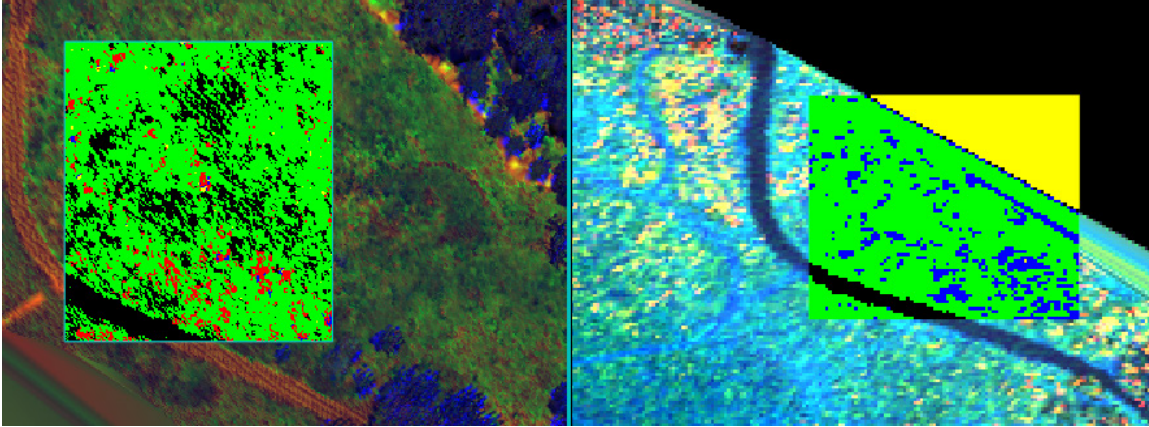
On the left is Optech data; on the right is AHAB data. Blue color indicates pine tree; green is low vegetation; red is cypress tree, and black is roads/trails. Ignore the yellow triangular area at the top.

Figure 51.   Evaluation of known low vegetation areas in the overlap area.

C.   RESULTS

From the beginning of the analytical processes until now, we obtained many results from various types of classification. The first result was from the unsupervised classification method. The K-means tool only distinguished the classes based on the number of classes we selected. It did not specify the type of each class. In the case of supervised classification, SAM achieved an accuracy of only 60 percent for both data sets; the percentage of unclassified points was lower than that for the SVM method. SVM increased the accuracy of classification (up to 100% on AHAB data). The percentage of unclassified points for the SVM method is very high, and it seems to increase the distribution in a specific area. As in Figure 48c, the white noise is distributed mostly at the top-left of the picture. In the case of considering only the low vegetation class, both methods can clearly distinguish this class from its surroundings. In terms of the overlap area, however, Optech data gave more satisfactory results for the AHAB data. One possible error that we could locate during the analyses is the spectral profile of each expected class. This spectral profile shows the characteristic of return pulses. Figure 52 depicts the spectral profile of all expected classes. Low vegetation (Figure 52a) is almost the same as the cypress tree (Figure 52d). This might explain why, when we classified the

overlap area, some red dots (cypress trees) appeared on the expected area. Also, similar to the Optech data, for the AHAB data set the area that should be assigned as the roads/ trails class (soil) was assigned as the pine tree class (Figure 53).
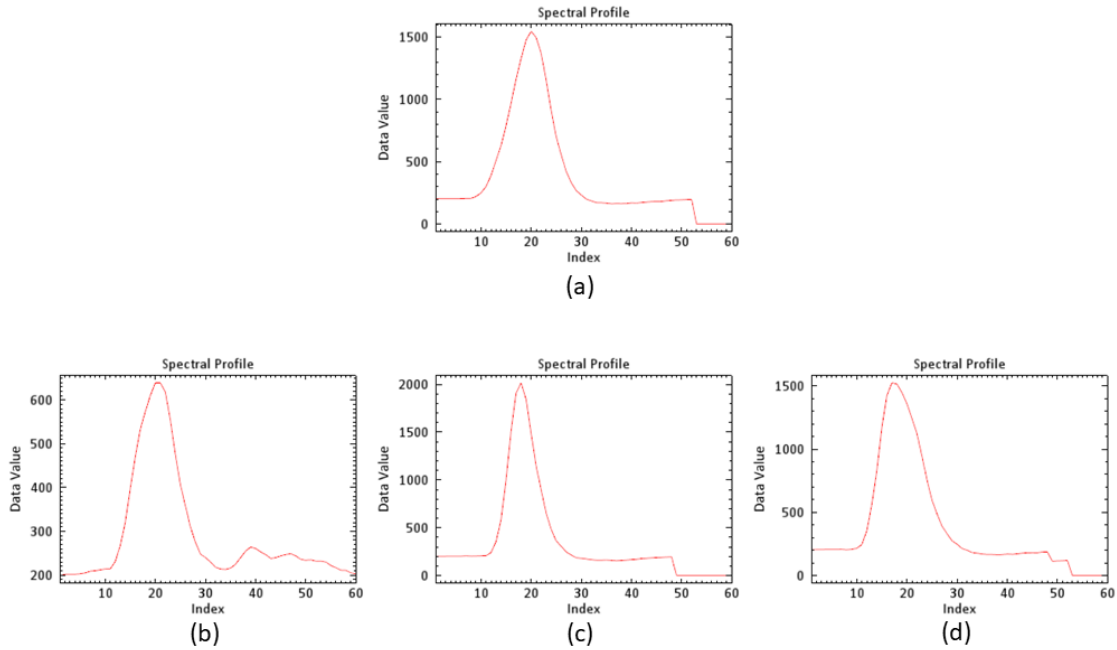
a) Low vegetation; b) pine tree; c) roads/trails, d) cypress tree.

Figure 52.   Spectral profile of Optech data.
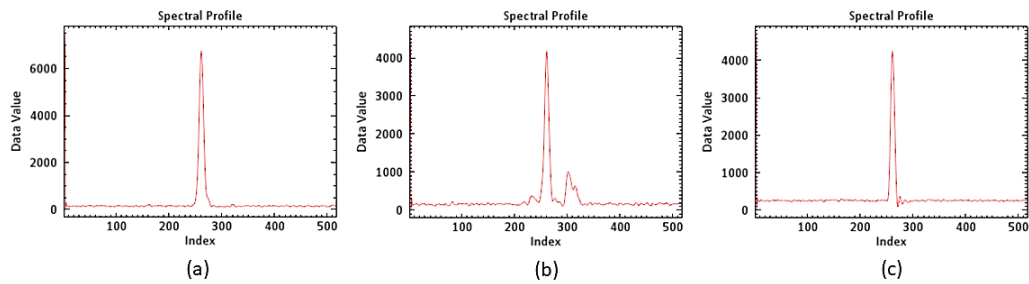
a) Low vegetation; b) pine tree; c) roads/trails.

Figure 53.   Spectral profile of AHAB data.

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION

This chapter summarizes the results of this study to answer the objectives stated in Chapter I, and notes the problems that occurred during the research. Furthermore, the second section offers some specific recommendations for future work.

A. SUMMARY

This thesis has focused on evaluating the Terrain Categorization (TERCAT) capability of LiDAR waveform data. The data we used was obtained from two different LiDAR systems and provided us with various case studies. We classified and evaluated the accuracy of LiDAR waveform data from both systems using ENVI software. The results from the analyses led us to conclude that waveform data can be used to study and classify low vegetation areas. In fact, low vegetation areas were found to be unique when compared to other classes. This study also found that waveform data provides the capability to locate various types of tree species, even in areas in which the sample types are densely distributed. Furthermore, when using waveform data, the SVM classification method was found to produce more accurate results than the SAM classification method. Nevertheless, the number of unclassified points produced from the SVM method was higher. The study also uncovered another problem when the waveform data within the overlap area of the two data systems provided a case study based on the wrong classification decision. We surmise that this problem might occur when classes have a narrow gap in data values and share similar pulse shape.

Even though we were able to complete the analysis and evaluate the results of this study, it was not without problems. First, we had to deal with the inherent limitations of each data system used in our research. The differences in the data collection systems, date, time, altitude, and scanning angle, can introduce bias in decisions when comparing the two data systems (Lewis and Hancock 2007). Therefore, we cannot confidently evaluate in detail the data collection techniques. We can only roughly review the difference between the two data (Optech and AHAB) systems in the overlap area. Moreover, ground observation data does not cover the entire area of interest. We

59

acknowledge that this limits the sample points selected for evaluating the accuracy of LiDAR waveform data from both systems, using ENVI software.

B.   FUTURE WORK

The recommendation for future researchers is to focus on classifying the tree species and low vegetation but from a different point of view. One area that this author would recommend is to consider the impact of ROIs on the decision making of the classifier tools provided in ENVI software. Specifically, future researchers should start transforming the actual ground area according to the actual size and location of the ROI. This should give more reliable results and fewer errors.

LIST OF REFERENCES

ArcGIS Desktop. n.d. "What Is Lidar Data?—Help ArcGIS for Desktop." ArcGIS Desktop. Accessed March 31, 2019. http://desktop.arcgis.com/en/arcmap/10.3/manage-data/las-dataset/what-is-lidar-data-.htm.

Baker Mobile LiDAR. 2016. "Esri International User Conference." June 23, 2016. http://www.mobilelidar.com/2016/06/esri-international-user-conference.html.

BCC Research (blog). 2018. "Brief History of LiDAR, Its Evolution and Market Definition." March 20, 2018. http://blog.bccresearch.com/brief-history-of-lidar-evolution-and-market-definition.

Bell, Jason. 2015. "Support Vector Machines." In *Machine Learning*, 139−160. Indianapolis, IN: John Wiley & Sons, Inc.

Blair, J., D. Rabine, and Michelle Hofton. 1999. "The Laser Vegetation Imaging Sensor (LVIS): A Medium-Altitude, Digitization-Only, Airborne Laser Altimeter for Mapping Vegetation and Topography." *ISPRS Journal of Photogrammetry and Remote Sensing* 54 (February): 115–22. https://doi.org/10.1016/S0924-2716(99)00002-7.

Bufton, J. L., D. J. Harding, and J. B. Garvin. 1999. "Shuttle Laser Altimeter: Mission Results and Pathfinder Accomplishments." In *Proceedings from the 1999 Shuttle Small Payloads Symposium* (Rept-99A01574), 337–345. https://ntrs.nasa.gov/search.jsp?R=19990087525.

City of Portland, Oregon. n.d. "How to Measure a Tree."Resources page. Accessed November 17, 2019. https://www.portlandoregon.gov/trees/article/424017.

Collin, Antoine, Bernard Long, and Phillippe Archambault. 2011. "Benthic Classifications Using Bathymetric LIDAR Waveforms and Integration of Local Spatial Statistics and Textural Features." *Journal of Coastal Research*, Special Issue 62, 86–98. https://www.jcronline.org/doi/abs/10.2112/SI_62_9.

Crutchley, Simon, and Peter Crow. 2010. *The Light Fantastic: Using Airborne LiDAR in Archaeological Survey*. Historic England. http://www.historicengland.org.uk/images-books/publications/light-fantastic/.

Dang, Nguyet. 2012. "Error Propagation in Carbon Estimation Using the Combination of Airborne LiDAR and Very High Resolution GEO-EYE Imagery in Ludhikhola Watershed, Nepal." Master's thesis, University of Twente, Netherlands. https://doi.org/10.13140/RG.2.2.20020.65920.

Davis, Andrew. 2018. "Forestry Identification with LIDAR Waveform and Point Clouds." Master's thesis, Naval Postgraduate School. http://hdl.handle.net/10945/59646.

Deems, Jeffrey S., Thomas H. Painter, and David C. Finnegan. 2013. "Lidar Measurement of Snow Depth: A Review." *Journal of Glaciology* 59 (215): 467–479. doi:10.3189/2013JoG12J154.

Geomatics Data Solutions, Inc. n.d. "Bathymetric & Topographic Lidar." Accessed August 15, 2019. http://www.geomaticsdatasolutions.com/bath-topo-lidar.

Google Earth. 2019a. "Point Lobos." Accessed November 20, 2019. https://earth.google.com/web/@36.5211219,-121.95198897,28.61020613a,488.38775837d,35y,32.48131457h,44.24832406t,360r/data=ChcaFQoNL2cvMTFiNmJja2ZoMxgCIAEoAg.

———. 2019b. "Point Lobos." Accessed November 20, 2019. https://earth.google.com/web/@36.51792028,-121.94727818,33.09664916a,1008.50555755d,35y,32.48495354h,44.2515877t,0r.

Gregersen, Erick. 2016. "Lidar − Scientific Technique." *Encyclopedia Britannica*. October 13, 2016. https://www.britannica.com/technology/lidar.

Guenther, Gary C., A. Grant Cunningham, Paul E. LaRocque, and David J. Reid. 2000. "Meeting the Accuracy Challenge in Airborne Bathymetry." In *Proceedings of the 20th EARSeL Symposium: Workshop on LiDAR Remote Sensing of Land and Sea*. https://www.semanticscholar.org/paper/the-accuracy-challenge-in-airborne-lidar-bathymetry-Guenther-Cunningham/7bdbfb3b12fd05365a52e0745fa4ad3222043b3f

Harding, David J., Geoffrey G. Parker, Michael A. Lefsky, and Warren B. Cohen. 2002. "Lidar Remote Sensing for Ecosystem Studies: Lidar, an Emerging Remote Sensing Technology That Directly Measures the Three-Dimensional Distribution of Plant Canopies, Can Accurately Estimate Vegetation Structural Attributes and Should Be of Particular Interest to Forest, Landscape, and Global Ecologists." *BioScience* 52 (1): 19–30. https://doi.org/10.1641/0006-3568(2002)052[0019:LRSFES]2.0.CO;2.

Harris Geospatial. n.d. "K-Means." Accessed November 18, 2019. https://www.harrisgeospatial.com/docs/KMeansClassification.html.

———. n.d. "Spectral Angle Mapper." Accessed November 18, 2019. https://www.harrisgeospatial.com/docs/spectralanglemapper.html.

———. n.d. "Support Vector Machine Background." Accessed November 20, 2019. https://www.harrisgeospatial.com/docs/backgroundsvmgeneral.html.

Hexagon Geosystems. n.d. "Leica ScanStation P50 – Long Range 3D Terrestrial Laser Scanner." HEXAGON GEOSYSTEMS. Accessed March 31, 2019. https://leica-geosystems.com/products/laser-scanners/scanners/leica-scanstation-p50.

Hofton, Michelle A., Bernard Minster, and J. Bryan Blair. 2000. "Decomposition of Laser Altimeter Waveforms" 38 (August): 1989–91. https://doi.org/10.1109/36.851780.

Hudak, Andrew T., Michael A. Lefsky, Warren B. Cohen, and Mercedes Berterretche. 2002. "Integration of Lidar and Landsat ETM+ Data for Estimating and Mapping Forest Canopy Height." *Remote Sensing of Environment* 82 (2–3): 397–416. https://doi.org/10.1016/S0034-4257(02)00056-1.

Kelly, Maggi, and Stefania Di Tommaso. 2015. "Mapping Forests with Lidar Provides Flexible, Accurate Data with Many Uses." *California Agriculture* 69 (1): 14–20. http://dx.doi.org/10.3733/ca.v069n01p14.

Neuenschwander, Amy L., Ramses Gutiérrez, B. Schutz, and Tim Urban. 2006. "Comparison of Small-Footprint and Large-Footprint Waveform Lidar for Terrestrial Surface Characterization." In *Proceedings from IEEE Geoscience and Remote Sensing Symposium*. https://doi.org/10.1109/IGARSS.2006.963.

Leeuwen, Martin, and M. Disney. 2017. "Vegetation Structure (LiDAR)." In *Reference Module in Earth Systems and Environmental Sciences*, 104–16. University College London, London, United Kingdom: Elsevier Inc. https://doi.org/10.1016/B978-0-12-409548-9.10543-3.

Lefsky, Michael A., David J. Harding, Michael Keller, Warren B. Cohen, Claudia C. Carabajal, Fernando Del Bom Espirito-Santo, Maria O. Hunter, and Raimundo de Oliveira Jr. 2005. "Estimates of Forest Canopy Height and Aboveground Biomass Using ICESat." *Geophysical Research Letters* 32 (22): 1–4.

Leigh, Holly W., and Lori A. Magruder. 2016. "Using Dual-Wavelength, Full-Waveform Airborne Lidar for Surface Classification and Vegetation Characterization." *Journal of Applied Remote Sensing* 10 (4): 1–19. https://doi.org/10.1117/1.JRS.10.045001.

Lewis, P., and S. Hancock. 2007. "LiDAR for Vegetation Applications." Teacher's notes/supplemental material for MSc Remote Sensing course: Airborne LiDAR Systems, University College of London, UK. https://pdfs.semanticscholar.org/916b/171fc753d9551a4dd0cf0a594667dadba562.pdf.

LiDAR News. 2016. "Dewberry Selected to Produce High-Resolution Lidar Topography and Bathymetry Data for Puerto Rico." LiDAR News. February 5, 2016. https://lidarnews.com/press-releases/dewberry-selected-to-produce-high-resolution-lidar-topography-and-bathymetry-data-for-puerto-rico/.

Marcu, Casiana, Florian Stătescu, and Nicoleta Iurist. 2017. "A GIS-Based Algorithm to Generate a Lidar Pit-Free Canopy Height Model." *Present Environment and Sustainable Development* 11 (2): 89–95. https://doi.org/10.1515/pesd-2017-0027.

McClung, F., and Robert Hellwarth. 1962. "Giant Optical Pulsations from Ruby." *Applied Optics* 1 (March): 103–5. https://doi.org/10.1063/1.1777174.

McIver, Charles A. 2017. "Spectral Lidar Analysis and Terrain Classification in a Semi-Urban Environment." Master's thesis, Naval Postgraduate School. https://calhoun.nps.edu/handle/10945/53017.

Miller, Barry. 1965. "Laser Altimeter May Aid Photo Mapping." *Aviation Week & Space Technology*, March 29, 1965.

Olsen, Richard. 2016. *Remote Sensing from Air and Space*. 2nd ed. Bellingham, WA: SPIE Press.

———. 2007. *Remote Sensing from Air and Space*. 1st ed. Bellingham, WA: SPIE Press.

Olsen, Richard, Andrew Davis, and Jeremy Metcalf. 2019. *Analysis and Exploitation of Lidar Waveform Data*. Bellingham, WA: SPIE Press. https://doi.org/10.1117/12.2519188.

Point Lobos Foundation. 2018. "Trees." Learn More page. Accessed November 17, 2019. https://www.pointlobos.org/learn-more/wildlife-plants-birds-and-geology/plants/trees.

Popescu, Sorin C., Kaiguang Zhao, Amy Neuenschwander, and Chinsu Lin. 2011. "Satellite Lidar vs. Small Footprint Airborne Lidar: Comparing the Accuracy of Aboveground Biomass Estimates and Forest Structure Metrics at Footprint Level." *DESDynI VEG-3D Special Issue* 115 (11): 2786–97. https://doi.org/10.1016/j.rse.2011.01.026.

Quadros, Nathan. 2013. "Unlocking the Characteristics of Bathymetric LiDAR Sensors." *LiDAR Magzine* 3 (6): 62–67. http://lidarmag.com/wp-content/uploads/PDF/LiDARMagazine_Quadros-BathymetricLiDARSensors_Vol3No6.pdf.

Reitberger, Josef, Peter Krzystek, and Uwe Stilla. 2006. "Analysis of Full Waveform Lidar Data for Tree Species Classification." *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 36 (Part 3): 228–233.

Rizaev, Igor G., Anatoliy V. Pogorelov, and Maria A. Krivova. 2016. "A Technique to Increase the Efficiency of Artifacts Identification in Lidar-Based Canopy Height Models." *International Journal of Remote Sensing* 37 (7): 1658–70. https://doi.org/10.1080/2150704X.2016.1160299.

Schawlow, Arthur, and Charles Townes. 1958. "Infrared and Optical Masers." *Physical Review* 112 (December): 1940–49. https://doi.org/10.1103/PhysRev.112.1940.

Shan, Jie, and Charles K. Toth. 2018. *Topographic Laser Ranging and Scanning: Principles and Processing*. 2nd ed. Boca Raton, FL: Taylor & Francis, CRC Press.

Stigermark, Carl Johan. 2012. "Airborne Hydrography AB LEVI." Terra Solid. Last modified February 15, 2012. https://www.terrasolid.com/download/presentations/2012/ahab.pdf.

Teledyne Optech. n.d. "Titan." Products page. Accessed August 15, 2019. https://www.teledyneoptech.com/en/products/airborne-survey/titan/.

Tyo, J., Athanasios Konsolakis, David Diersen, and Richard Olsen. 2003. "Principal-Components-Based Display Strategy for Spectral Imagery." In *IEEE Transactions on Geoscience and Remote Sensing* 41 (April): 708–18. https://doi.org/10.1109/TGRS.2003.808879.

Wasser, Leah. 2018. "Canopy Height Models, Digital Surface Models & Digital Elevation Models - Work with LiDAR Data in Python." Earth Data Science - Earth Lab. February 5, 2018. https://www.earthdatascience.org/courses/earth-analytics-python/lidar-raster-data/lidar-chm-dem-dsm/.

Wasser, Leah A. n.d. "The Basics of LiDAR - Light Detection and Ranging - Remote Sensing." (Online tutorial.) National Science Foundation - Neon Science. Accessed March 6, 2019. https://www.neonscience.org/lidar-basics.

Webster, Timothy, Kevin Mcguigan, Nathan Crowell, Kate Collins, and Candace Macdonald. 2016. "Optimization of Data Collection and Refinement of Post-Processing Techniques for Maritime Canada's First Shallow Water Topographic-Bathymetric Lidar Survey." *Journal of Coastal Research* 76 (December): 31–43. https://doi.org/10.2112/SI76-004.

Wikipedia. 2019. S.v. "Global Positioning System." Accessed October 30, 2019. https://en.wikipedia.org/w/index.php?title=Global_Positioning_System&oldid=923690288.

———. 2019. S.v. "ICESat." Accessed October 30, 2019. https://en.wikipedia.org/w/index.php?title=ICESat&oldid=918278564.

Yilmaz, Volkan, C. Serifoglu Yilmaz, Levent Tasci, and Oguz Gungor. 2017. "Determination of Tree Crown Diameters with Segmentation of a UAS-Based Canopy Height Model." *IPSI Bgd Transactions on Internet Research* 13 (2): 63–67.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California