



**NAVAL
POSTGRADUATE
SCHOOL**

MONTEREY, CALIFORNIA

THESIS

**TERRAIN CLASSIFICATION USING MULTI-
WAVELENGTH LIDAR DATA**

by

Judson J. C. Thomas

September 2015

Thesis Advisor:
Second Reader:

Richard Olsen
Jeremy Metcalf

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			<i>Form Approved OMB No. 0704-0188</i>
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 September 2015	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE TERRAIN CLASSIFICATION USING MULTI-WAVELENGTH LIDAR DATA		5. FUNDING NUMBERS	
6. AUTHOR(S) Thomas, Judson J. C.		8. PERFORMING ORGANIZATION REPORT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		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. IRB Protocol number ____N/A____.	
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		12b. DISTRIBUTION CODE A	
13. ABSTRACT (maximum 200 words) <p>With the arrival of Optech's Titan multispectral LiDAR sensor, it is now possible to simultaneously collect three different wavelengths of LiDAR data. Much of the work performed on multispectral LiDAR data involves gridding the point cloud to create Digital Elevation Models and multispectral image cubes. Gridding and raster analysis can have negative implications with respect to LiDAR data integrity and resolution. Presented here is a method of attributing the Titan LiDAR point cloud with the spectral information of all three lasers and the potential improvement of performing all analysis within the point cloud.</p> <p>Data from the Optech Titan are analyzed for purposes of terrain classification, adding the spectral component to the LiDAR data point cloud analysis. The approach used here combines the three spectral sensors into one point cloud, integrating the intensity information from the 3 sensors. Nearest-neighbor sorting techniques are used to create the merged point cloud. Standard LiDAR and spectral classification techniques are then applied.</p> <p>The ENVI spectral tool "n-Dimensional Visualizer" is used to extract spectral classes from the data, which can then be applied using supervised classification functions. The Maximum Likelihood classifier provided consistent results demonstrating effective terrain classification for as many as eleven classes.</p>			
14. SUBJECT TERMS remote sensing, LiDAR, multi-wavelength, terrain classification, spectral		15. NUMBER OF PAGES 85	
		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

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

TERRAIN CLASSIFICATION USING MULTI-WAVELENGTH LIDAR DATA

Judson J. C. Thomas
Lieutenant, United States Navy
B.S., United States Naval Academy, 2008

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SPACE SYSTEMS OPERATIONS

from the

**NAVAL POSTGRADUATE SCHOOL
September 2015**

Author: Judson J. C. Thomas

Approved by: Richard Olsen
Thesis Advisor

Jeremy Metcalf
Second Reader

Rudy Panholzer
Chair, Space Systems Academic Group

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

With the arrival of Optech's Titan multispectral LiDAR sensor, it is now possible to simultaneously collect three different wavelengths of LiDAR data. Much of the work performed on multispectral LiDAR data involves gridding the point cloud to create Digital Elevation Models and multispectral image cubes. Gridding and raster analysis can have negative implications with respect to LiDAR data integrity and resolution. Presented here is a method of attributing the Titan LiDAR point cloud with the spectral information of all three lasers and the potential improvement of performing all analysis within the point cloud.

Data from the Optech Titan are analyzed here for purposes of terrain classification, adding the spectral component to the LiDAR data point cloud analysis. The approach used combines the three spectral sensors into one point cloud, integrating the intensity information from the 3 sensors. Nearest-neighbor sorting techniques are used to create the merged point cloud. Standard LiDAR and spectral classification techniques are then applied.

The ENVI spectral tool "n-Dimensional Visualizer" is used to extract spectral classes from the data, which can then be applied using supervised classification functions. The Maximum Likelihood classifier provided consistent results demonstrating effective terrain classification for as many as eleven classes.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
A.	PURPOSE OF RESEARCH	1
B.	OBJECTIVE	2
II.	BACKGROUND	3
A.	HISTORY OF LIDAR.....	3
B.	LIDAR PRINCIPLES	4
1.	Terrain Classification using LiDAR.....	5
2.	Efforts to Utilize Spectral Intensity and Reflectance Data	6
a.	<i>Correcting LiDAR Intensity Data.....</i>	<i>10</i>
3.	Multi-Wavelength LiDAR Efforts.....	11
III.	DATA SET.....	19
A.	INSTRUMENT	19
1.	Optech Titan LiDAR	19
2.	Data Collection	20
B.	DATA PREPARATION.....	21
1.	Wavelength Merging and Noise Clipping.....	21
2.	Developing a Standard Classification Model	22
3.	Manual Classification	25
4.	Flightline Merging and Flagging	28
IV.	DATA ANALYSIS	29
A.	ANALYSIS WORKFLOW	29
B.	VEGETATION INDICES.....	31
C.	UTILIZING MULTI-SPECTRAL LIDAR INTENSITIES TO DERIVE SUB-CLASSIFICATIONS	31
1.	Generating Spectral Training Sets	31
2.	Alternate Method for Training Set Derivation	38
V.	RESULTS	41
A.	MAXIMUM LIKELIHOOD.....	41
1.	Spectral Classes Derived From Training Areas.....	43
2.	Spectral Classes Derived From Random Subset Method	52
B.	OTHER CLASSIFICATION TOOLS.....	55
VI.	CONCLUSION	57
A.	STEPS TOWARD OPERATIONAL UTILITY	57
B.	SPACE APPLICATION	57
C.	FUTURE WORK	58
	LIST OF REFERENCES.....	61
	INITIAL DISTRIBUTION LIST	65

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

Figure 1.	Residential building footprints developed from LiDAR.	6
Figure 2.	Terrain classification conducted using LiDAR intensity, a generated DSM, normalized height, and return number.	8
Figure 3.	“LiDAR near infrared intensity and Sitka spruce volume.”	9
Figure 4.	Intensity raster image a) before intensity correction and b) after intensity correction.	11
Figure 5.	“Schematic green LiDAR waveform showing the three principal signal components.”	12
Figure 6.	(a) MWCL Targets; (b) three dimensional reconstruction of MWCL LiDAR detection; and (c) MWCL classification results.....	14
Figure 7.	Hyperspectral LiDAR waveforms at various stages of processing, with the thick black line representing the mean waveform of all channels. In plot c) trigger and target parts of the waveforms are normalized in different scale. The negative overshoot is visible, e.g., after the trigger pulse.....	15
Figure 8.	“Comparison of spectra collected from the Norway spruce using the hyperspectral LiDAR and a passive spectrometer.”	16
Figure 9.	Pseudo-NDVI of three layers within the vertical structure of the forest. (Top) First return from the LiDAR instrument, including the ground layer; (Middle) The mid-section of the forest canopy; (Bottom) Low section of the forest canopy and the ground layer. This demonstrates the differences in NDVI throughout the vertical canopy.	17
Figure 10.	Optech Titan operating wavelengths and typical spectral responses.....	19
Figure 11.	Optech Titan multi-wavelength LiDAR system.	20
Figure 12.	False color image of the Optech Titan data set in Quick Terrain Modeler.....	22
Figure 13.	Results of LASground program separating points into “ground” and “non-ground” shown in QTM.	23
Figure 14.	Results of LASclassify program separating points into unclassified, ground, vegetation, and buildings shown in QTM.	24
Figure 15.	Training Region #1 – Middle School shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.	25
Figure 16.	Training Region #2 – Wooded Power Station/Antenna shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.	26
Figure 17.	Training Region #3 – Beachfront shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.....	26
Figure 18.	Training Region #1 - Raw result from LASclassify showing both classification errors and unclassified areas.	27
Figure 19.	Training Region #1 - Manually classified data set.	27
Figure 20.	Analysis workflow for spectral classification of multi-wavelength LiDAR. ...	30
Figure 21.	Ground points within the training regions viewed in spectral space in N-D VIS, prior to spectral classes being identified.	33

Figure 22.	Ground points in training regions viewed in spectral space in N-D VIS, with six spectral classes having been identified and reassigned as other classes.	34
Figure 23.	Ground points in Training Region #1 viewed in “X,Y” space in N-D VIS.....	35
Figure 24.	Google Earth imagery of Training Region #1.	36
Figure 25.	Spectral Library Plot of classes developed in N-D VIS.	38
Figure 26.	Study area with training regions boxed in red.	42
Figure 27.	Demonstration of ML Results with no probability threshold showing all 11 classes.	43
Figure 28.	Demonstration of ML Results with no probability threshold showing classes 2–4 and 11, highlighting roads and road paint.....	44
Figure 29.	Demonstration of ML Results with no probability threshold showing spectral classes 7–9, highlighting buildings.....	44
Figure 30.	Demonstration of ML Classifier results with no probability threshold showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.	45
Figure 31.	Demonstration of ML Classifier results with a high probability threshold (0.85) showing all 11 spectral classes.....	46
Figure 32.	Demonstration of ML Classifier results with a high probability threshold (0.85) showing spectral classes 2–4 and 11, highlighting roads and road paint.....	46
Figure 33.	Demonstration of ML Classifier results with a high probability threshold (0.85) showing classes 7–9, highlighting buildings.....	47
Figure 34.	Demonstration of ML Classifier results with a high probability threshold (0.85) showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.	47
Figure 35.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing all 11 spectral classes.....	48
Figure 36.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 2–5 and 11, highlighting roads, railroads and road paint.....	49
Figure 37.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 7–9, highlighting buildings.	49
Figure 38.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.....	50
Figure 39.	Detail view of spectral classes 1 and 6 which represent vegetation, and errors in the spectral classification process from one neighborhood (yellow) and the lake water (orange).	51
Figure 40.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing all 9 spectral classes derived by the random subset method.....	53
Figure 41.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 2–4, highlighting roads, driveways and railroad.....	53

Figure 42.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 6–9, highlighting buildings,	54
Figure 43.	Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 1 and 5, highlighting vegetation separated by grass and trees.....	54
Figure 44.	Detailed view of performance advantage for random subset classification in the instance of one neighborhood.	55

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1.	Comparison of multi-wavelength sensors developed for vegetation analysis.....	14
Table 2.	List of classes produced by LASground and LASclassify software.....	23
Table 3.	Distribution of point classification from LASclassify program.	25
Table 4.	Listing of Spectral Classes generated in N-D VIS using training regions.....	37
Table 5.	Listing of Spectral Classes generated in N-D VIS using random sampling. ...	40

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

3D	Three Dimensional
A2/AD	Anti-Access/Area Denial
AGL	Above Ground Level
ASCII	American Standard Code for Information Interchange
DSM	Digital Surface Model
DTM	Digital Terrain Model
EO	Electro-optical Imagery
GDVI	Green Difference Vegetation Index
GLAS	Geoscience Laser Altimeter System
GNDVI	Green Normalized Difference Vegetation Index
GRVI	Green Ratio Vegetation Index
GPS	Global Positioning System
IMU	Internal Motion Unit
INS	Inertial Navigation System
IR	Infrared
kHz	Kilo-Hertz
LEO	Low-Earth Orbit
LiDAR	Light Detection and Ranging
LMS	LiDAR Mapping Suite
ML	Maximum Likelihood
NIR	Near Infrared
N-D VIS	n-Dimensional Visualizer
NDVI	Normalized Difference Vegetation Index
nm	Nanometers
OPLAN	Operations Plan
PNT	Positioning, Navigation, and Timing
PRI	Photochemical Reflectance Index
QTM	Quick Terrain Modeler
ROI	Region of Interest
SAM	Spectral Angle Mapper

TCARI

Transformed Chlorophyll Absorption in Reflectance Index

ACKNOWLEDGMENTS

A special thank you to

Dr. R. C. Olsen,
Jeremy Metcalf
Chad Miller
Angela Kim

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. PURPOSE OF RESEARCH

As operational planning staffs endeavor to better their procedures for developing coherent courses of action, one area that can provide substantial benefit is accurate knowledge of the operating environment. Reliable operations plans (OPLANS) are based heavily on the ability to correctly determine the nature of the surroundings in which operations will take place. Remote sensing is the discipline of observing the Earth from satellite or aircraft for the purpose of characterizing the environment. Specifically identifying buildings, roads, trails, features, and vegetation is termed terrain classification. In the field of remote sensing, electro-optical (EO) imaging, radar imaging, and light detection and ranging (LiDAR) are examples of platforms employed for the purpose of terrain classification.

While EO imagery offers advantages for the terrain classification process in terms of spectral fidelity, and radar imagery has capabilities unique to its sensors, neither offer the same vertical resolution provided by a LiDAR point cloud. LiDAR, like radar, is an active sensor and can be employed regardless of lighting conditions. Efforts and progress have been made primarily in the field of merging photogrammetric imagery with LiDAR data sets in order to support better terrain classification and evaluation. There are issues, however, stemming mainly from the complexities of merging data in differing formats which hamper the effectiveness of the product. Multi-spectral LiDAR may provide a way to maintain the analysis entirely in the LiDAR point cloud, and would only require one sensor.

Separately, researchers have begun to investigate the utility of analyzing not only the spatial components of LiDAR data, but also the spectral data. With motivations primarily driven by vegetation analysis, first-time experiments are being conducted to attempt to classify vegetation on the species level using spectral returns. Multi-wavelength LiDAR sensors have been implemented in many of these studies, as well as being historically useful in bathymetric efforts.

Given that multi-wavelength LiDAR systems are now being fielded, there is a potential to better the terrain classification techniques currently employed by taking advantage of the spectral portion of the data collected by the sensor. Classification using spectral information may be able to act as an error correction following implementation of current methods. Furthermore, the spectral components have the potential to further sub-classify individual points. The advantage provided to the operations planning process by accurately knowing the material makeup of the environment from roads to buildings, or the species of a given stand of vegetation, could be tremendous, particularly if only LiDAR is available.

B. OBJECTIVE

The objective of this thesis is to test the idea that the spectral intensity components of a multi-wavelength LiDAR data set can lend accuracy to the terrain classification process, and to evaluate the effectiveness of sub-classification techniques based on the spectral data. To meet this objective, data collected by the Optech Titan LiDAR sensor over the city of Toronto will be analyzed.

II. BACKGROUND

A. HISTORY OF LIDAR

LiDAR is but one fascinating utilization of the laser, or Light Amplification by Stimulated Emission of Radiation. What was theorized by Townes and Schawlow¹ and named by Gould,² quickly became reality as Maiman and the Hughes Research Laboratory created the first working laser in 1960.³ Fiocco and Smullen worked to analyze the upper atmosphere through a pulsed probing technique, arguably the first instance of a laser being employed as a remote sensor.⁴ Shortly thereafter, the orientation was reversed and LiDAR was used to measure the topographic profile of a football stadium in Philadelphia.⁵ These measurements took advantage of the inherently short wavelength of LiDAR. Laser pulses are directed at a target, usually the Earth, and then the sensor measures the time required for the pulse to return. This time is then used to calculate a distance, the accuracy of which is heavily dependent on the fidelity of the clock used to measure the time between pulse and return. When the laser pulse is returns to the LiDAR sensor it is collected by a detector with the purpose of gathering the desired data set. LiDAR detectors usually implement a photo-multiplier tube, or an avalanche photodiode in order to amass the photons for recognition. Detectors may be designed to obtain returns which contain the full waveform, individual photon, or a discrete number of critical ranges.

¹ Arthur L. Schawlow and Charles H. Townes, "Infrared and Optical Masers," *Physical Review* 112, no. 6, (15 December 1958): 1948–1949, doi: <http://dx.doi.org/10.1103/PhysRev.112.1940>.

² Gordon R. Gould, "The LASER, Light Amplification by Stimulated Emission of Radiation," *The Ann Arbor conference on optical pumping, the University of Michigan*, (15 June 1959), quoted in Richard Olsen, *Remote Sensing from Air and Space* (SPIE Press, 2007), 229.

³ Theodore H. Maiman, "Stimulated Optical Radiation in Ruby," *Nature* 187, no. 4736, (6 August 1960): 493–494, doi: 10.1038/187493a0.

⁴ G. Fiocco and L.D. Smullen, "Detection of Scattering Layers in the Upper Atmosphere (60–140 km) by Optical Radar," *Nature* 199, no. 4900, (28 September 1963): 1275–1276, doi: 10.1038/1991275a0.

⁵ Barry Miller, "Laser Altimeter May Aid Photo Mapping," *Aviation Week & Space Technology* 83, no. 13, (29 March 1965): 60–61.

As LiDAR technology has continued to develop over the past 50 years, three major components of the data have been analyzed. The primary resource that LiDAR provides is extremely accurate three dimensional point clouds, or a set of data points with x, y, and z coordinates. This aspect of the data alone has been the basis for massive advances in the accurate production of digital elevation models, or DEMs. LiDAR data also contains intensity values, which are a function of the reflectance of the material they hit. Originally, intensity or reflectance data were used for the production of rasterized images which provided some utility based on their precision regardless of illumination conditions. Finally, while the production of most topographical products make use of only the first and last returns for a given pulse transmission, modern LiDAR sensors can record a full waveform return, which provides the analyst more detail into the power received from each pulse transmitted.⁶

B. LIDAR PRINCIPLES

LiDAR detectors have progressed rapidly since the inception of the red ruby laser. Most modern LiDAR systems operate at one frequency and use one of a variety of media including solid state, gas, excimer, dye, and semiconductors. As in all lasers, photons are pumped into the lasing medium, and consequently stimulated and amplified in an optical cavity. The collimated resulting beam is both coherent and polarized, which add to its low divergence to make it operationally valuable as a remote sensing tool.

Topographic LiDAR systems are typically mounted underneath an aircraft and flown over the area of interest in order to gather the desired data set. One of the critical advancements making LiDAR operationally valuable is the advent of Positioning, Navigation, and Timing (PNT) systems such as the U.S. GPS constellation. Not only does PNT provide the geolocation of the LiDAR sensor so that its data are geographically referenced, it also provides precision timing on the order of nanoseconds to the detector and central processing unit which measure the LiDAR pulse travel time. Additionally, internal motion units (IMU) are flown with the LiDAR sensor which measure and

⁶ Linda Nordin, "*Analysis of Waveform Data from Airborne Laser Scanner System*," (master's thesis, Lulea University of Technology, 2006): 38.

account for the rotational motion of the host platform on three axes. These two systems make up the internal navigation system (INS), and are critically important to both the accuracy and utility of LiDAR as a remote sensing tool. The basic output of a LiDAR sensor is a geographically referenced range determined by:

$$R = c\tau/2$$

where:

R = Range in meters

c = the speed of light, 3×10^8 m/s

τ = laser pulse travel time in seconds

1. Terrain Classification using LiDAR

The first and most prominent method of terrain classification using LiDAR data makes use of the highly accurate three-dimensional spatial data set to conduct feature extraction or object recognition. These methods often rely on a set algorithms developed for use in specific endeavors. For building footprint construction, Zhang used a morphological filter, a region-growing algorithm, and a noise removal algorithm to produce footprint maps like the one shown in Figure 1.⁷

⁷ Keqi Zhang, "Automatic Construction of Building Footprints from Airborne LiDAR Data," *IEEE Transactions on Geoscience and Remote Sensing* 44, no. 9, (September 2006): 2532, doi: 10.1109/TGRS.2006.874137.



Figure 1. Residential building footprints developed from LiDAR.⁸

In 2005, Helt investigated the ability to distinguish between three different tree species using statistical analysis of both height information and foliage dispersion.⁹ He was able to accurately map the locations of California Scrub Oaks, California Live Oaks, and Eucalyptus trees using a LiDAR data set taken over the Elkhorn Slough Wetlands.

2. Efforts to Utilize Spectral Intensity and Reflectance Data

LiDAR intensity measurements have traditionally been considered secondary data when compared to the spatial components of the three dimensional high-resolution point cloud which can now be modeled and displayed with modern computing resources. Many of the most interested parties in utilizing intensity returns for analysis are focused on in-depth analysis of forest canopies. With goals like species classification, tree crown measurement, and sub-canopy health evaluation, LiDAR sensors can provide both access

⁸ Ibid., 2531.

⁹ Michael Helt, “*Vegetation Analysis with LiDAR*,” (master’s thesis, Naval Postgraduate School, September 2005): 30–40, retrieved from Calhoun: <http://hdl.handle.net/10945/2030> on April 28, 2015.

to the variable levels of forested areas, but also an intensity return which can be indicative of biological tendencies.

As early as 2002, efforts were being made to use intensity data from LiDAR systems to perform terrain classification.¹⁰ Researchers used various techniques, many of which involved meshed classification using both intensity data, and physical return metrics to differentiate between specific species of trees.¹¹ In 2006 and 2007, multiple attempts were made to assess land coverage and classification using LiDAR intensities,¹² with varying levels of success.

¹⁰ Jeong-Heon Song, Soo-Hee Han, Kiyun Yu, and Yong-II Kim, "Assessing the Possibility of Land-Cover Classification Using LIDAR Intensity Data," *ISPRS Commission III, "Photogrammetric Computer Vision" Graz, Austria* 34, no. 3B, (9-13 September, 2002): 259–262, retrieved from: <http://www.isprs.org/proceedings/XXXIV/part3/>.

¹¹ Johan Holmgren and Åsa Persson, "Identifying Species of Individual Trees Using Airborne Laser Scanner," *Remote Sensing of Environment* 90, no. 4, (21 May 2003): 415–423, doi: 10.1016/S0034-4257(03)00140-8. ; Tomas Brantberg, "Classifying Individual Tree Species Under Leaf-off and Leaf-on Conditions Using Airborne LiDAR," *ISPRS Journal of Photogrammetry and Remote Sensing* 61, no. 5, (January 2007): 325–340, doi: 10.1016/j.isprsjprs.2006.10.006.

¹² R. Brennan and T. L. Webster, "Object-oriented Land Cover Classification of LiDAR-derived Surfaces," *Canadian Journal of Remote Sensing* 32, no. 2, (April 2006): 162–172, doi: 10.5589/m06-015.



Figure 2. Terrain classification conducted using LiDAR intensity, a generated DSM, normalized height, and return number.¹³

¹³ Ibid., 171.

Donaghue was able to utilize intensity and the coefficient of variation to dissect the species populations in coniferous plantations.¹⁴ At multiple sites, he accurately mapped the Sitka Spruce tree fraction, providing a functional use for the intensity data. Figure 3 shows that the intensity data for NIR LiDAR is strongly related to the spruce fraction, particularly in the case of the Laurieston forest block. He also studied how the flight geometry affects the return intensities, and how to manage the effects of off-nadir scan angles and their varying path lengths.

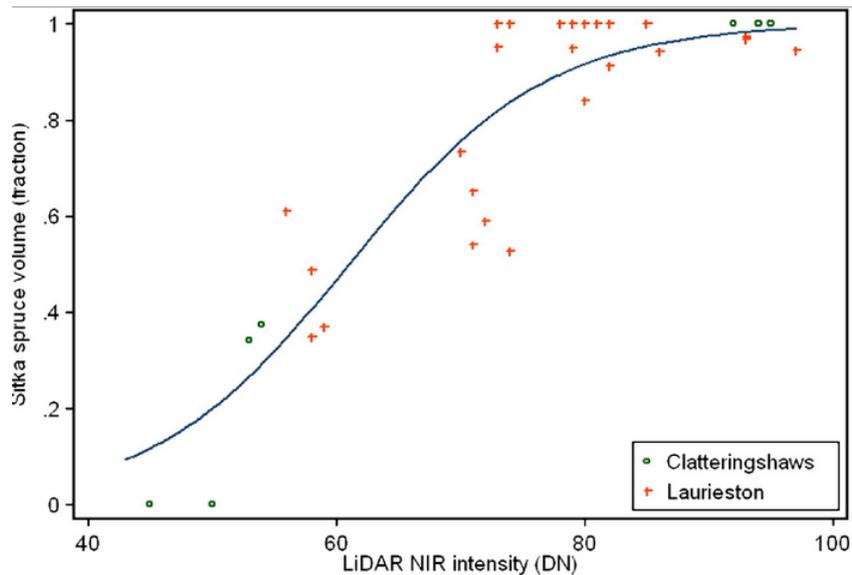


Figure 3. “LiDAR near infrared intensity and Sitka spruce volume.”¹⁵

Gaulton points out that there are issues with using intensity data to measure target properties, reflectance specifically, which originate from the scattering effects and incidence angles experienced by the laser.¹⁶ The idea of dual or multi-wavelength LiDAR sensors to correct these issues has been closely studied over the last ten years. Specific just to vegetation applications, there have been several prototype sensors

¹⁴ Daniel N. Donoghue, Peter J. Watt, Nicholas J. Cox, & Jimmy Wilson, “Remote Sensing of Species Mixtures in Conifer Plantations using LiDAR Height and Intensity Data,” *Remote Sensing of Environment* 110, no. 4, (10 October 2007): 509–522, doi: 10.1016/j.rse.2007.02.032.

¹⁵ Ibid.

¹⁶ Rachel Gaulton, “The Potential of Dual-Wavelength Laser Scanning for Estimating Vegetation Moisture Content,” *Remote Sensing of Environment* 132, (15 May 2013): 35.

developed with between two and eight different wavelengths. Dual wavelength sensors tend to be two separate lasers separated by an angular offset, while multi-wavelength systems may employ a tunable or supercontinuum laser.

a. Correcting LiDAR Intensity Data

A major consideration in utilizing intensity data procured by laser scanners is that intensity, as recorded, is not necessarily an accurate measure of the target reflectance. The major discrepancies between recorded intensity and actual reflectance are due to spherical losses, topographic, and atmospheric effects.¹⁷ Hölfe researched different methods of correcting for these errors using a data-driven correction method and a model-driven correction method.¹⁸ Figure 4 shows the effects of his data-driven correction on an intensity raster image.

¹⁷ Bernhard Hölfe, “Correction of Laser Scanning Intensity Data: Data and Model-driven Approaches,” *ISPRS Journal of Photogrammetry and Remote Sensing* 62, no. 6, (December 2007): 415, doi: 10.1016/j.isprsjprs.2007.05.008.

¹⁸ *Ibid.*, 415–433.

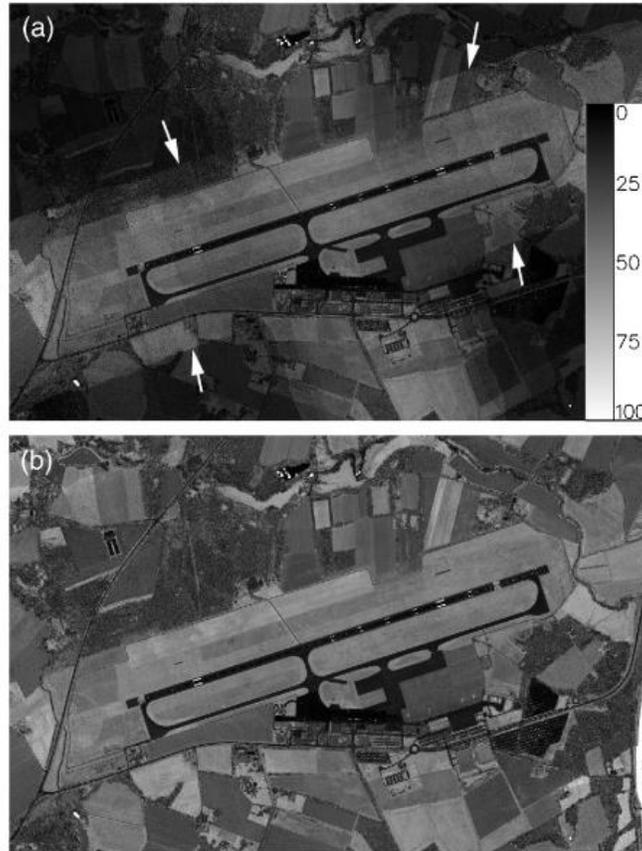


Figure 4. Intensity raster image a) before intensity correction and b) after intensity correction.¹⁹

3. Multi-Wavelength LiDAR Efforts

Multi-wavelength LiDAR has been implemented since well before it was considered for use in vegetation analysis. As early as 1969, the study of dual wavelength LiDAR systems for use in shallow water bathymetry began.²⁰ Airborne LiDAR bathymetry (ALB) takes advantage of the fact that a green wavelength laser will penetrate water without attenuating significantly, while a second laser, often infrared, will

¹⁹ Ibid.

²⁰ G. Daniel Hickman and John E. Hogg, "Application of an Airborne Pulsed Laser for Near-shore Bathymetric Measurements," *Remote Sensing of Environment* 1, no. 1, (March 1969): 47–58, doi: 10.1016/S0034-4257(69)90088-1.

hardly penetrate the surface at all. The result is a matched pair of return signatures which delineate the surface from the ocean bottom.²¹

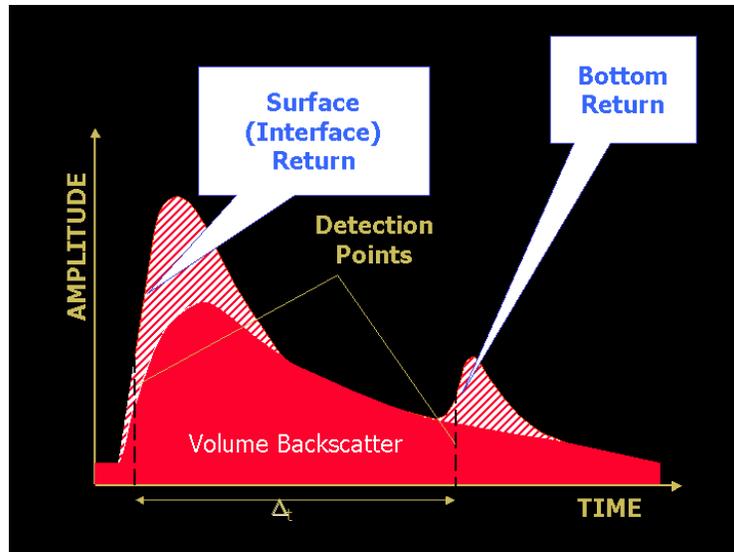


Figure 5. “Schematic green LiDAR waveform showing the three principal signal components.”²²

The bathymetric utilization of dual wavelength LiDAR systems over the past 45 years has been instrumental in developing the techniques for managing return data from two different wavelengths which arrive at the detector at differing intervals.²³ However, bathymetric exploitation of LiDAR capabilities does not make use of the intensity data, which makes it dissimilar from the goal of refined terrain classification based on intensity returns.

In the mid-2000s, researchers were exploring new ways to utilize the wealth of data returned by LiDAR systems. These efforts are summarized in Table 1. Tan and Narayanan took the Airborne Laser Polarimetric Sensor (ALPS), a NASA Goddard

²¹ Gary C. Guenther, “Meeting the Accuracy Challenge in Airborne LiDAR Bathymetry,” *Proceedings of European Association of Remote Sensing Laboratories Workshop LIDAR Dresden, Germany*, (16-17 June 2000): 3–4, accessed online at: http://www.e proceedings.org/static/vol01_1/01_1_guenther1.pdf?SessionID=63c3bb4a5564b5f62d99e on May 26, 2015.

²² Ibid.

²³ Ibid., 8–10.

system designed for hovering helicopter flight, and rebuilt it into the Multi-wavelength Airborne Polarimetric LiDAR (MAPL). MAPL used an infrared and green laser system similar to those used for bathymetry, but focused on analyzing the polarimetric returns and cross-polarization ratios for the purpose of evaluating vegetation canopies.²⁴

Woodhouse sought to exploit two existing phenomena in his study, one being a change in light use efficiency and reflectance centered at 531 nanometers, the other being the “red edge” change in reflectance of vegetation at approximately 700 nanometers. Implementing a LiDAR sensor with four wavelengths, two each straddling the critical wavelengths for the phenomena, he demonstrated the value of a LiDAR signature that could be analyzed for both photochemical reflectance index (PRI) and normalized difference vegetation index (NDVI).²⁵ The Multispectral Canopy LiDAR (MSCL) developed by Woodhouse only provided a laboratory demonstrator, but it further stimulated research into utilization of intensity and reflectance data coupled with the fidelity of a three dimensional point cloud.

This study was shortly followed by the similar multi-wavelength canopy LiDAR (MWCL) study by Wei, who demonstrated that in a controlled environment, a four wavelength LiDAR could distinguish not only between vegetation and non-vegetation, but also between plants of differing health based solely on reflectance indices.²⁶ A sample of his results is shown in Figure 6.

²⁴ Songxin Tan and Ram Narayanan, “Design and Performance of a Multiwavelength Airborne Polarimetric LiDAR for Vegetation Remote Sensing,” *Applied Optics* 43, no. 11, (10 April 2004): 2362, 2367.

²⁵ Iain Woodhouse, “A Multispectral Canopy LiDAR Demonstrator Project,” *IEEE Geoscience and Remote Sensing Letters* 8, no. 5, (21 April 2011): 839.

²⁶ Gong Wei, “Multi-wavelength Canopy LiDAR for Remote Sensing of Vegetation: Design and System Performance,” *ISPRS Journal of Photogrammetry and Remote Sensing* 69, (April 2012): 1–9, doi: 10.1016/j.isprsjprs.2012.02.001.

Researcher	System	No. of Wavelengths	Wavelengths	Focus
Tan & Narayanan	MAPL	2	532 nm 1064 nm	Polarimetric Vegetation Analysis
Woodhouse	MSCL	4	531 nm 550 nm 690 nm 780 nm	NDVI PRI
Wei	MWCL	4	555 nm 670 nm 700 nm 780 nm	NDVI GNDVI TCARI

Table 1. Comparison of multi-wavelength sensors developed for vegetation analysis.

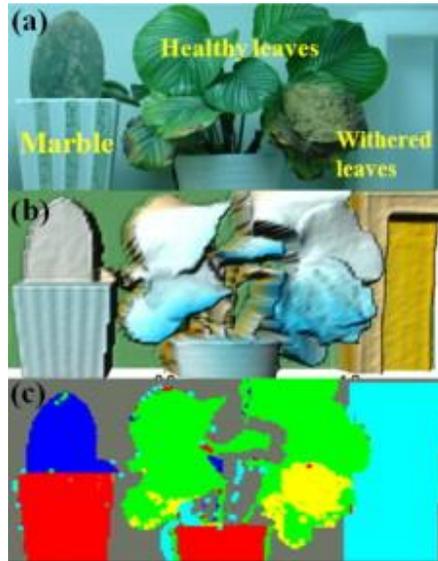


Figure 6. (a) MWCL Targets; (b) three dimensional reconstruction of MWCL LiDAR detection; and (c) MWCL classification results.²⁷

Perhaps the most complex research with multi-wavelength LiDAR was conducted by Hakala, who recognized that gathering intensity data, hyperspectral specifically, while simultaneously recording LiDAR range data represented a significant step towards efficiency in terms of analysis. Such a data set would be analytically flexible compared to

²⁷ Ibid.

monochromatic LiDAR, illumination independent compared to photogrammetric methods, and absolved of inherent registration errors when forced into data matching.²⁸

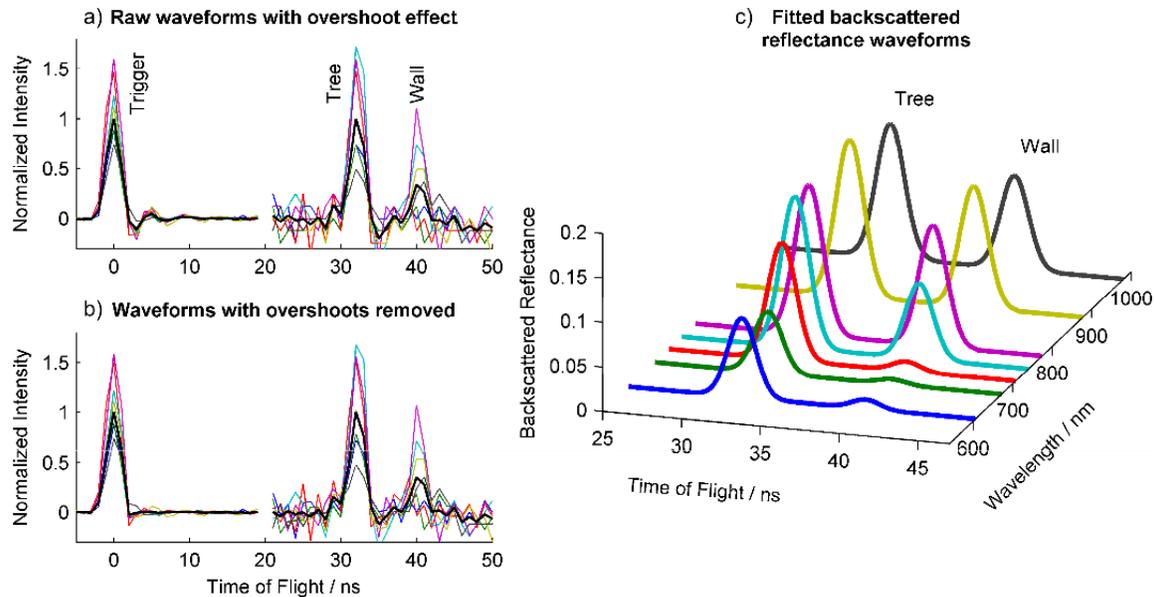


Figure 7. Hyperspectral LiDAR waveforms at various stages of processing, with the thick black line representing the mean waveform of all channels. In plot c) trigger and target parts of the waveforms are normalized in different scale. The negative overshoot is visible, e.g., after the trigger pulse.²⁹

Hakala employed a supercontinuum laser which produced broadband light collimated into eight wavelengths between 542 and 981 nanometers, providing a spectral intensity response optimized for vegetation analysis. The prototype sensor was capable of collecting multiple waveforms for investigation with the results shown in Figure 7. Hakala also demonstrated the validity of the backscattered reflectance data by comparing it to a passive spectrometer in identical laboratory conditions (Figure 8). Finally, the data was analyzed for common vegetative indices, which plainly identified the regions of the test subject spruce tree which were dead or dying.

²⁸ Teemu Hakala, "Full Waveform Hyperspectral LiDAR for Terrestrial Laser Scanning," *Optics Express*, 20, no. 7, (26 March 2012): 7120.

²⁹ *Ibid.*, 7123.

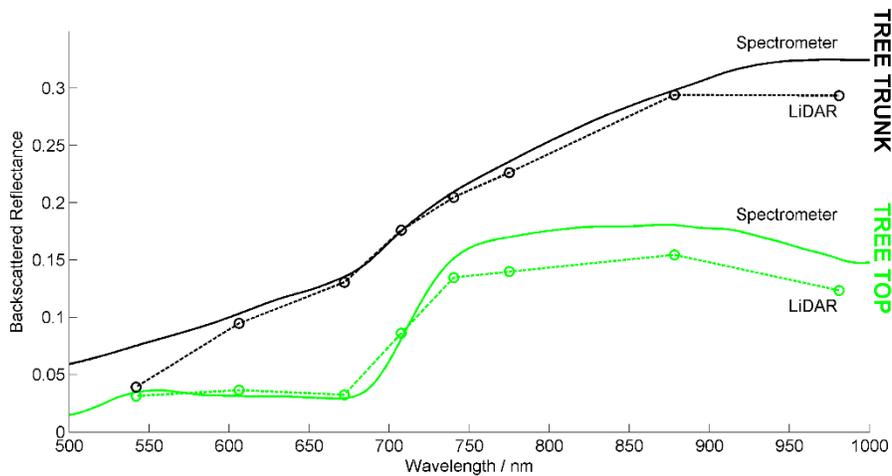


Figure 8. “Comparison of spectra collected from the Norway spruce using the hyperspectral LiDAR and a passive spectrometer.”³⁰

Making use of multi-wavelength LiDAR where the spectral points are not coincident requires a gridding and rasterization process according to a Fleming and Woodhouse article published in February 2015.³¹ Specifically referencing the Optech Titan sensor used for this thesis, they illustrate a false color rasterized image. Their own analysis describes the first real operational test of a multi-wavelength LiDAR system with the intent of fully exploiting the reflectance data for analytical purposes. Instruments flown by Riegl in June of 2013 took data in three wavelengths of 532 nm, 1064 nm, and 1550 nm. The collection involved flying three single wavelength sensors, on one airframe, over the same area with two flights. The three resultant data sets were analyzed for forestry evaluation, primarily by using a pseudo-NDVI, and stratifying the index by elevation as shown in Figure 9.³² The pseudo-NDVI is similar to the GNDVI employed by Wei. While most NDVI make use of a red channel as the visible (VIS) variable, both

³⁰ Ibid., 7126.

³¹ Sam Fleming, Iain Woodhouse, and Antoine Cotton, “Bringing Colour to Point Clouds,” *GIM International* 29, no. 2, (February 2015): 2, accessed online: <http://www.gim-international.com/content/article/bringing-colour-to-point-clouds?output=pdf>.

³² Sam Fleming, Antoine Cotton, and Iain Woodhouse, “The First Spectral Map of a Forest Understory from Multi-spectral LiDAR,” *LiDAR News Magazine* 5, no. 1, (2015): 3, accessed online: <https://carbomap.wordpress.com/2015/02/23/lidar-news-first-multispectral-lidar-map-of-a-forest-understory/>.

Fleming's pseudo-NDVI and Wei's GNDVI employ a green laser, 532 nm and 556 nm, respectively, as the visible variable.

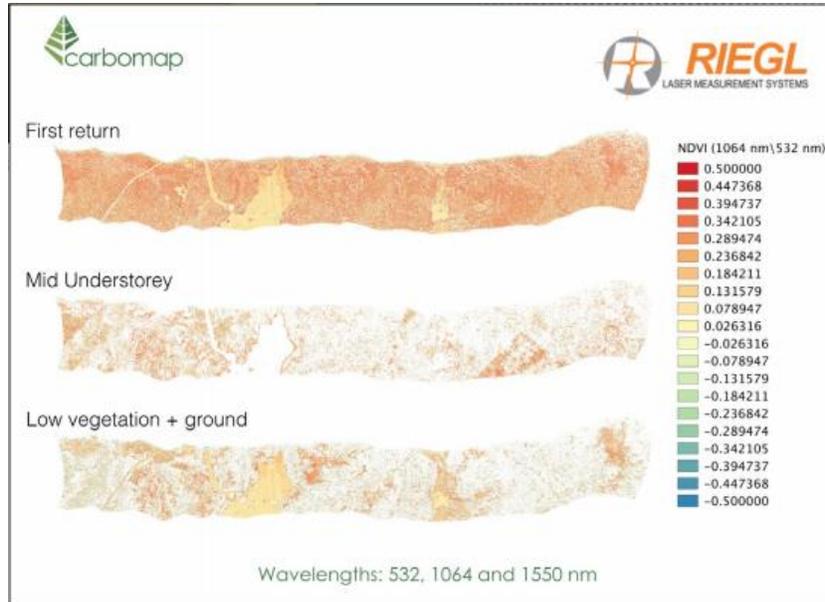


Figure 9. Pseudo-NDVI of three layers within the vertical structure of the forest. (Top) First return from the LiDAR instrument, including the ground layer; (Middle) The mid-section of the forest canopy; (Bottom) Low section of the forest canopy and the ground layer. This demonstrates the differences in NDVI throughout the vertical canopy.³³

³³ Ibid.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA SET

A. INSTRUMENT

1. Optech Titan LiDAR

The data to be analyzed in this thesis were collected by the Titan LiDAR System developed by the commercial enterprise Teledyne Optech. The Titan is a single sensor with three active lasers at 532 nm, 1064 nm, and 1550 nm. Each beam samples at 300 kHz, for an incredibly dense point cloud sampled at 900 kHz combined sample rate. The 1064 nm beam is assigned as channel two, and has zero degrees of offset from vertical, or nadir facing. Channel one is assigned to the 1550 nm beam and has a 3.5° forward offset, while channel three is assigned to the 532 nm beam with a 7° forward offset for bathymetric applications. The Titan additionally carries a range of electro-optical cameras for coincident imagery. Designed to exploit the unique capability of multi-wavelength LiDAR, its advertised uses include 3D land cover classification, vegetation mapping, shallow-water bathymetry, and dense topography mapping.³⁴

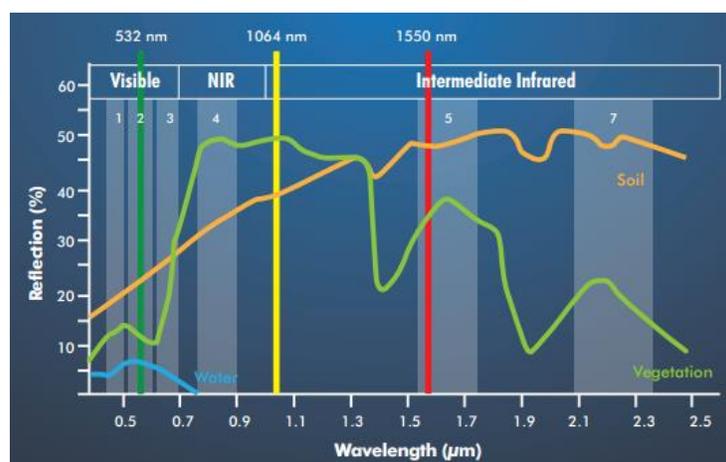


Figure 10. Optech Titan operating wavelengths and typical spectral responses.³⁵

³⁴ Optech Titan Multispectral LiDAR System: High Precision Environmental Mapping, (Brochure), n.d., 3, retrieved June 4, 2015 from <http://www.teledyneoptech.com/wp-content/uploads/Titan-Specsheet-150515-WEB.pdf>.

³⁵ Ibid., 2.



Figure 11. Optech Titan multi-wavelength LiDAR system.³⁶

2. Data Collection

In October of 2014 Optech flew the Titan sensor for an operational take in a local area which would demonstrate its unique capabilities. Port Union is a suburb of Toronto, and the collection consists largely of housing neighborhoods, butted against the western shore of Lake Ontario. Consisting of nine data files corresponding to three passes and three wavelengths, there are ample opportunities to challenge current classification methods, including various vegetation, buildings, water, sand beaches and power lines. Once the data are collected it is post-processed by Optech, which consists of running the data through their LiDAR Mapping Suite (LMS). LMS takes the raw data and assigns geographic coordinates to the LiDAR range samples through PNT and IMU correction. Post-processing in LMS does not include any form of radiometric correction to the data.

³⁶ Ibid., 1.

B. DATA PREPARATION

1. Wavelength Merging and Noise Clipping

In order to make the data set more analytically pliable, the data set was then run through a locally developed program which “merged” the different wavelengths. Starting with the nadir facing near infrared (NIR) laser at 1064 nm, the program then identifies the nearest infrared (IR) return at 1550 nm and green return at 532 nm. This search for the closest points takes place in three dimensions. The nearest neighbor algorithm then assigns the three intensities at the original 1064 nm location. This has the effect of reducing the point cloud density by two thirds, as only the points with all three intensities populated are retained. While decreasing the density of the point cloud has deleterious effect on the performance of basic geometry driven classification tools, it has no effect on the potential spectral classification techniques as all the spectral information is catalogued in the remaining points. The result is a false color point cloud in which each point contains a spectral return from each of the three laser beams. Acknowledging that this introduces a minute discrepancy between where the green and IR points originally were recorded, and where they exist in the analyzed point cloud, it is considered worth the analytical flexibility.

Approximately a third of the data set consisted of water, which is analytically insignificant and the majority of which was discarded. Additionally, a significant number of noise points were manually edited out. Noise was identified both above and below the realistic point cloud data. The resultant data set can be viewed in Quick Terrain Modeler (QTM), a program designed to handle three-dimensional visualization of LiDAR point clouds as shown below.

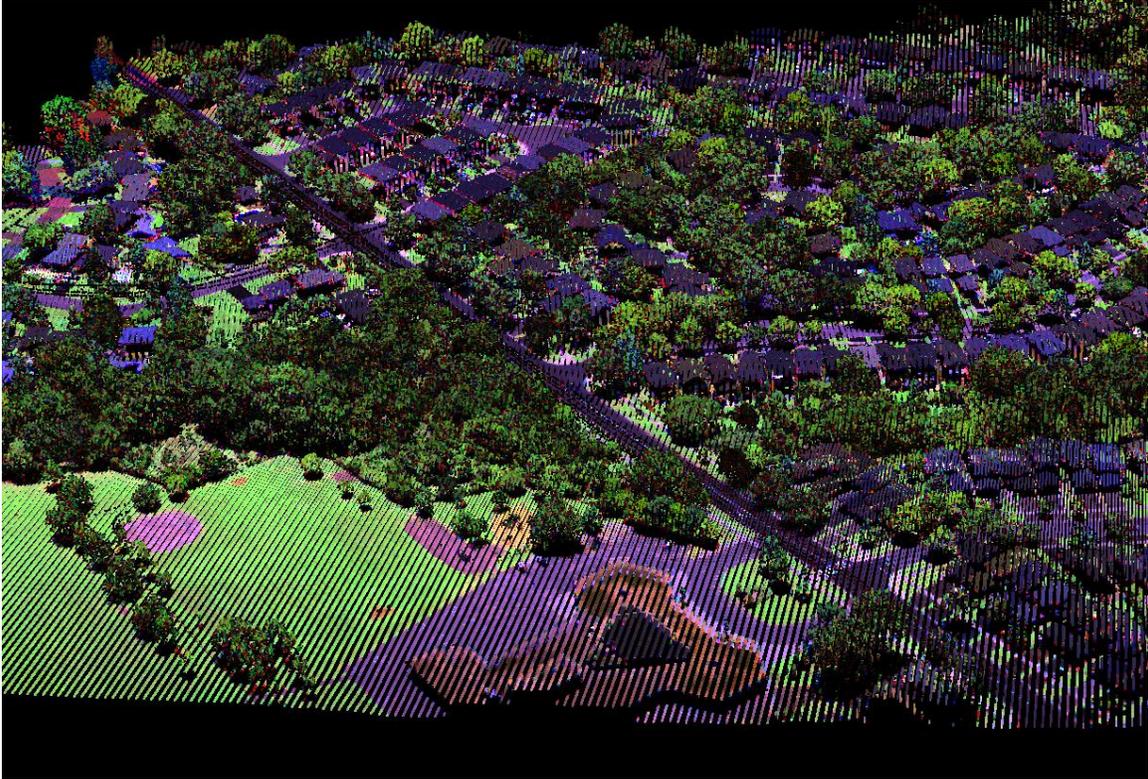


Figure 12. False color image of the Optech Titan data set in Quick Terrain Modeler.

2. Developing a Standard Classification Model

Before analyzing how the spectral components of the LiDAR point cloud improved existing classification techniques, the existing techniques were explored to determine their strengths and weaknesses. A software suite called LAStools, developed by Martin Isenburg, was used to classify the data set to represent the best available automatic solution. The process consists of two steps. First, the LASground program is used to classify points as either ground or unclassified. LASground uses the geometry of the points in a given area to determine where the ground level exists, and calculates the height above ground for each point. Height above ground will also be referred to as above ground level, or AGL. The AGL calculation is optional, and is generally useful in the creation of DTMs. In this case AGL was calculated and stored for use in later analysis. If there are multiple returns from a given laser pulse, the LASground algorithm will only consider the lowest or “last return” for possible classification as ground. Once the data set has been classified as either ground or other, and

the height of each point assigned, the data set is fed into the LASclassify program. LASclassify looks at points above a tunable threshold height, and evaluates them against their neighboring points as either planar, or rugged. Planar points are classified as buildings, and rugged points are classified as vegetation. The result of this process is a data set consisting of four values shown in the table.

Classification Number	Classification Definition	Default Classification Color
1	Unclassified	Red
2	Ground	Green
5	Vegetation	Dark Blue
6	Building	Light Blue

Table 2. List of classes produced by LASground and LASclassify software.

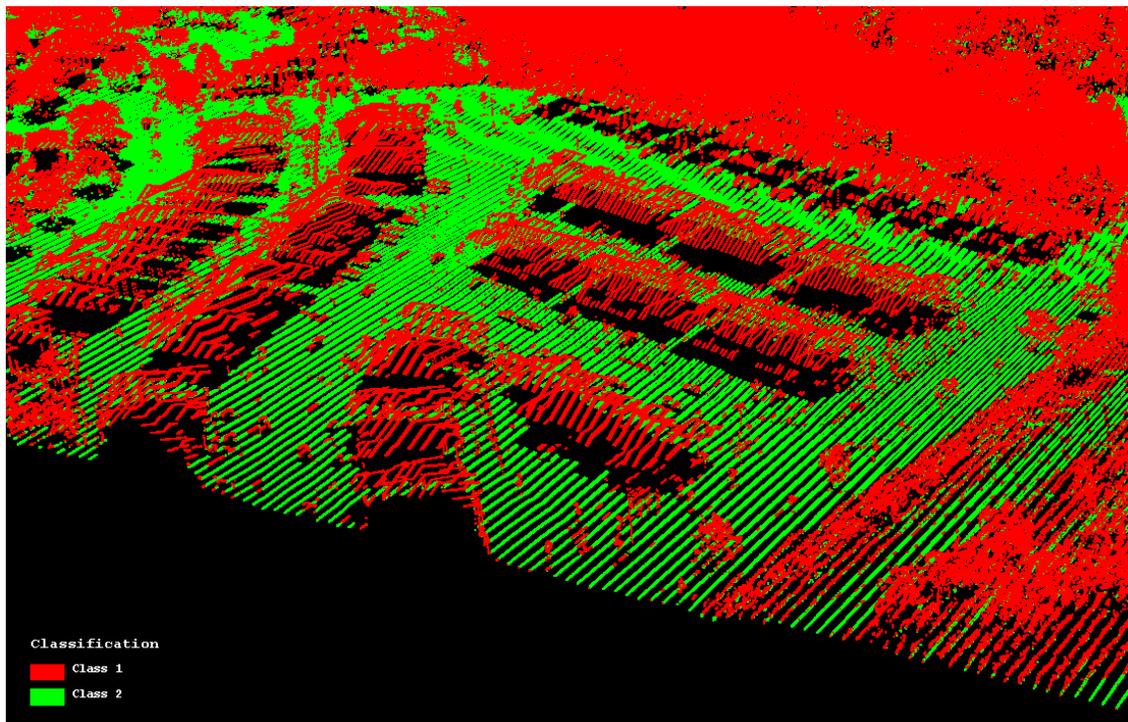


Figure 13. Results of LASground program separating points into “ground” and “non-ground” shown in QTM.

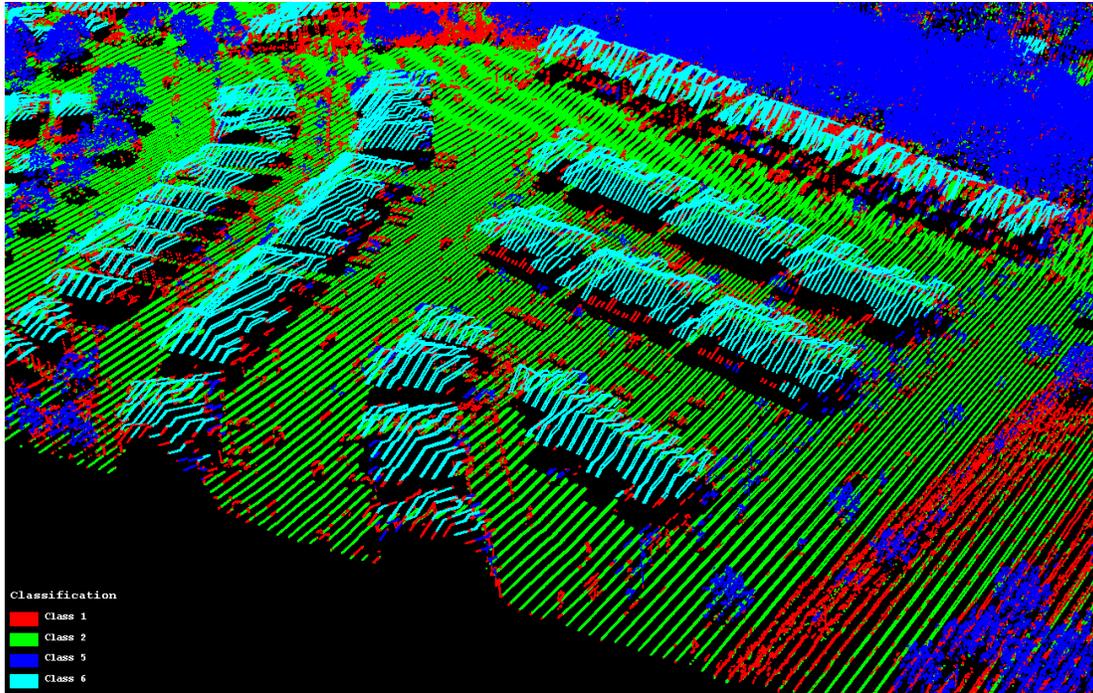


Figure 14. Results of LASclassify program separating points into unclassified, ground, vegetation, and buildings shown in QTM.

While the LASground and LASclassify programs provide a reasonable amount of classification fidelity for automated functions, they still leave much to be desired in terms of creating an operationally valuable environmental picture. The primary shortfall lies in the limited number of discernable classes when using point geometry only. LASground cannot tell the difference between ground and water, for instance, if they both appear relatively flat. Similarly, a field that is grass on one end, and sand on the other appears the same to most automated classification tools. Secondly, automated tools which rely on geometry only struggle to classify a significant number of points. In the case of LASclassify, if the points do not meet metrics to be considered vegetation or buildings they go unclassified.

Classification No.	Classification Type	No. of Points	% of Total
1	Unclassified	1,838,670	16.2
2	Ground	3,224,968	28.5
5	Vegetation	5,353,407	47.2
6	Buildings	914,934	8.1
		11,331,979	100

Table 3. Distribution of point classification from LASclassify program.

3. Manual Classification

In order to provide increased fidelity to the follow on spectral classification techniques, the results of the automatic classification tool, LASclassify, required manual correction. Manual classification of a LiDAR point cloud is time intensive, and requires imagery in order to prove the ground truth. For this reason, three subset regions were selected in order to be developed into training sets for later analysis. The regions were chosen to represent some of the unique features of the Optech Titan data collection and are shown below.

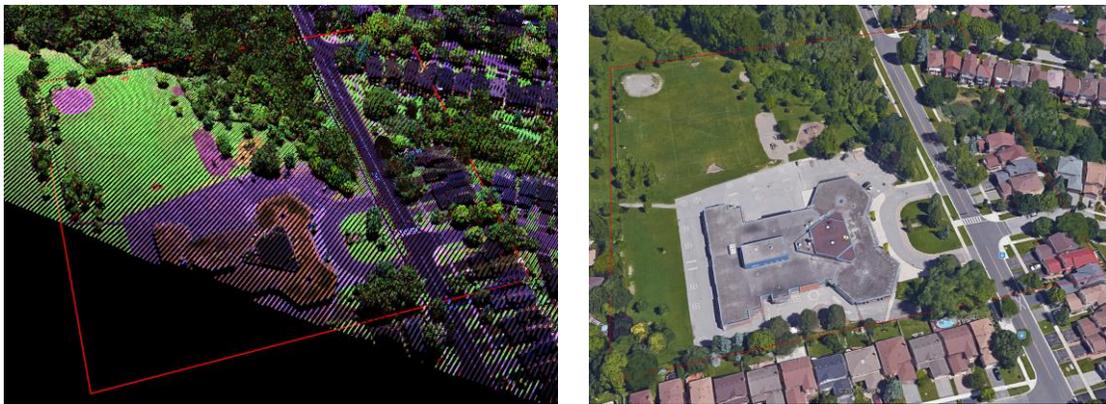


Figure 15. Training Region #1 – Middle School shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.



Figure 16. Training Region #2 – Wooded Power Station/Antenna shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.

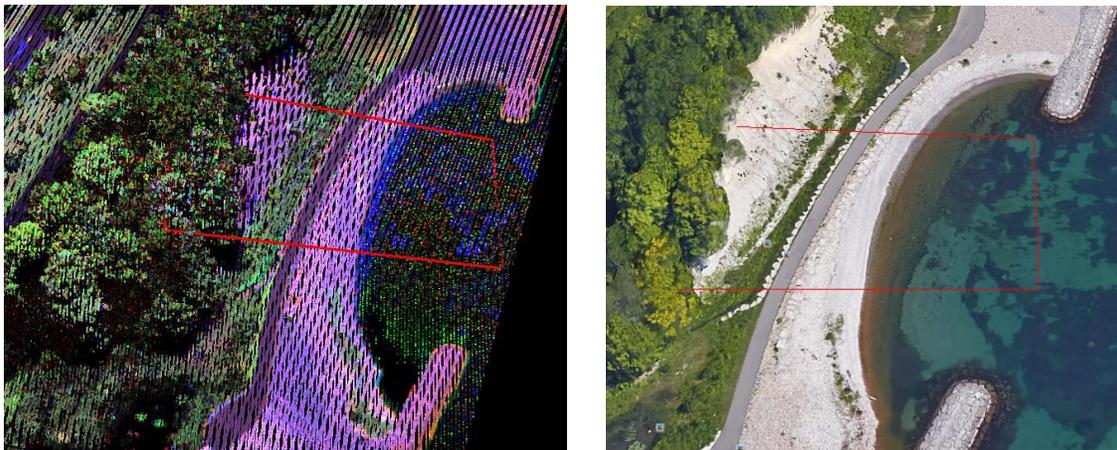


Figure 17. Training Region #3 – Beachfront shown in the false color LiDAR point cloud and the corresponding Google Earth imagery.

During the process of manual classification, the goal was to eliminate unclassified and misclassified points by placing them into the most appropriate classification. To help achieve this end, three additional classifications were implemented: water, power lines, and a miscellaneous man-made classification. One clear example of how geometrically driven classifiers can fail is illustrated in Figure 18. The roof of the school building (bottom right of Figure 18) pictured is so large and flat it is mistakenly classified as ground (Class 2). Figure 19 shows the manually classified result.

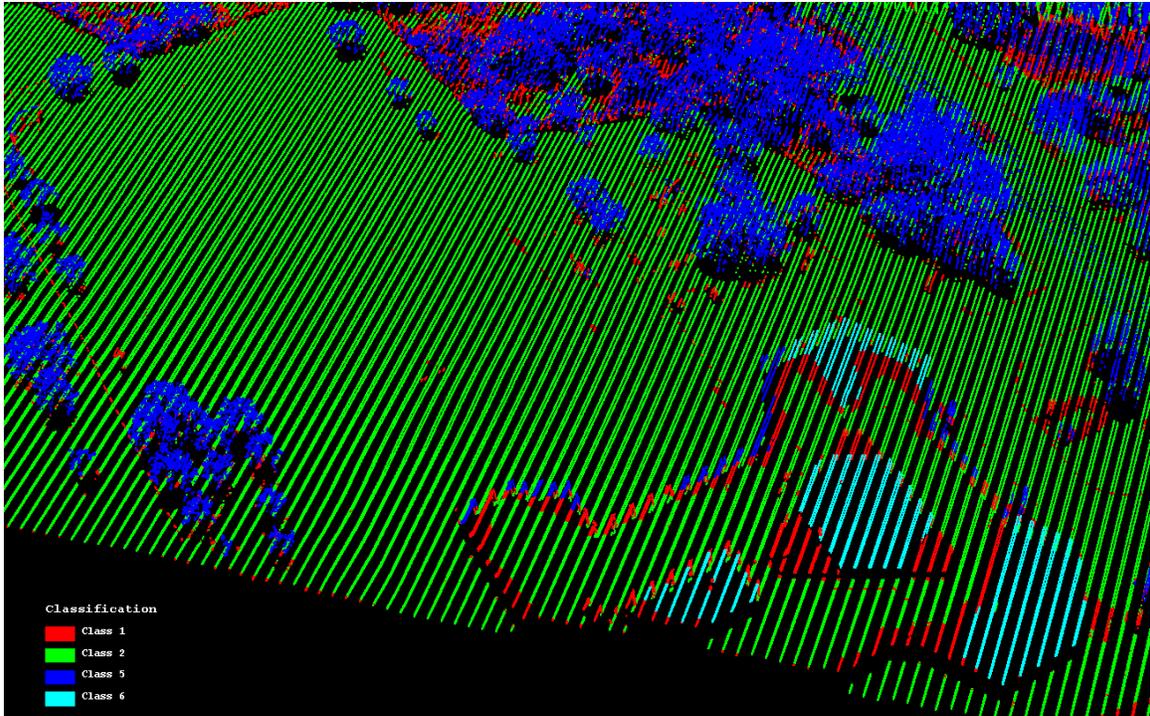


Figure 18. Training Region #1 - Raw result from LASclassify showing both classification errors and unclassified areas.

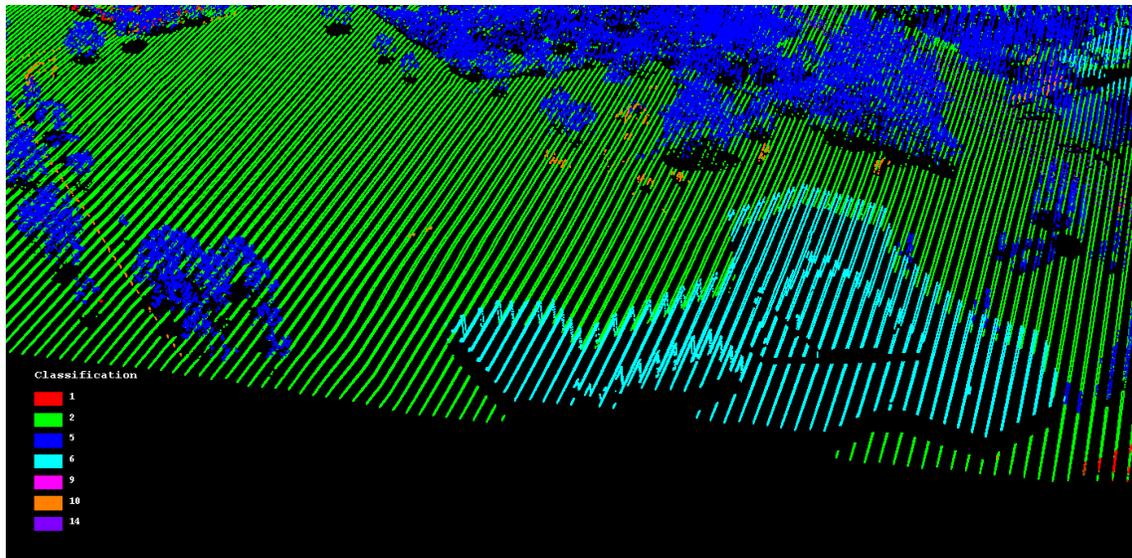


Figure 19. Training Region #1 - Manually classified data set.

4. Flightline Merging and Flagging

Several steps were taken at the end of the data preparation phase to ease the analysis phase. Up until this point in the workflow, preparation was being performed on three files representing the three flightlines of data supplied by Optech. While less cumbersome than the nine original files (three flightlines by three wavelengths), one file is more ideal. A LAStools program named LASmerge combined the three flightlines into one file totaling over 11 million points.

Points in the training regions developed during the manual classification phase needed to be easily identified and segregated during follow on analysis. For this purpose the training regions were turned into shape files and used to flag the points inside the regions. This was completed in a software called ArcGIS, which also allowed the flagged points to be populated in the “Keypoint” attribute field of a standard LAS file. The Keypoint attribute is sometimes used in the development of DTMs when file sizes are too large. Since DTMs were not created as part of this process, the Keypoint attribute was not needed for spectral analysis and could be overwritten.

As an alternative to the training region method, a random subset was implemented as a means to manage the number of points being included in the spectral analysis. The data set was reduced by a factor of twenty through a random sampling code. Subsequently the randomly sampled points were flagged in the Keypoint attribute field for later identification.

IV. DATA ANALYSIS

A. ANALYSIS WORKFLOW

The utilization of multi-wavelength LiDAR intensity data for terrain classification has not been rigorously explored. This means that tools to conduct such an analysis have to either be derived or adapted from pre-existing tools. A workflow was developed in order to manipulate existing software for the purpose of conducting spectral analysis on a multi-wavelength LiDAR point cloud. Three programs were primarily utilized in this workflow, LAStools, Quick Terrain Modeler (QTM), and ENVI.

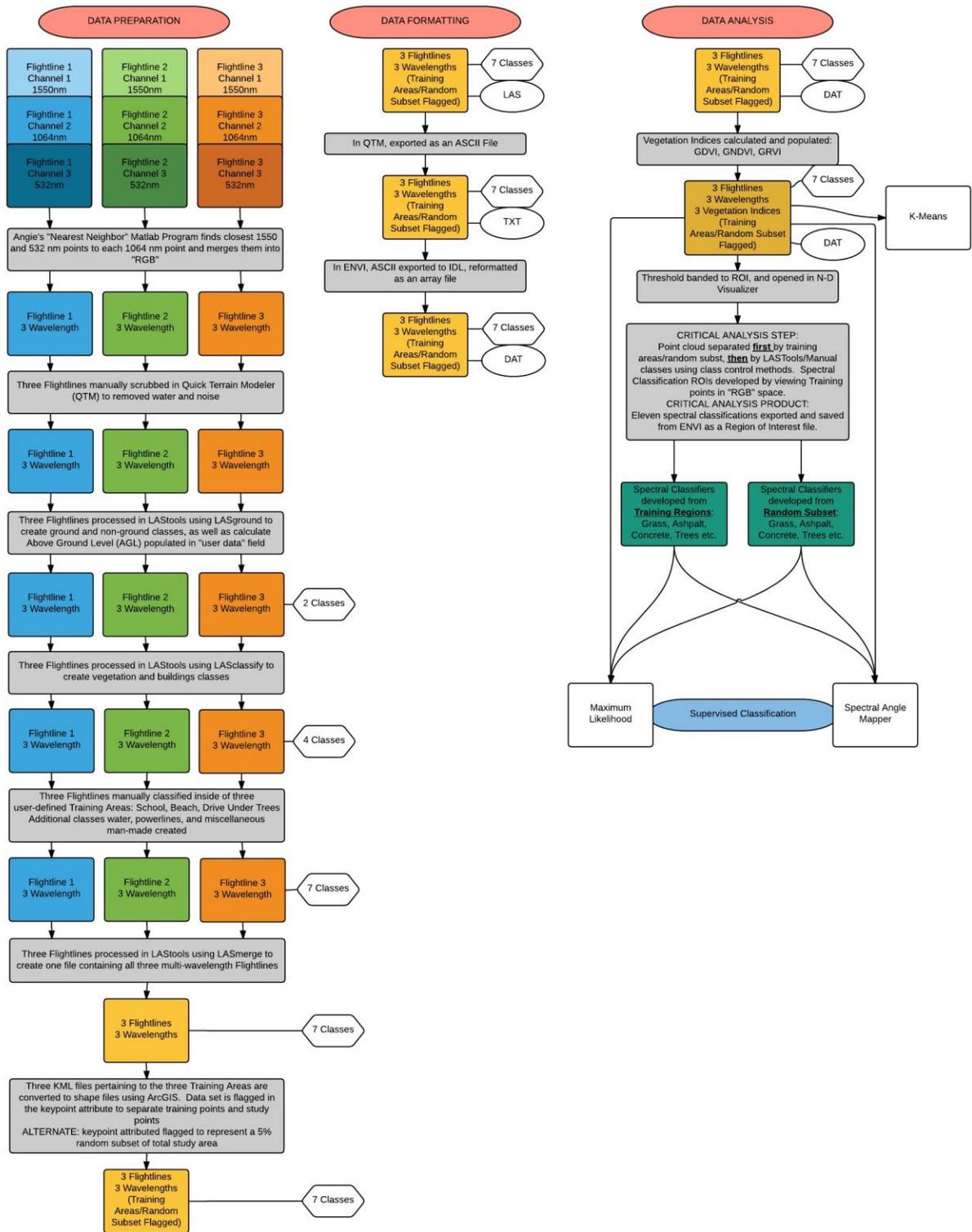


Figure 20. Analysis workflow for spectral classification of multi-wavelength LiDAR.

B. VEGETATION INDICES

One of the proven analytical indices employed in spectral analysis is the Normalized Difference Vegetation Index (NDVI). As discussed in Chapter II, without a red channel contributing to the spectral return, a true NDVI is not attainable. Instead, a green NDVI or GNDVI was employed in the case of the Optech Titan data. Along with GNDVI, two additional indices were calculated for inclusion in the spectral analysis. These indices excel at segregating data based on the absence or presence of vegetation, but they do experience limitations performing deeper in vegetation canopies where intensity values may vary.

Index	Formula
Green Normalized Difference Vegetation Index (GNDVI)	$GNDVI = (NIR - Green)/(NIR + Green)$
Green Difference Vegetation Index (GDVI)	$GDVI = NIR - Green$
Green Ratio Vegetation Index (GRVI)	$GRVI = NIR/Green$

C. UTILIZING MULTI-SPECTRAL LIDAR INTENSITIES TO DERIVE SUB-CLASSIFICATIONS

1. Generating Spectral Training Sets

Having manually classified three training regions, the data set next required spectral analysis to discern potential sub-classifications. The primary tool used for extracting sub-classifications from the spectral intensity data was the “n-Dimensional Visualizer” (N-D VIS) analysis tool in the ENVI software suite. Although ENVI does publish a separate LiDAR analysis, ENVI LiDAR, the N-D VIS tool in the suite designed for imagery processing provided the analysis tool best suited for spectral evaluation. For this reason the LiDAR point cloud was manipulated in the ENVI programming interface, IDL, to resemble a structured array rather than a point cloud. LiDAR point clouds are typically formatted as an LAS or .las file, which consists of twenty attributes like

Keypoint mentioned previously. In this case the LAS files were merged and converted into an ASCII format, which is convertible in IDL to be properly accepted by N-D VIS. Using regions of interest (ROI) to ease the burden of display time, the ASCII file was converted into an array which subsequently can be imported in N-D VIS.

N-D VIS provides a tool exactly as its name suggests. In the case of the reformatted LiDAR point cloud, this means “n” dimensions can be visualized at a time and viewed in a staggering number of combinations. By increasing dimensions input to the visualizer, its utility becomes apparent. First, training region points were visualized and separated using the Keypoint flag as a dimension. By working within the manually classified training regions, the classifications were treated as reliable. Secondly, the points were separated by geometrically derived class, predominantly the ground and building classes. Working within the ground class, inside the training regions, spectral classification could be conducted.

The spectral analysis took place by observing the three spectral components in three dimensional space, or “spectral space” displayed in Figure 21. Similar to a 3D scatter plot, the visualizer animates the plot through various orientations. By carefully identifying where the three spectral components grouped or “clumped” spectral classes which were not directly tied to geometry could be derived. The process consists of first recognizing the groupings, selecting them and assigning them a new class, and then trimming the new class to eliminate outliers. The result of this process is shown in Figure 22.

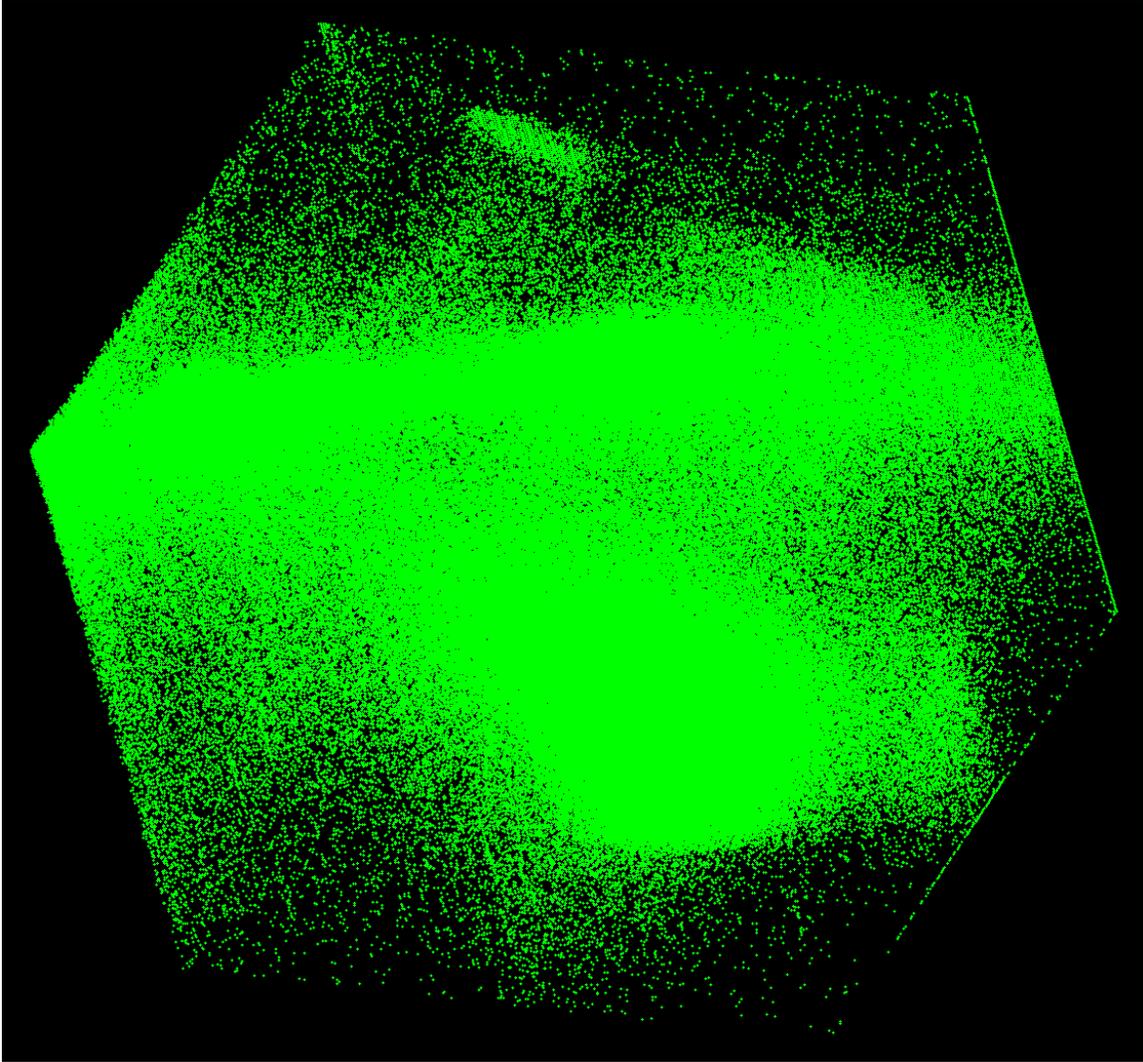


Figure 21. Ground points within the training regions viewed in spectral space in N-D VIS, prior to spectral classes being identified.

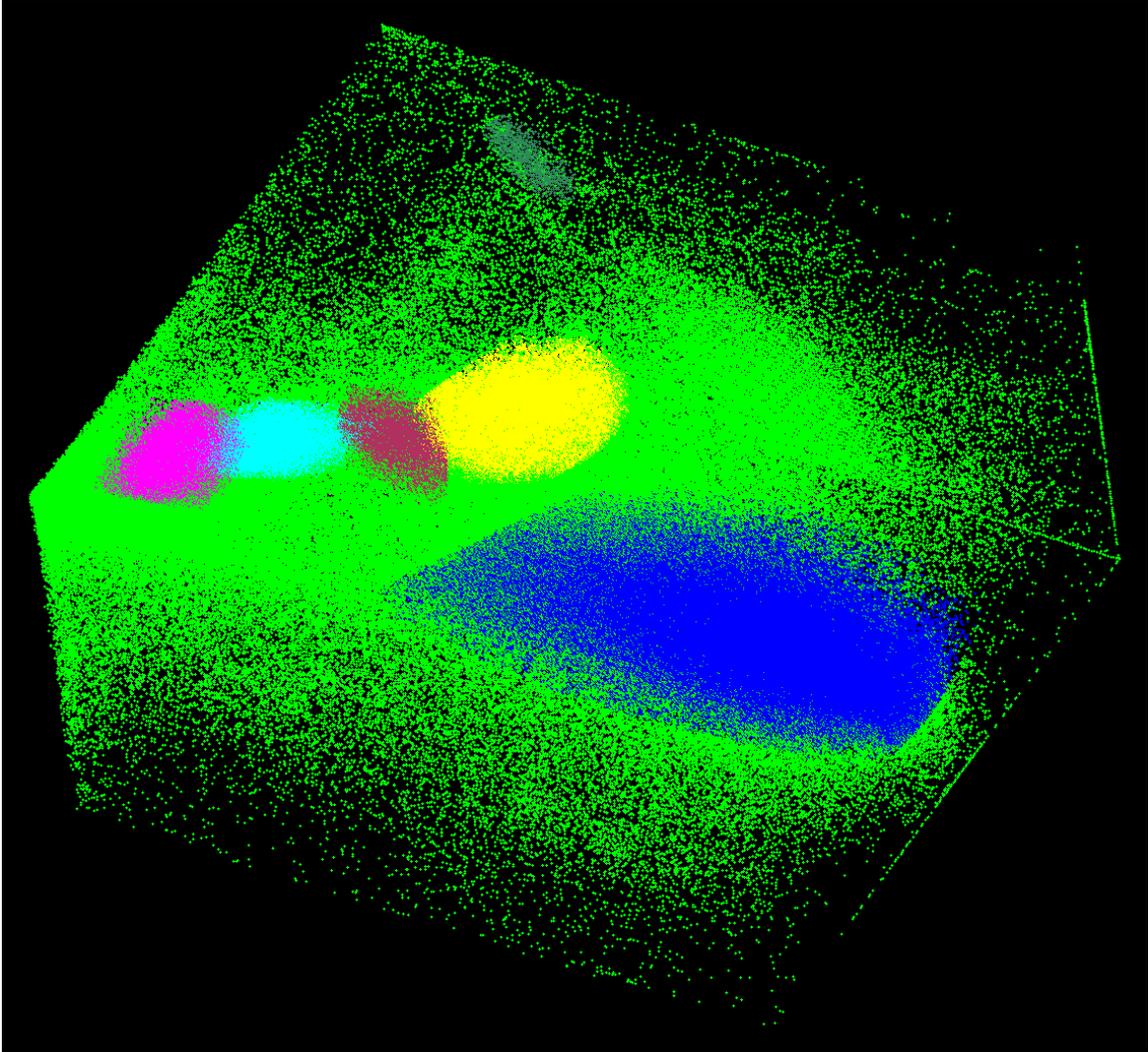


Figure 22. Ground points in training regions viewed in spectral space in N-D VIS, with six spectral classes having been identified and reassigned as other classes.

Throughout the process of spectral classification, the classes were validated by viewing them in the X, Y space, or a two-dimensional view. Comparing this two-dimensional view to the original false color point cloud, and Google Earth imagery, spectral classes were kept or discarded.

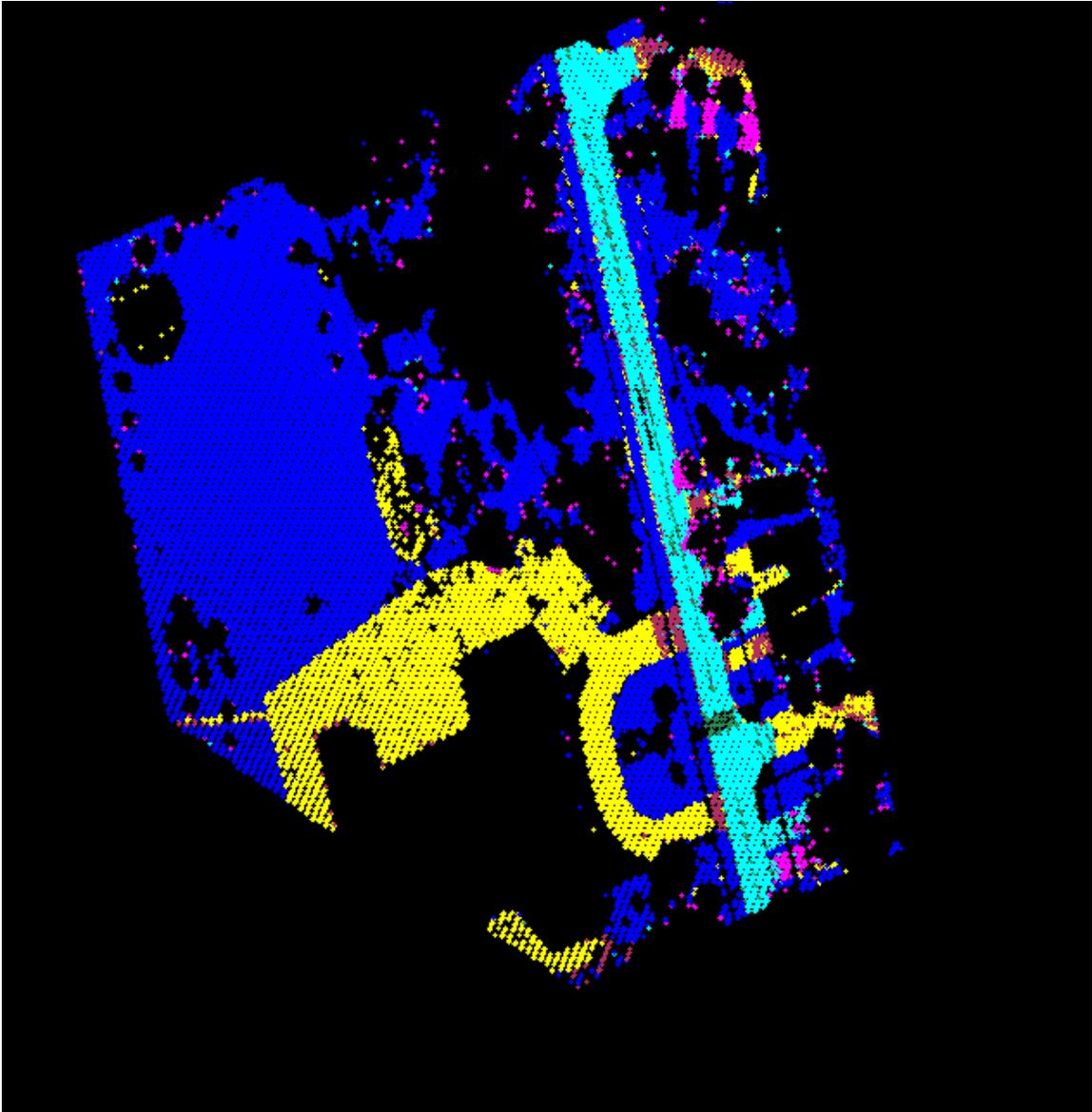


Figure 23. Ground points in Training Region #1 viewed in "X,Y" space in N-D VIS.



Figure 24. Google Earth imagery of Training Region #1.

As shown in Figure 23, the spectral classes have an increased level of fidelity and detail over the classifications derived by geometry alone. In this case all the points in Figure 23 would have been geometrically classified as “ground” at best. Applying the spectral classification technique pulls out areas that are covered in grass, asphalt, concrete, etc. Furthermore close comparison between Figure 23 and Figure 24 reveals that *different types* of asphalt have been separated and classified differently (highlighted in yellow, maroon and light blue in Figure 23).

The useful output of the spectral classification process in N-D VIS, is a set of classes grouped in Regions of Interest (ROI). ROIs can be saved and applied to supervised classification tools in the ENVI software suite. In the case of this analysis, eleven spectral classes were extracted. The classes represent grass, three types of asphalt, trees, railroad, four building materials, and paint used on streets.

Spectral Class Number	Previous Class	Sub-Class	Unique Features	Spectral Plot Color
1	Ground (2)	Grass		White
2	Ground (2)	Pavement 1	School Parking Lot and Substation Access Road	Red
3	Ground (2)	Pavement 2	Side streets and Lake Footpath	Light Green
4	Ground (2)	Pavement 3	Main thoroughfares	Blue
5	Ground (2)	Railroad Tracks		Yellow
6	Vegetation (5)	Vegetation		Light Blue
7	Building (6)	Roof 1	Trapezoidal Segment of School Roof	Magenta
8	Building (6)	Roof 2		Maroon
9	Building (6)	Roof 3		Dark Green
10	Building (6)	Roof 4	School Roof	Purple
11	Ground (2)	Road Paint	Cross walks	Coral

Table 4. Listing of Spectral Classes generated in N-D VIS using training regions.

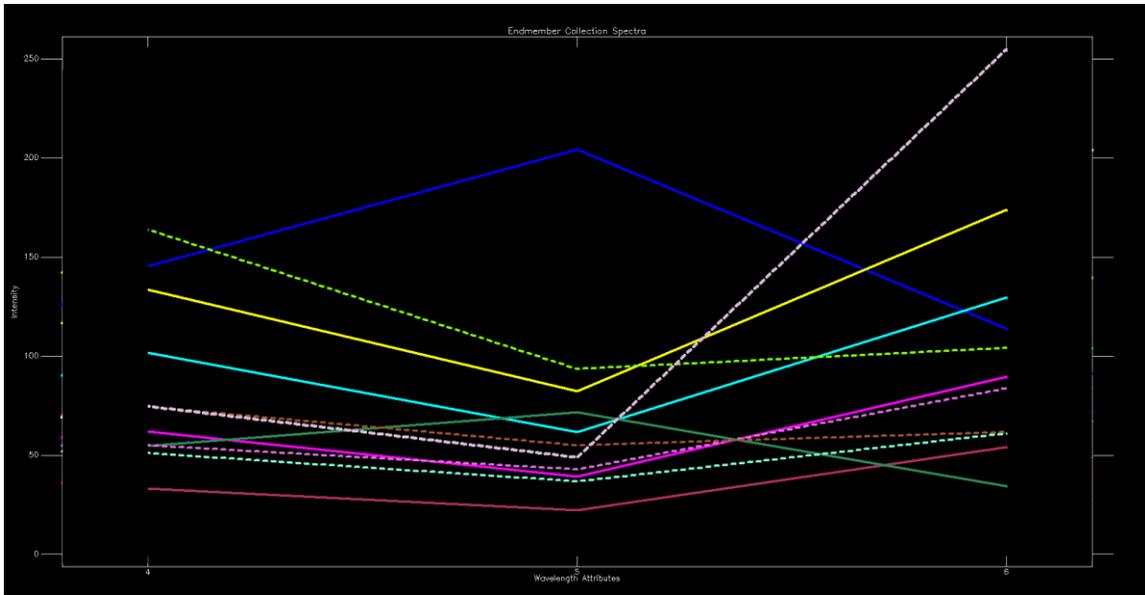


Figure 25. Spectral Library Plot of classes developed in N-D VIS.

The classes developed in N-D VIS are shown in Figure 25 to show the spectral similarities and differences between the 11 classes. The x-axis is numbered by band, and in this case bands 4, 5, and 6 correspond to 1550 nm, 1064 nm, and 532 nm respectively. The y-axis is scaled from 0 to 255 to encompass the eight bit values of the respective intensities. It is worth noting that some of the spectra are very similar, for example the Pavement 3 (blue) and Roof 2 (maroon) classes plot similarly. This is an example of why filtering the points by geometric classification prior to spectral grouping was valuable. Had the points not been filtered, it would have been difficult to tell the two classes apart in spectral space.

2. Alternate Method for Training Set Derivation

As an alternative to conducting spectral classification in the training regions, the full data set was reduced through random sampling and evaluated in N-D VIS for spectral classes. Using a reduction factor of twenty adequately downsized the LiDAR point cloud so as to allow spectral groupings to be recognized in N-D VIS. On the other hand it was not so great a reduction as to prevent clear spatial feature recognition when examining the spectral groupings. This provides a noteworthy advantage to the analyst. Using a random subset ensures that all largely populated spectral classes are represented in the N-D VIS

process. This ensures any robust spectral groupings which are spatially outside of the training regions are not missed. An example of a class which is missed by the training region method is the roof classification which is noticeably lacking as shown in Figure 39. This class was strongly represented in the randomly sampled data set. There are two disadvantages to working in a random subset of the entire study area. First, since the manual classification to correct the geometric classification only took place in the training regions, the spectral classification is reliant on the imperfect results of the LASclassify program for filtering. Secondly, because the data is reduced by a factor of twenty, and the remaining points are spread across the entire study area, some smaller spectral classes which are adequately represented in the training regions are not discernable. Examples of classes which are missed in the due to under-sampling are the footpath pavement along the lake, and the paint used to mark roads.

Spectral Class Number	Previous Class	Sub-Class	Unique Features
1	Ground (2)	Grass	
2	Ground (2)	Pavement 1	School Parking Lot and Substation Access Road
3	Ground (2)	Pavement 2	Main thoroughfares
4	Ground (2)	Railroad Tracks	
5	Vegetation (5)	Vegetation	
6	Building (6)	Roof 1	Neighborhood unaccounted for by training regions
7	Building (6)	Roof 2	Trapezoidal Segment of School Roof
8	Building (6)	Roof 3	
9	Building (6)	Roof 4	

Table 5. Listing of Spectral Classes generated in N-D VIS using random sampling.

V. RESULTS

A. MAXIMUM LIKELIHOOD

Applying the output of the spectral classification process to the entire study area was conducted using supervised classification tools in ENVI. The first of these tools was the Maximum Likelihood (ML) classifier. The ML classification tool starts by calculating statistics for each region of interest applied in the tool. Once calculated, the program evaluates the likelihood that a given pixel or point is part of a given class, which are assumed to be normally distributed. If he or she chooses, the analyst may select a probability threshold, which will result in points remaining unclassified if they do not exceed the threshold of likeliness in any class. If no threshold is specified, all points will be assigned a class based on their highest probability.³⁷ The primary thresholds used in this analysis were high, medium, and none, corresponding to settings of 0.85, 0.5, and none. Probability is calculated based on the following discriminant function³⁸:

$$g_i(x) = \ln p(\omega_i) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (x - m_i)^T \Sigma_i^{-1} (x - m_i)$$

Where:

$g_i(x)$ = the probability a data point, x, falls in a given class i

i = class

x = n-dimensional data (n is number of bands)

$p(\omega_i)$ = probability that class ω_i occurs in the image and is assumed the same for all classes

$|\Sigma_i|$ = determinant of the covariance matrix of the data in class ω_i

Σ_i^{-1} = its inverse matrix

m_i = mean vector

³⁷ "Maximum Likelihood," Excelis Visual Information Systems, n.d., accessed online at <http://www.exelisvis.com/docs/MaximumLikelihood.html> on 17 Aug. 2015.

³⁸ John A. Richards, *Remote Sensing Digital Image Analysis: An Introduction*, (Berlin, Germany: Springer, 2006), 197.

Optimizing the effectiveness of the ML tool requires the analyst to appropriately select the bands, or attributes in the LiDAR point cloud's case, for inclusion in the process. For all runs of the ML tool in this analysis the following bands were selected: 1550 nm intensity, 1064 nm intensity, 532 nm intensity, above ground level (AGL), and GNDVI. It is important to note that while the purpose of this study is to investigate what the spectral aspects of multi-wavelength LiDAR add to the terrain classification process, it is not intended to do so absent of the other data collected by the sensor. With that in mind, AGL was included in the ML classifier to as an aid to the spectral classifier. Including AGL means the classification results were not derived by spectral means entirely, but are primarily spectrally divided. The probability threshold was varied to populate several results from which the best could be selected.

The two methods of deriving spectral classes described in Chapter IV resulted in two different training sets. Both sets were analyzed by the ML classifier separately and their results were compared and contrasted. While many classes derived from training regions and the random subset method appeared similar, their performance as training sets had some marked differences.

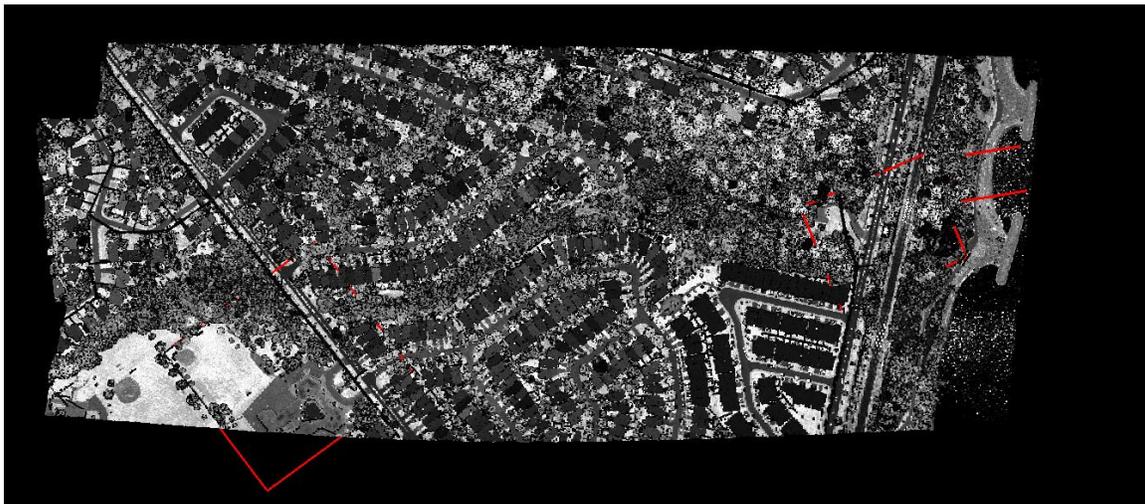


Figure 26. Study area with training regions boxed in red.

1. Spectral Classes Derived From Training Areas

In general, when the probability threshold was omitted or a low value, the classifier was less discriminant, leaving no or few points unclassified respectively. This resulted in classes which were discernable, but heavily interspersed with errors of commission. Conversely, as the probability threshold approached one, the classifier is more discriminant, and in the case of the threshold being set at 0.85, left 10.2 million out of 11.3 million points unclassified. The resultant classes, however, are distinct and are characterized by errors of omission when the threshold is high.

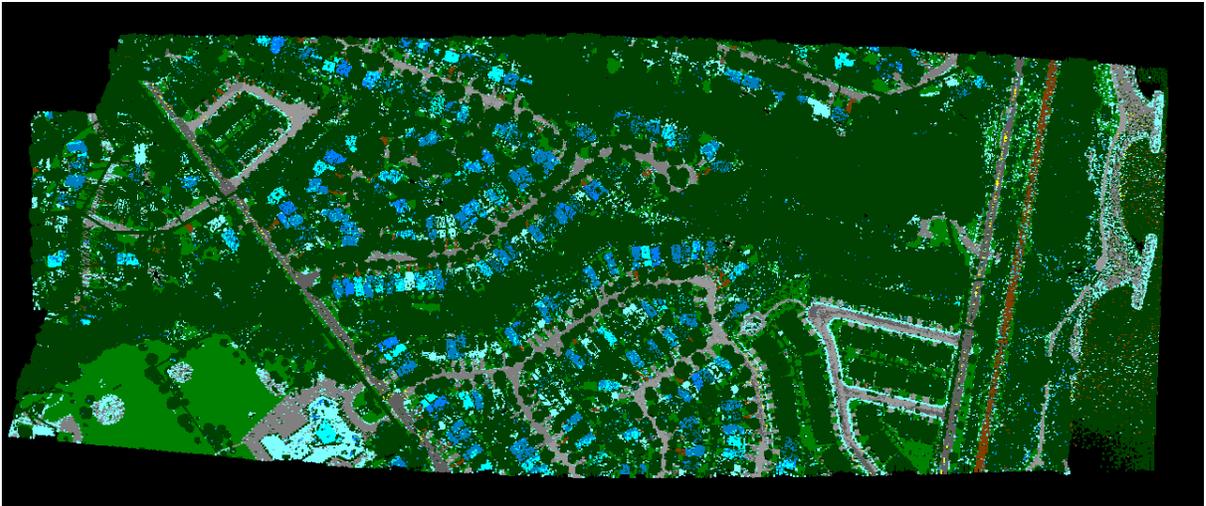


Figure 27. Demonstration of ML Results with no probability threshold showing all 11 classes.

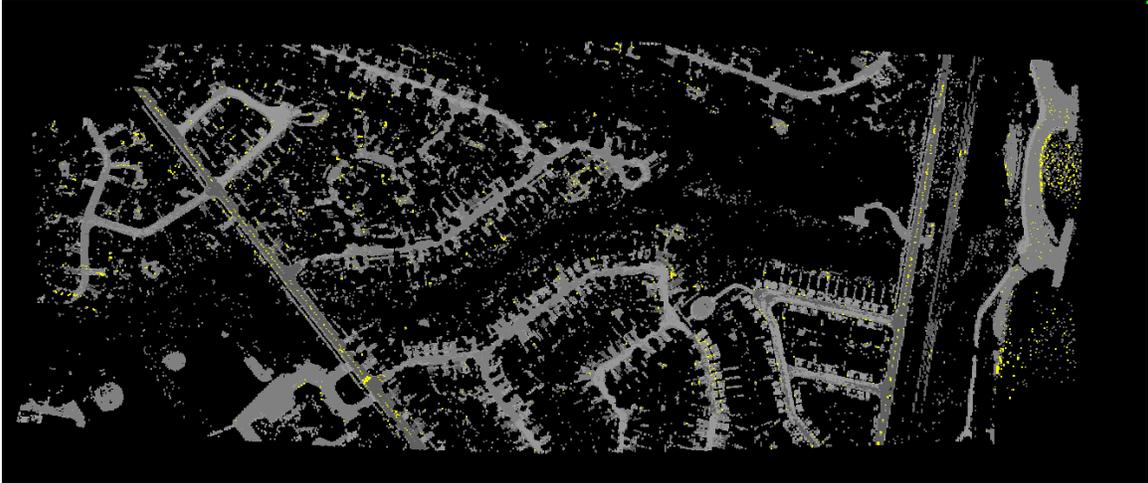


Figure 28. Demonstration of ML Results with no probability threshold showing classes 2–4 and 11, highlighting roads and road paint.

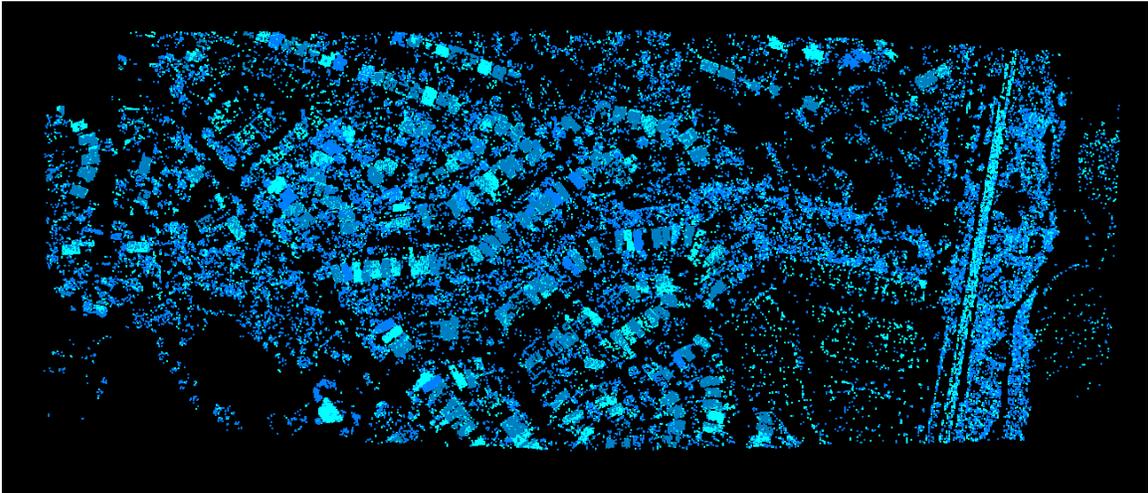


Figure 29. Demonstration of ML Results with no probability threshold showing spectral classes 7–9, highlighting buildings.

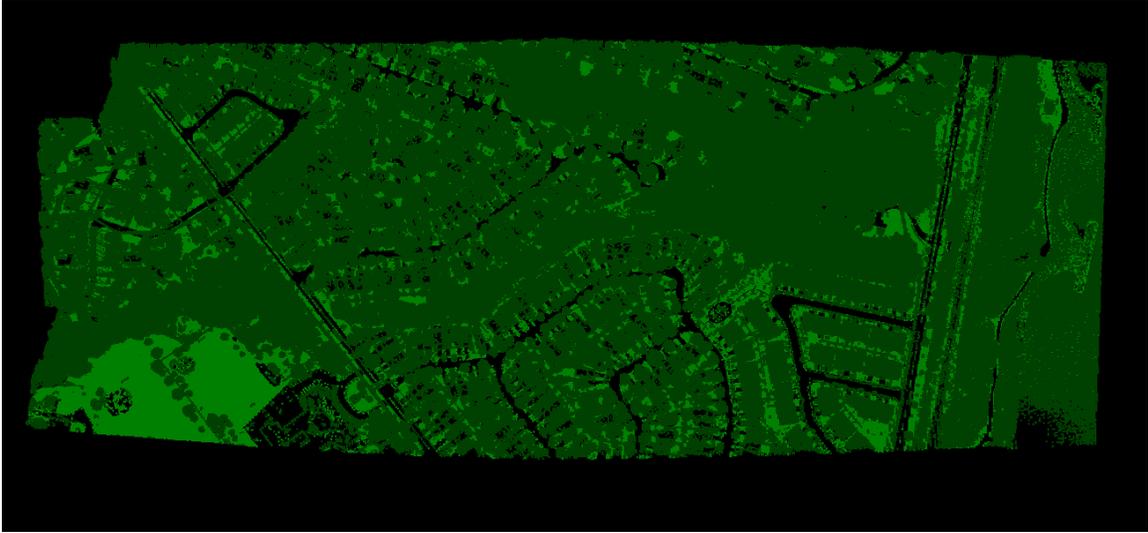


Figure 30. Demonstration of ML Classifier results with no probability threshold showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.

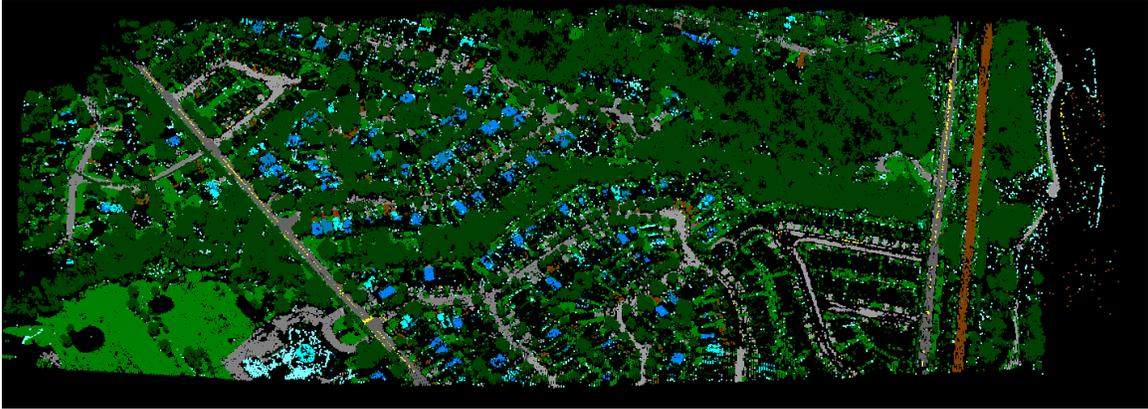


Figure 31. Demonstration of ML Classifier results with a high probability threshold (0.85) showing all 11 spectral classes.

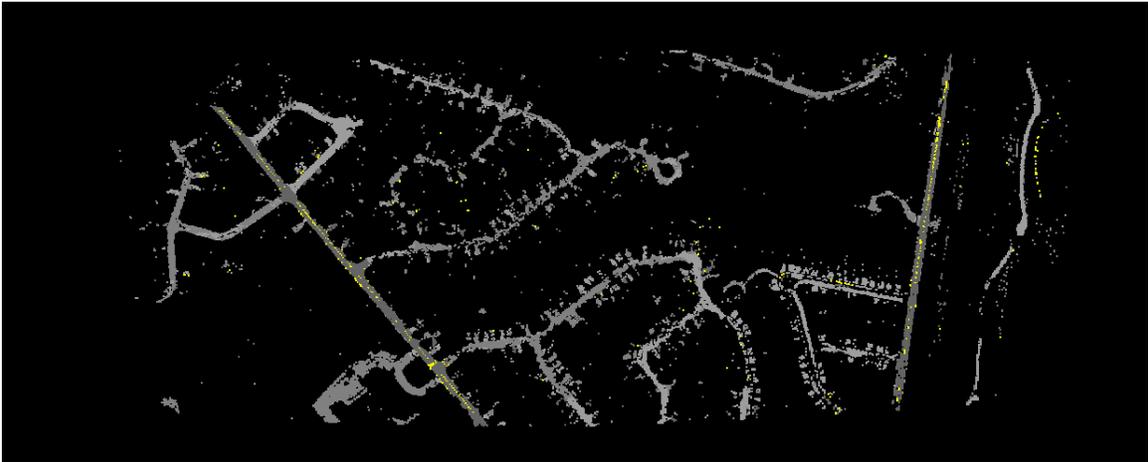


Figure 32. Demonstration of ML Classifier results with a high probability threshold (0.85) showing spectral classes 2–4 and 11, highlighting roads and road paint.

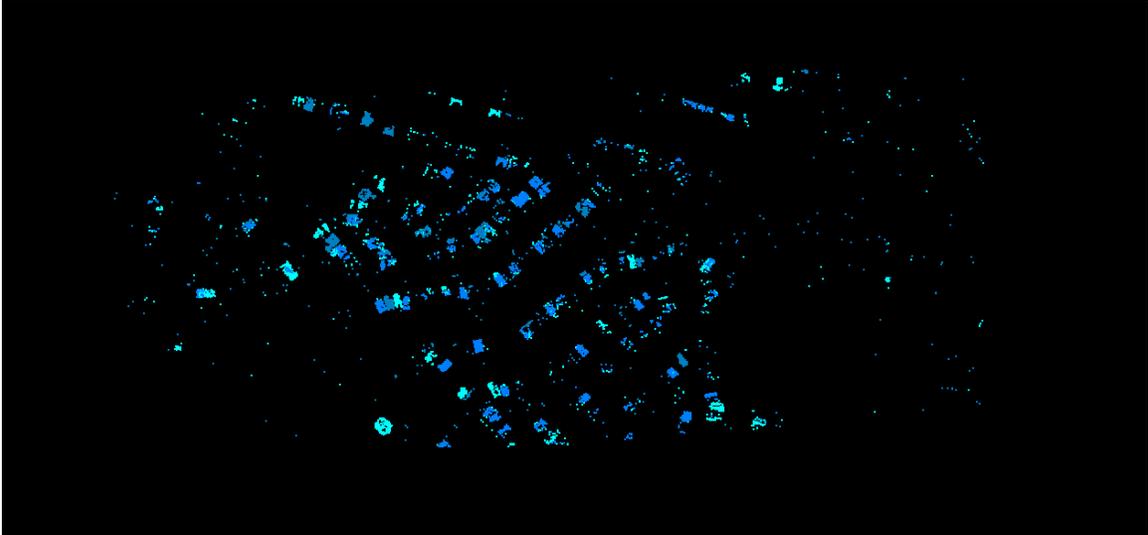


Figure 33. Demonstration of ML Classifier results with a high probability threshold (0.85) showing classes 7–9, highlighting buildings.

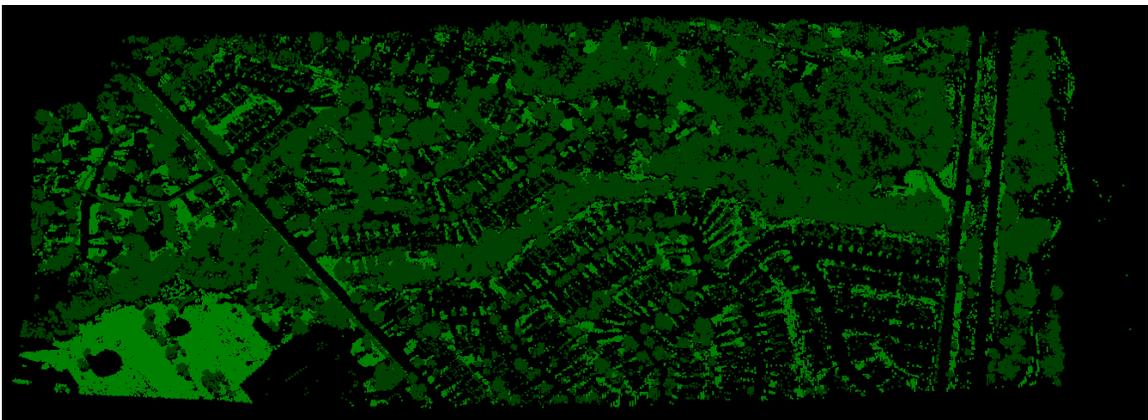


Figure 34. Demonstration of ML Classifier results with a high probability threshold (0.85) showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.

These two cases show that applying the appropriate threshold to the ML classifier is critical to obtaining the desired result from the analysis. Both ends of the spectrum perform well in identifying the paved surfaces. In the case of no threshold, there is discernable noise, or misclassified points, however the boundaries and forms of the roads are still easily identifiable to the casual observer. Absent the noise, the high probability threshold case (Figures 31–34) does not significantly improve the mapping of the paved surfaces. Looking at the building classes, the errors of commission do begin to mask the

nature of the neighborhood structures in the case of no probability threshold. On the other hand the high probability threshold case clearly leaves a significant number of structures unclassified, providing a poor portrayal of the structure density. Unsurprisingly, the optimal performance of the ML tool may be achieved by finding a middle ground. Setting the probability threshold at 0.5 results in 7.4 million unclassified points of the 11.3 million point cloud. The almost 4 million remaining points provide a reasonable portrayal of the study area. Streets and roads are still easily discernable, and building performance leaves fewer structures out without becoming too cluttered to see. One of the main failures of the ML classifier using the spectral classes is the misclassification of the lake and one unique neighborhood as vegetation. Highlighted in Figure 39, neither feature performs well regardless of the probability threshold setting.

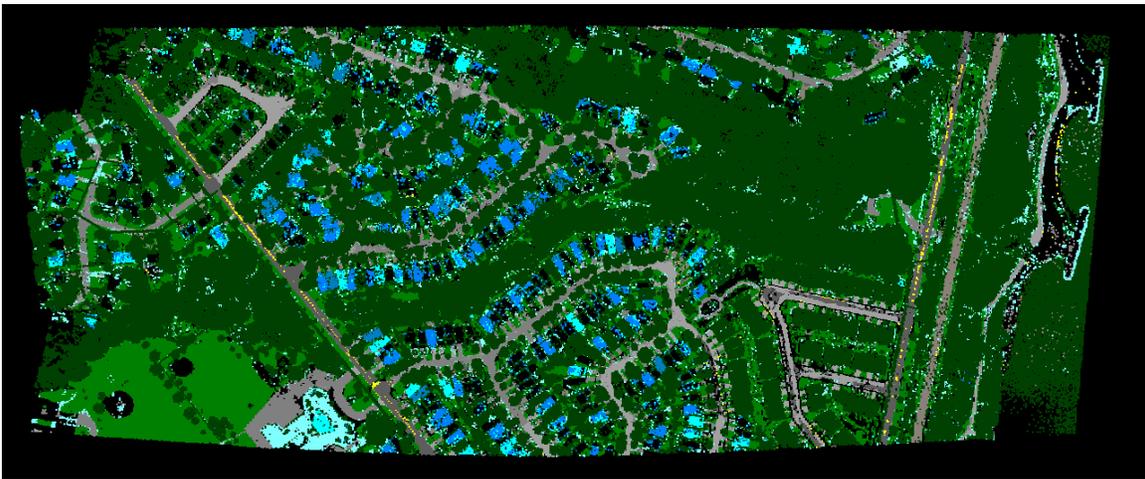


Figure 35. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing all 11 spectral classes.



Figure 36. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 2–5 and 11, highlighting roads, railroads and road paint.

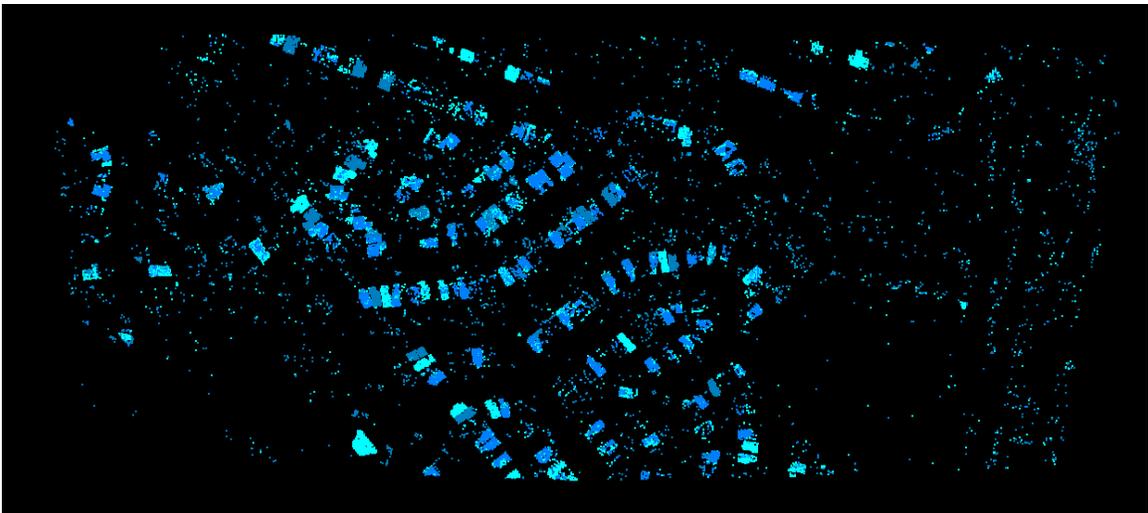


Figure 37. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 7–9, highlighting buildings.

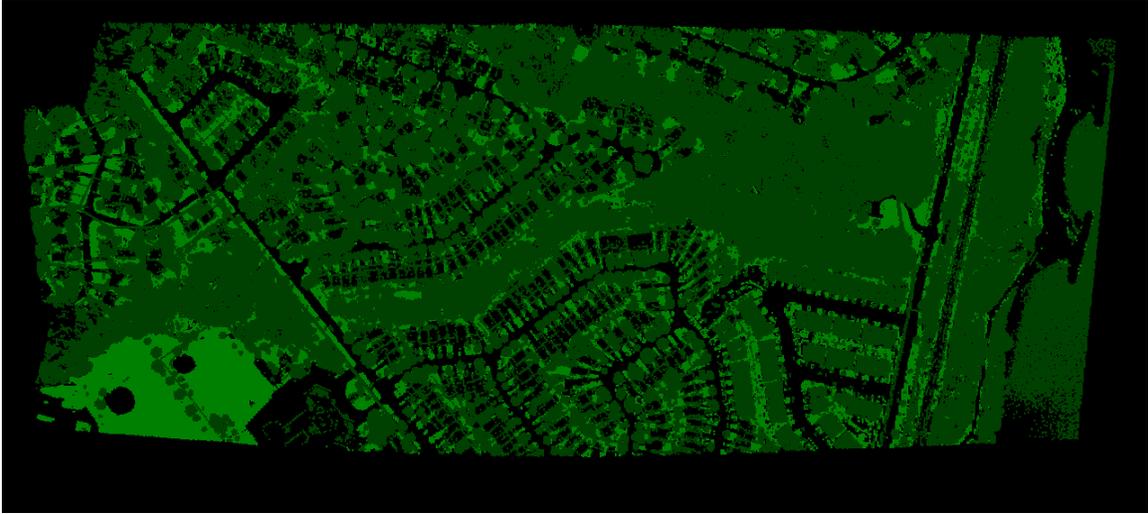


Figure 38. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 1 and 6, highlighting vegetation separated by grass and trees.

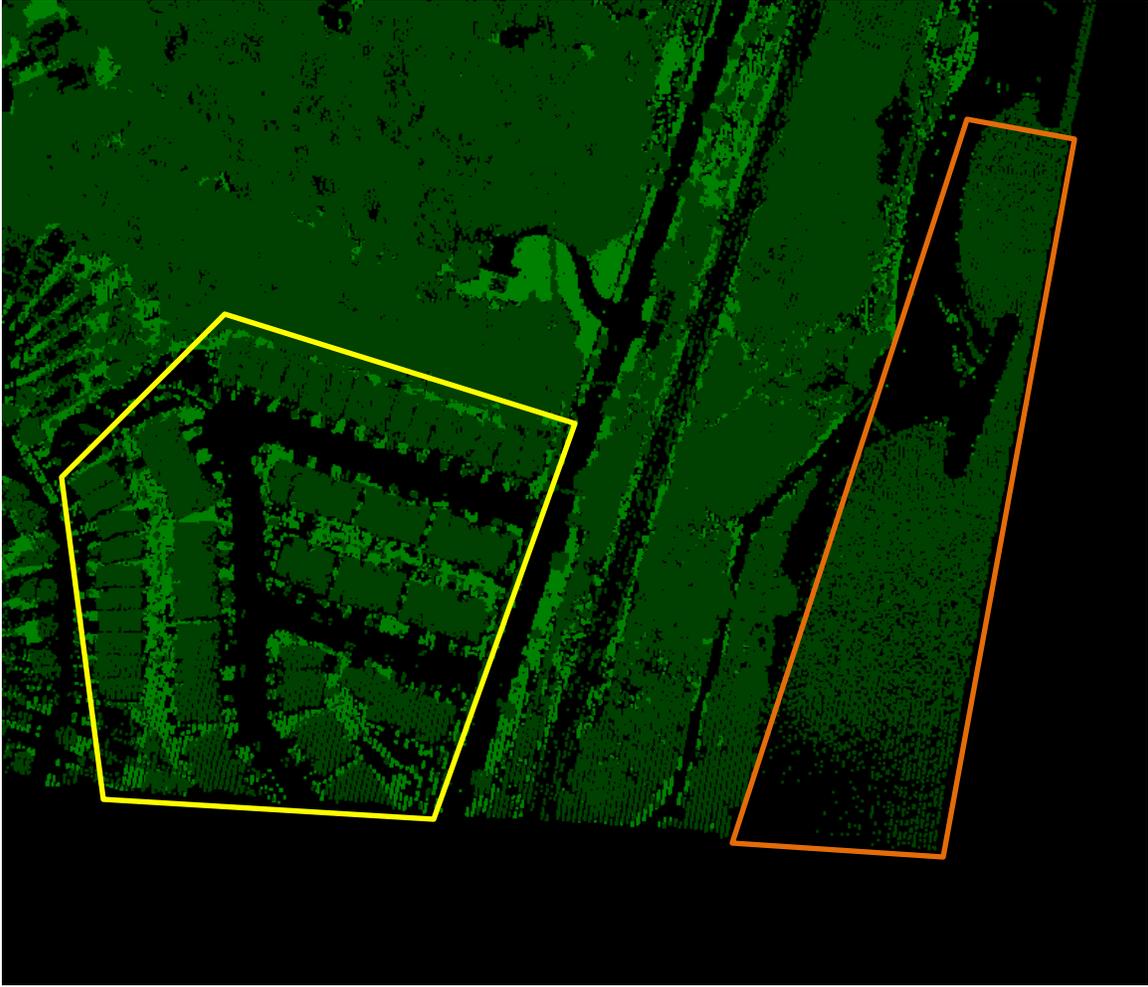


Figure 39. Detail view of spectral classes 1 and 6 which represent vegetation, and errors in the spectral classification process from one neighborhood (yellow) and the lake water (orange).

The poor performance of the spectral classifier in one of these two cases may be directly linked to how the data was reduced in order to support the identification of spectral classes. In the training area method, points outside of the training regions were not evaluated for spectral grouping. This allowed for a less dense cloud in N-D VIS, as well as deliberate work within the geometrically derived classes. Despite best effort being made to carefully select the training areas so as to obtain a representative subset, the neighborhood in Figure 39 was not captured. Satellite imagery shows that the roofing material in this particular neighborhood as being significantly different from the rest of

the study area. Given that only one of these roofs made it into the training regions, it is not surprising that it was unable to be separated as its own class.

As only the first step in the process to make multi-wavelength LiDAR data operationally valuable as a terrain classification resource, this process demonstrates the validity of using multi-wavelength intensities to derive spectral classes. In order to quantitatively evaluate the spectral classification method, ground truth data must be taken locally and compared to the classification results. Absent proper ground truth, this method can only be evaluated by a subjective or qualitative test. More simply, does utilizing multi-spectral intensity provide a “good” product in terms of terrain classification? Given that the process is capable of revealing grassy ground regions, clearly identifying multiple types of pavement and building materials, all while remaining in the LiDAR point cloud makes it a successful investigation. The spectral components of the multi-wavelength LiDAR data provide better terrain classification than relying on geometry alone.

2. Spectral Classes Derived From Random Subset Method

Although the classes were derived in a different manner, the characteristics of the ML classifier remained the same. In this case, setting a probability threshold of 0.5 resulted in 8.8 million points remaining unclassified. The strengths of the random subset method over the training region method described in Chapter IV are highlighted here. In addition to inclusion of all the major spectral classes, the subset method also results in ROIs that are more spectrally broad, with variation in the environment across the whole study area. This appears to aid the ML classifier in accurately classifying points. Specifically in terms of buildings, the result is drastically better due to the inclusion of a new building class, and better performance on commission errors. The weakness of the random subset method is more easily distinguished by comparing the lists of spectral classes in Tables 4 and 5. Less classes were extracted via random subset, with smaller classes like road paint not being sufficiently represented in N-D VIS.

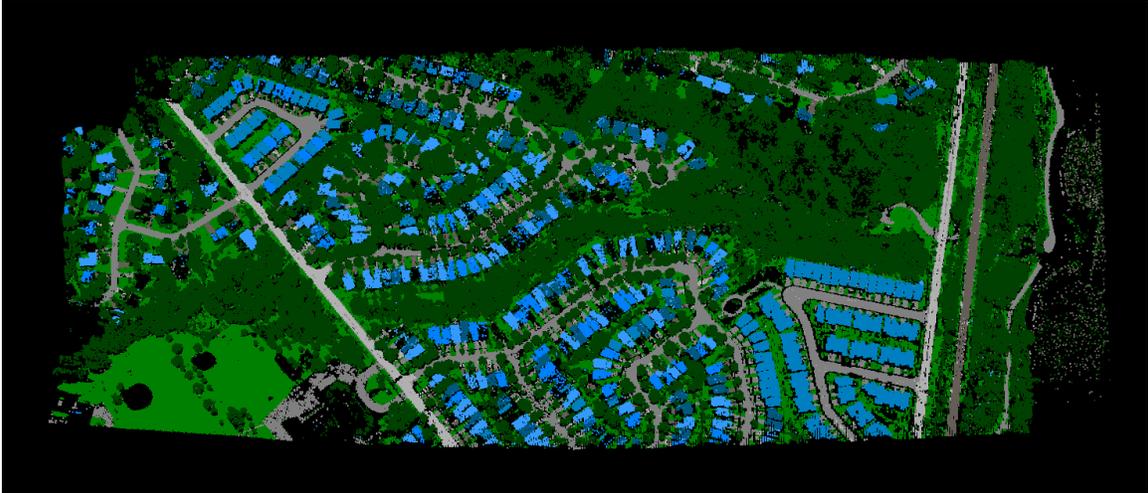


Figure 40. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing all 9 spectral classes derived by the random subset method.



Figure 41. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 2–4, highlighting roads, driveways and railroad.

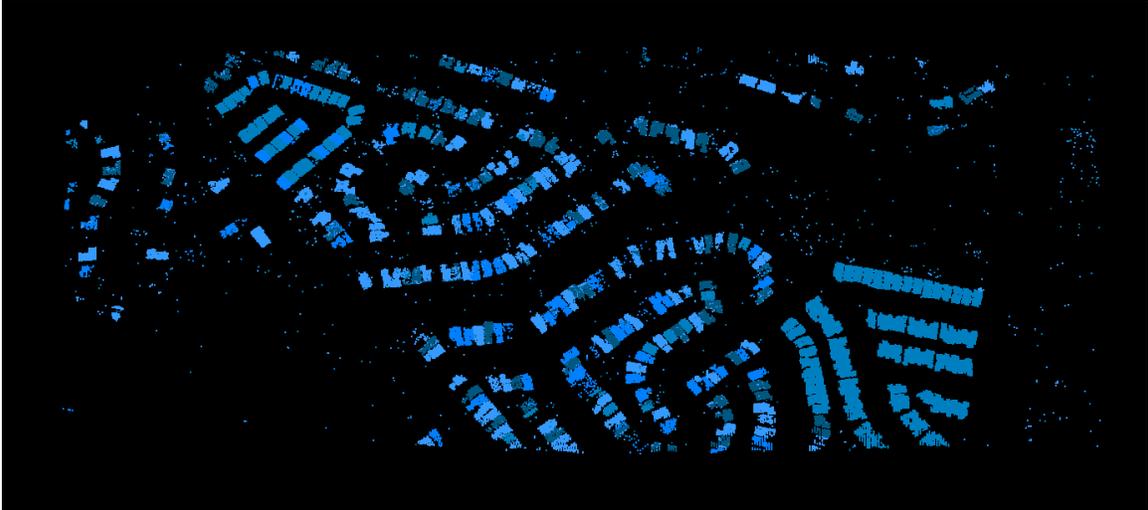


Figure 42. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 6–9, highlighting buildings,

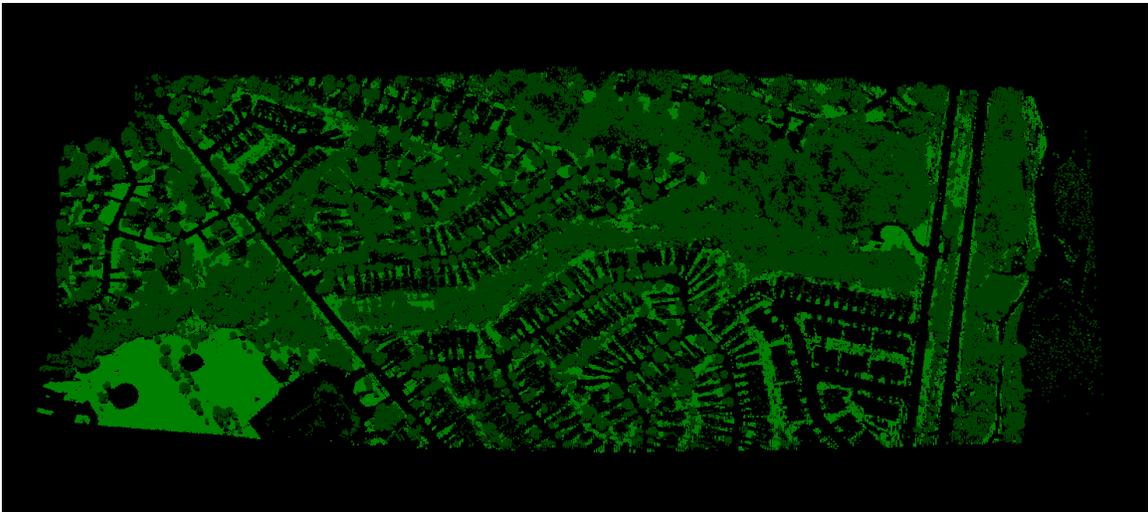


Figure 43. Demonstration of ML Classifier results with a medium probability threshold (0.50) showing spectral classes 1 and 5, highlighting vegetation separated by grass and trees.

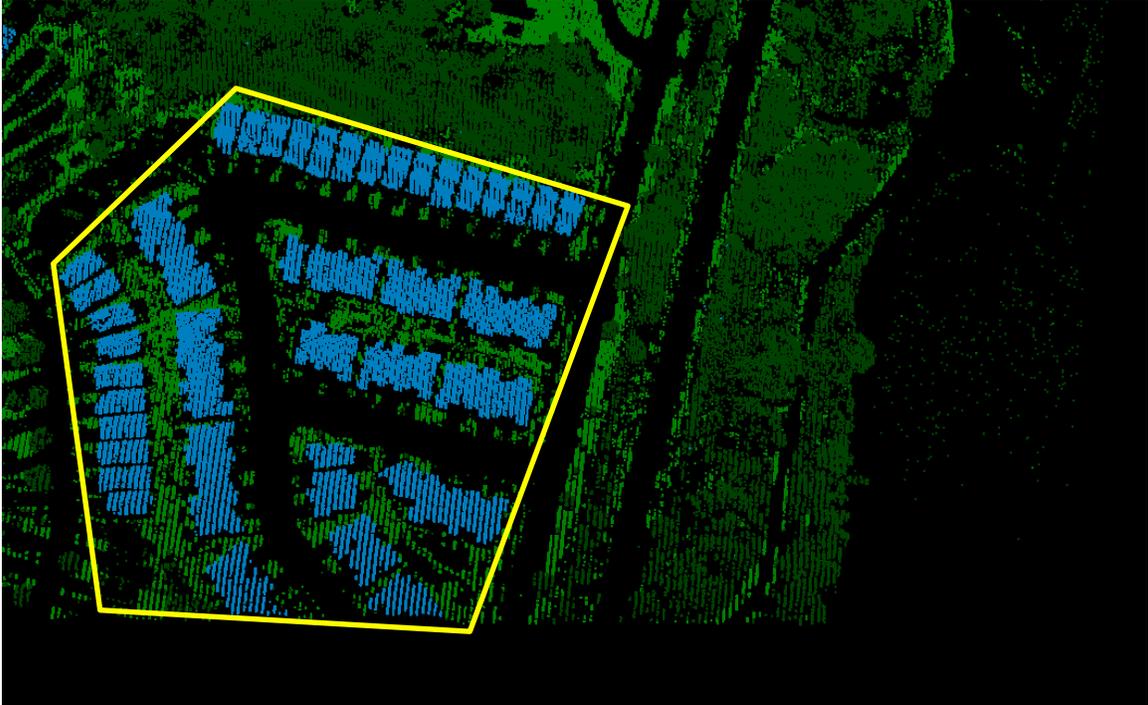


Figure 44. Detailed view of performance advantage for random subset classification in the instance of one neighborhood.

B. OTHER CLASSIFICATION TOOLS

Besides Maximum Likelihood, a handful of other classification tools were tested for their ability to spectrally classify the LiDAR point cloud. The K-Means classifier was run as a baseline to evaluate the ability of an unsupervised classification method. In addition to Maximum Likelihood, the Spectral Angle Mapper (SAM) classifier was run using the same ROIs as ML, as a supervised classification alternative. Neither the K-Means nor the SAM classifiers performed well. In the case of the SAM, errors were driven largely by the absence of the AGL metric to better the classification results.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION

A. STEPS TOWARD OPERATIONAL UTILITY

The Optech Titan sensor is a unique product that offers more spectral fidelity to the LiDAR point cloud than any previous commercial sensor. With that in mind this analysis was aimed at evaluating the potential advantages of a multi-wavelength sensor over a mono-chromatic LiDAR, specifically in terms of terrain classification. The analysis indicates that having multiple wavelengths makes the sensor much more useful for classification over simple geometry driven LiDAR classification. It is impossible to tell without conducting a ground truth campaign whether multi-spectral can out-perform efforts to conduct terrain classification using single wavelength LiDAR. It is likely that processes like the one employed by Brennan, using intensity, return number, AGL, and a DSM could be improved by implementing more conservative spectral groups bounded by three wavelengths.

The biggest challenge to future utilization of a multi-wavelength LiDAR sensor for terrain classification is the global development of anti-access/area-denial (A2/AD) architectures. LiDAR is almost universally hosted on aerial platforms which are vulnerable in an A2/AD environment. Only slightly less of a challenge is creating the processing network which could make terrain classified products available to the warfighter in a timely manner. Since LiDAR is a relatively new technology compared to electro-optical imagery and synthetic aperture radar, no such network currently exists for LiDAR data. Additionally, the critical analysis of this thesis was conducted in a software designed for image processing, not LiDAR point clouds. In order to perform much of the analysis, data files were manipulated to make use of tools like N-D VIS in ways that they were not intended. As the benefits of multi-wavelength LiDAR become apparent, the software to more efficiently exploit various intensities must also be developed.

B. SPACE APPLICATION

Currently LiDAR instruments are not widely used in space based remote sensing. One of the few on orbit assets is the Geoscience Laser Altimeter System (GLAS). GLAS

was developed by NASA Goddard to measure topography and temporal changes in ice-sheets, as well as take atmospheric measurements.³⁹ The issue with using LiDAR from space lies in diffusion and atmospheric effects. GLAS produces a spot 70 m in diameter on the Earth's surface, and the distance between spots is over 150 meters.⁴⁰ This is clearly inadequate for meeting the need of terrain mapping and classification for operational purposes. As the technology of lasers continues to evolve, higher power, more finely collimated sensors may be able to be employed from Low Earth Orbit (LEO) for the purpose of observing the Earth's surface features. There are significant benefits to having a space based LiDAR sensor which could accomplish terrain classification in addition to accurately detailed three dimensional modeling. It could provide a level of intelligence capable of limiting the domestic advantage of the adversary, but more importantly as an active sensor it can operate in any lighting condition. This broadens the operational window of the sensor, and does not limit the host satellite to a sun-synchronous orbit. For the near future however, LiDAR will continue to be dependent on space in terms of GPS, but flown in the atmosphere.

C. FUTURE WORK

Multi-wavelength LiDAR represents a new dimension in the field of LiDAR technology. Companies like Optech will need to generate products which justify the complexity of a multi-wavelength sensors over traditional LiDAR assets. One such product could be a process for conducting terrain classification with high fidelity and timely results. Validation of the results of the spectral classification techniques could be achieved in a number of ways. First, spectral data could be taken to obtain ground truth in the area of the data which is already on hand. Alternatively, obtaining data where ground truth is already known or readily attainable may be more feasible. In either case a quantitative analysis could be conducted to provide a hard metric describing the accuracy of multi-spectral classification. Another step toward progressing the spectral classification would be to investigate different means of radiometric correction of the

³⁹ "ICESat Cryospheric Sciences Lab, Code 615," NASA Goddard Space Flight Center, n.d., accessed online at: <http://icesat.gsfc.nasa.gov/icesat/glas.php> on 27 August, 2015.

⁴⁰ Ibid.

LiDAR intensities. One of the issues with this analysis is that while a major motivation for a multi-wavelength LiDAR sensor is vegetation analysis, separate vegetation classes were indistinguishable. Radiometric correction may have allowed different vegetation classes to coalesce in an otherwise spectrally defuse 3D scatter plot. With no correction applied to this data, the negative intensity effects common not only in vegetation could be a source of inaccuracy throughout the analysis process. The classification process should be expanded to be more robust, taking advantage of the full range of data provided by the LiDAR collection. A multi-tiered analysis leveraging geometry, return number, AGL, spectral intensities, and derived indices provides the best opportunity to accurately classify terrain.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

- Brantberg, Tomas. "Classifying Individual Tree Species under Leaf-off and Leaf-on Conditions Using Airborne LiDAR," *ISPRS Journal of Photogrammetry and Remote Sensing* 61, no. 5, (January 2007): 325–340, doi: 10.1016/j.isprsjprs.2006.10.006.
- Brennan, R., and T. L. Webster. "Object-oriented Land Cover Classification of Lidar-Derived Surfaces," *Canadian Journal of Remote Sensing* 32, no. 2, (2006): 162–172, doi: 10.5589/m06-015.
- Donoghue, Daniel N., Peter J. Watt, Nicholas J. Cox, & Jimmy Wilson. "Remote Sensing of Species Mixtures in Conifer Plantations Using LiDAR Height and Intensity Data," *Remote Sensing of Environment* 110, no. 4, (10 October 2007): 509–522, doi: 10.1016/j.rse.2007.02.032.
- Fiocco, G., and L.D. Smullen. "Detection of Scattering Layers in the Upper Atmosphere (60–140 km) by Optical Radar," *Nature* 199, no. 4900, (28 September 1963): 1275–1276, doi: 10.1038/1991275a0.
- Fleming, Sam, Antoine Cotton, and Iain Woodhouse. "The First Spectral Map of a Forest Understory from Multi-spectral LiDAR," *LiDAR News Magazine* 5, no. 1, (2015): 3, accessed online: <https://carbomap.wordpress.com/2015/02/23/lidar-news-first-multispectral-lidar-map-of-a-forest-understory/>.
- Fleming, Sam, Iain Woodhouse, and Antoine Cotton. "Bringing Colour to Point Clouds," *GIM International* 29, no. 2, (February 2015): 1–3, accessed online: <http://www.gim-international.com/content/article/bringing-colour-to-point-clouds?output=pdf>.
- Gaulton, Rachel. "The Potential of Dual-Wavelength Laser Scanning for Estimating Vegetation Moisture Content," *Remote Sensing of Environment* 132, (15 May 2013): 35.
- Guenther, Gary C.. "Meeting the Accuracy Challenge in Airborne LiDAR Bathymetry," *Proceedings of European Association of Remote Sensing Laboratories Workshop LIDAR Dresden, Germany*, (16-17 June 2000): 3–4, accessed online at: http://www.e proceedings.org/static/vol01_1/01_1_guenther1.pdf?SessionID=63c3bb4a5564b5f62d99e on May 26, 2015.
- Hakala, Teemu. "Full Waveform Hyperspectral LiDAR for Terrestrial Laser Scanning," *Optics Express* 20, no. 7, (26 March 2012): 7120.

- Helt, Michael. "Vegetation Analysis with LiDAR," (master's thesis, Naval Postgraduate School, September 2005): 30–40, retrieved from Calhoun: <http://hdl.handle.net/10945/2030> on April 28, 2015.
- Hickman, G. Daniel, and John E. Hogg. "Application of an Airborne Pulsed Laser for Near-shore Bathymetric Measurements," *Remote Sensing of Environment* 1, no. 1, (March 1969): 47–58, doi: 10.1016/S0034-4257(69)90088-1.
- Hölfe, Berhard. "Correction of Laser Scanning Intensity Data: Data and Model-driven Approaches," *ISPRS Journal of Photogrammetry and Remote Sensing* 62, no. 6, (December 2007): 415, doi: 10.1016/j.isprsjprs.2007.05.008.
- Holmgren, Johan, and Åsa Persson. "Identifying Species of Individual Trees Using Airborne Laser Scanner," *Remote Sensing of Environment* 90, no. 4, (21 May 2003): 415–423, doi: 10.1016/S0034-4257(03)00140-8.
- "ICESat Cryospheric Sciences Lab, Code 615," NASA Goddard Space Flight Center, n.d., accessed online at: <http://icesat.gsfc.nasa.gov/icesat/glas.php> on August 27, 2015.
- Maiman, Theodore H.. "Stimulated Optical Radiation in Ruby," *Nature* 187, no. 4736, (6 August 1960): 493–494, doi: 10.1038/187493a0.
- "Maximum Likelihood," Excelis Visual Information Systems, n.d., accessed online at <http://www.exelisvis.com/docs/MaximumLikelihood.html> on 17 Aug. 2015.
- Miller, Barry. "Laser Altimeter May Aid Photo Mapping," *Aviation Week & Space Technology* 83, no. 13, (29 March 1965): 60–61.
- Nordin, Linda. "Analysis of Waveform Data from Airborne Laser Scanner System," (master's thesis, Lulea University of Technology, 2006): 38.
- Olsen, Richard. *Remote Sensing from Air and Space*, (SPIE Press, 2007), 229.
- Optech Titan Multispectral LiDAR System: High Precision Environmental Mapping, (Brochure), n.d., 3, retrieved June 4, 2015 from <http://www.teledyneoptech.com/wp-content/uploads/Titan-Specsheet-150515-WEB.pdf>.
- Richards, John A.. *Remote Sensing Digital Image Analysis: An Introduction*, (Berlin, Germany: Springer, 2006), 193–200.
- Schawlow, Arthur L., and Charles H. Townes. "Infrared and Optical Masers," *Physical Review* 112(6), (26 August 1958), 1948–1949, doi: <http://dx.doi.org/10.1103/PhysRev.112.1940>.

- Song, Jeong-Heon, Soo-Hee Han, Kiyun Yu, and Yong-II Kim. "Assessing the Possibility of Land-cover Classification Using LIDAR Intensity Data," *ISPRS Commission III, "Photogrammetric Computer Vision" Graz, Austria* 34, no. 3B, (9-13 September, 2002): 259–262, retrieved from: <http://www.isprs.org/proceedings/XXXIV/part3/>.
- Tan, Songxin and Ram Narayanan. "Design and Performance of a Multiwavelength Airborne Polarimetric LiDAR for Vegetation Remote Sensing," *Applied Optics* 43, no. 11, (10 April 2004): 2362, 2367.
- Wei, Gong. "Multi-wavelength Canopy LiDAR For Remote Sensing of Vegetation: Design and System Performance," *ISPRS Journal of Photogrammetry and Remote Sensing* 69, (April 2012): 1–9, doi: 10.1016/j.isprsjprs.2012.02.001.
- Woodhouse, Iain. "A Multispectral Canopy LiDAR Demonstrator Project," *IEEE Geoscience and Remote Sensing Letters* 8, no. 5, (21 April 2011): 839.
- Zhang, Keqi. "Automatic Construction of Building Footprints from Airborne LiDAR Data," *IEEE Transactions on Geoscience and Remote Sensing* 44, no. 9, (September 2006): 2532, doi: 10.1109/TGRS.2006.874137.

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