



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**POINT DENSITY EFFECTS ON DIGITAL ELEVATION
MODELS GENERATED FROM LIDAR DATA**

by

Richard L. Duldulao

June 2009

Thesis Advisor:
Second Reader:

R.C. Olsen
David M. Trask

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 June 2009	3. REPORT TYPE AND DATES COVERED Master's Thesis	
4. TITLE AND SUBTITLE Point Density Effects on Digital Elevation Models Generated from LiDAR Data			5. FUNDING NUMBERS	
6. AUTHOR(S) R. L. Duldulao				
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	
13. ABSTRACT (maximum 200 words) The use of Airborne LiDAR Systems (ALS) to obtain topographical information of the earth's surface and generate Digital Elevation Models (DEMs) has grown extensively in the field of Remote Sensing. Selected areas of point cloud LiDAR data collected from Honduras in 2008 was used to produce DEMs with varying densities to show the effects of lower resolution LiDAR data. An IDL code was utilized to reduce the selected LiDAR point cloud data to 90%, 66%, 50%, 30%, 10%, 5%, 3%, 1%, 0.5%, 0.3%, 0.1%, 0.05%, 0.03%, and 0.01% of its original density to obtain lower resolution data sets. The software Quick Terrain Modeler (QTM) and its ILAP Bare Earth Extractor Plug-in was used to generate DEMs from the varying point cloud density data sets and the software ENVI was used to perform DEM analysis. It was found that LiDAR point cloud density data set of at least 0.6 points per square meter is necessary to generate an accurate Digital Elevation Model for the test environment.				
14. SUBJECT TERMS DEM, Digital Elevation Model, Point Density, LiDAR, QTM, Quick Terrain Modeler, ILAP Bare Earth Extractor, ENVI.			15. NUMBER OF PAGES 73	
			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. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited.

**POINT DENSITY EFFECTS ON DIGITAL ELEVATION MODELS
GENERATED FROM LIDAR DATA**

Richard L. Duldulao
Lieutenant, United States Navy
B.S., University of Idaho, 2003

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN APPLIED PHYSICS

from the

**NAVAL POSTGRADUATE SCHOOL
June 2009**

Author: Richard L. Duldulao

Approved by: R. C. Olsen
Thesis Advisor

COL (Ret) Dave Trask
Second Reader

James Luscombe
Chairman, Physics Department

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

The use of Airborne LiDAR Systems (ALS) to obtain topographical information of the earth's surface and generate Digital Elevation Models (DEMs) has grown extensively in the field of Remote Sensing. Selected areas of point cloud LiDAR data collected from Honduras in 2008 was used to produce DEMs with varying densities to show the effects of lower resolution LiDAR data. An IDL code was utilized to reduce the selected LiDAR point cloud data to 90%, 66%, 50%, 30%, 10%, 5%, 3%, 1%, 0.5%, 0.3%, 0.1%, 0.05%, 0.03%, and 0.01% of its original density to obtain lower resolution data sets. The software Quick Terrain Modeler (QTM) and its ILAP Bare Earth Extractor Plug-in was used to generate DEMs from the varying point cloud density data sets and the software ENVI was used to perform DEM analysis. It was found that LiDAR point cloud density data set of at least 0.6 points per square meter is necessary to generate an accurate Digital Elevation Model for the test environment.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	BACKGROUND	3
A.	LIGHT DETECTION AND RANGING (LiDAR)	3
B.	AIRBORNE LIDAR SYSTEM (ALS)	4
1.	Components of an Airborne LiDAR System.....	5
a.	Laser Ranging Unit.....	5
b.	Position and Orientation System (POS).....	6
c.	Synchronization	8
2.	Registration	8
3.	Point Density.....	12
4.	LAS LiDAR Data Standard	13
C.	DIGITAL ELEVATION MODEL (DEM) GENERATION PROCESS ..	14
1.	Filtering.....	14
2.	Model Selection	15
3.	DEM Interpolation.....	15
D.	PREVIOUS DATA ANALYSIS RESULTS	15
III.	OBSERVATIONS.....	19
A.	LOCATION.....	19
B.	POST-PROCESSING AND DATA ANALYSIS SOFTWARE.....	20
1.	Quick Terrain Modeler (QTM) Version 6.0.6.....	20
2.	ILAP Bare Earth Extractor Version 1.0.....	21
3.	Environment for Visualizing Images (ENVI) Version 4.5	21
C.	METHODS	21
1.	Generation of the Base Model.....	21
2.	Generation of Digital Elevation Models of Varying Densities	22
IV.	ANALYSIS	27
A.	PREPARATION OF DIGITAL ELEVATION MODELS FOR ANALYSIS	27
1.	Warping	27
2.	Masking.....	29
B.	CORRELATION ANALYSIS IN ENVI.....	29
1.	Reference Model of Size 1625 Meters by 875 Meters (Smaller Reference Model)	30
2.	Reference Model of Size 3005 Meters by 844 Meters (Larger Reference Model)	36
C.	VALIDATION OF DECIMATION APPROACH	45
V.	CONCLUSION	51
	LIST OF REFERENCES	53
	APPENDIX: RANDOM_PTS_FROMXYZ_V2.PRO.....	55
	INITIAL DISTRIBUTION LIST	57

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

Figure 1.	LiDAR Pulse-Echo Range Finder (From Petrie & Toth, 2009)	3
Figure 2.	LiDAR Continuous Wave Range Finder (From Petrie & Toth, 2009).....	4
Figure 3.	Airborne LiDAR System (From Burtch, 2002)	5
Figure 4.	Sample Layout of a Laser Ranging Unit (From Petrie & Toth, 2009)	6
Figure 5.	Multiple returns from a forested area (From Harding, 2009)	6
Figure 6.	Position and Orientation System Components (From Wehr, 2009)	7
Figure 7.	IMU mounted on top of Laser Ranging Unit (From Wehr, 2009).....	7
Figure 8.	Sample PCU and LCU of an ALS (From Wehr, 2009)	8
Figure 9.	Registration of LiDAR data points (From Wehr, 2009)	9
Figure 10.	IMU, GPS and LiDAR Configuration (From Wehr, 2009)	10
Figure 11.	Roll, pitch, and heading of an aircraft carrying an ALS (From Wehr, 2009)..	11
Figure 12.	Scanning lines over a flat (a) and sloping (b) terrain (From Wehr, 2009).....	13
Figure 13.	Accuracy Measurement of Data Reduction using Root Mean Square and Standard Deviation (From Liu, et al., 2007)	17
Figure 14.	Google earth image of 1625 meters by 875 meters (a) and 3005 meters by 844 meters b) taken from Mocoron, Honduras.	19
Figure 15.	Sequoia National Park LiDAR data coverage	20
Figure 16.	QTC format of the reference model (1625 meters by 875 meters).....	22
Figure 17.	DEM generated from the reference model.....	23
Figure 18.	Digital Elevation Models in QTT format visualized using QTM software. DEMs from each of the reduced LiDAR data set of 1625 meters by 875 meters reference model.	25
Figure 19.	Warp parameters	28
Figure 20.	Warped DEM from 0.01% of the smaller reference model (right image)	28
Figure 21.	100% (left) and 0.01% (right) data DEMs after mask has been applied	29
Figure 22.	Correlation of each DEM to the DEM generated from the smaller reference model. These are the DEMs depicted in Figures 17 and 18.	30
Figure 23.	Different correlation factors obtained from two similar studies. (a) Results of previous study (Anderson, 2008). (b) Results from this study.	32
Figure 24.	Correlation factors of the first set of DEMs (1625 meters by 875 meters) labeled “First Run” and the additional DEMs generated from 3%, 0.3%, and 0.1% of the smaller reference model.....	33
Figure 25.	Calculated Mean Correlation Factor of DEM generated from 0.1% of 1625 meters by 875 meters reference model.	34
Figure 26.	Percentage of Total Point Cloud classified as Surface Points	35
Figure 27.	DEM generated from the larger (3005 meters by 844 meters) reference model of Mocoron, Honduras.	37
Figure 28.	Digital Elevation Models in QTT format visualized using QTM software. DEMs from each of the reduced LiDAR data set of 3005 meters by 844 meters reference model.	41

Figure 29.	Correlation analysis of DEMs generated from the larger reference model plotted on the same plot with the first set of DEMs generated from the smaller reference model.	43
Figure 30.	Percentage of Classified Surface Points with Diminishing Point Cloud Density. First set of DEMs (Red diamonds) and new set of DEMs (blue circle)	44
Figure 31.	High (a) and standard (b) resolution LiDAR data of Sequoia National Park visualized in QTC format using QTM	47
Figure 32.	Selected (white box) LiDAR Point Cloud (QTC) of Standard and High Resolution data viewed in QTM	48
Figure 33.	High (a) and Standard (b) Resolution DEMs generated from the selected LiDAR Point Cloud data sets (Figure 33)	49

LIST OF TABLES

Table 1.	Samples of LiDAR data attributes contained in a standard LAS file (From Graham, 2009)	14
Table 2.	Number of Points and Densities of each DEM.....	26
Table 3.	Percentage of Total Point Cloud of the reference model classified as Surface Points	36
Table 4.	Number of Points and Densities of each DEM generated from a larger reference model.....	42
Table 5.	Percentage of Total Point Cloud of the larger reference model classified as Surface Points	45

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

Special Thanks to:

Dr. R. C. Olsen, Naval Postgraduate School

COL (Ret) David M. Trask, Naval Postgraduate School

Angela Puetz, Naval Postgraduate School

Nancy Ann Budden, Naval Postgraduate School

Office of the Secretary of Defense Rapid Technology Transition Office (OSD/RTTO)

Eric Adint, Naval Postgraduate School

My fellow Combat Systems Classmates

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Strategic planning of a military operation is critical to mission success and knowledge of the terrain is a key contributor to the development of proper strategic planning. Military leaders use the lay of the land to determine where to initiate an attack and where to establish a perimeter for a defensive stance that would provide the strategic advantage.

An increasingly popular method is the Airborne LiDAR Systems (ALS), which has the ability to detect objects under tree canopies, provide data to generate digital elevation models, and provide three-dimensional models of non-surface or manmade structures. ALS has high-density data and is able to provide accurate, detailed digital representations of terrain and of targets of interest. However, to ensure the high accuracy of digital models most efforts in LiDAR data collection lead to oversampling, which result in excessive LiDAR data density for the intended purpose. The present software and hardware equipments used for LiDAR data processing are limited in terms of their capabilities to generate large area of digital models from high-density LiDAR data, and excessive data compounds computational demands.

The purpose of this study is to determine the impact of point density on the accuracy of DEMs, with the goal of defining the minimum necessary point density for a given environment. The LiDAR data used in this study was collected over the jungle of Honduras in 2008. To represent lower resolution LiDAR data sets, this study conducted subsequent reductions of a reference model. These lower density LiDAR data sets were used to generate digital elevation models which were compared against the digital elevation model generated from the reference model.

Chapter II presents a broad technical background of LiDAR technology and the support systems necessary in order to conduct Airborne LiDAR System surveys. It also briefly discusses the steps taken, and the various algorithms involved in each step, to

generate digital elevation models. A similar study of identifying the lowest data density necessary to generate an accurate digital elevation model using standard deviation and root mean square is also presented.

Chapter III provides a description of the locations used in this study, the post-processing software Quick Time Modeler and ILAP Bare Earth Extractor utilized in generating DEMs, and the software (ENVI) used in analysis. It also describes in detail the process of generating the digital elevation models with varying densities.

Chapter IV describes the process used in ENVI to analyze each of the digital elevation models and the evaluation processes to obtain the lowest point-cloud density that would generate an accurate digital elevation model for this environment.

II. BACKGROUND

A. LIGHT DETECTION AND RANGING (LiDAR)

Light Detection and Ranging is a remote sensing technology used in finding information of a target by measuring the properties of returned or scattered light transmitted by a laser system. It uses much shorter wavelengths, usually in the near infrared section of the electromagnetic spectrum, providing better spatial resolution than RADAR technology.

LiDAR is implemented in two ways. The more common approach utilizes discrete pulses to determine the range of a target. LiDAR measures the time of flight (TOF) of a pulse from the transmitter to the target and its reflected signal received from the target to the detector. LiDAR calculates range using the TOF and the pulse's speed, the speed of light (Petrie & Toth, 2009). As expressed in equation 2.1, c is the speed of light, Δt is the TOF and R is the range of the object. Equation 2.1 has a factor of one-half to account only for the pulse's travel from the LiDAR transmitter to the target.

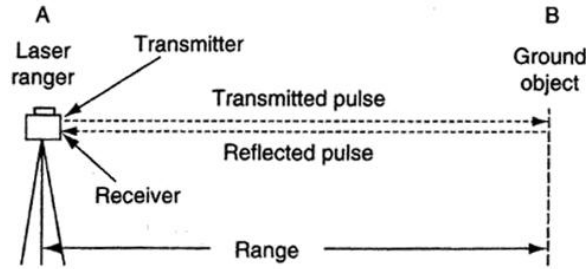


Figure 1. LiDAR Pulse-Echo Range Finder (From Petrie & Toth, 2009)

$$R = c * \frac{\Delta t}{2} \quad (2.1)$$

Another method (more commonly used at distances less than 100 meters) utilizes continuous wave (CW) or beam of laser to determine a target's range. Here, LiDAR determines range by measuring the integer number of wavelengths ($M\lambda$) and the phase difference ($\Delta\lambda$) between the transmitted and received waveforms of the emitted beam.

The M number of wavelengths is measured by varying the modulation frequencies of the emitted beam (Petrie & Toth, 2009). Figure 2 shows the phase difference between the transmitted and received signals measured at point A. Point B is the location of the target requiring two full wavelengths. The range of the target is calculated using equation 2.2. The factor one-half is included for the same reason as in equation 2.1.

$$R = \frac{(M\lambda + \Delta\lambda)}{2} \quad (2.2)$$

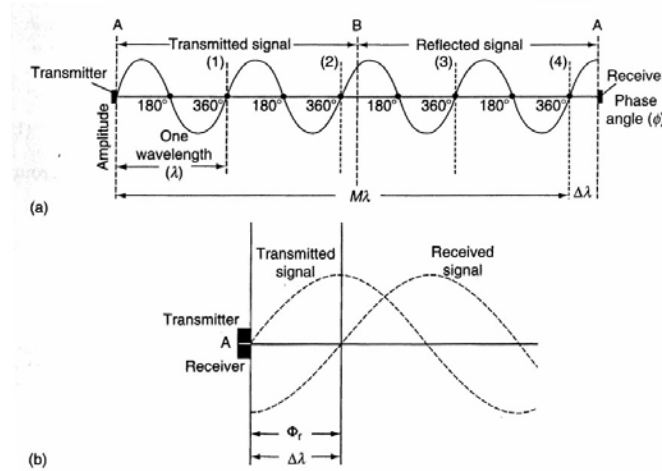


Figure 2. LiDAR Continuous Wave Range Finder (From Petrie & Toth, 2009)

B. AIRBORNE LIDAR SYSTEM (ALS)

The advancements in Global Positioning Systems (GPS) and its integration with Inertial Navigation Units (INS) have provided the means to use LiDAR systems onboard an aircraft, later named Airborne LiDAR Systems or ALS. Over the past decades, more reliable and accurate ALS have been developed. This led to a significant increase in the use of LiDAR data in generating Digital Elevation Models (DEMs) (Liu, 2008). Figure 3 depicts an aircraft mapping the shape of the terrain and showing the three major components of an ALS: a laser ranging unit; GPS; and INS.

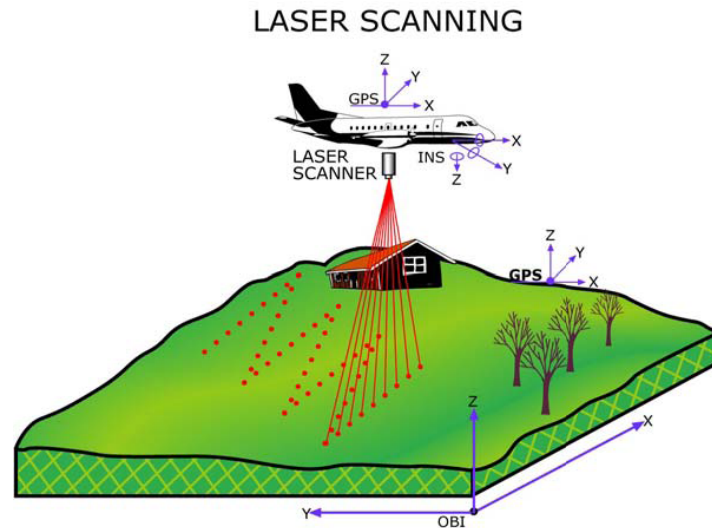


Figure 3. Airborne LiDAR System (From Burtch, 2002)

1. Components of an Airborne LiDAR System

a. Laser Ranging Unit

The laser ranging unit consists of a diode-pumped solid-state laser commonly made of neodymium-doped yttrium aluminum garnet (Nd:YAG). It transmits pulses with wavelengths between $0.8 \mu m$ and $1.6 \mu m$ (typical wavelengths used are $1.064 \mu m$ or $1.500 \mu m$). Laser pulses usually have pulse widths of 4 to 15 ns with peak energies of several millijoules and are emitted at a rate of up to 250 kHz (Liu, 2008). A photodiode detector made of silicon (for wavelengths up to 1100 nm) or germanium (for wavelengths 1000 to 1650 nm) is used to detect scattered and reflected pulses from targets and converts them to electrical signals (Wehr, 2009). Figure 4 shows a principle layout of a Laser Ranging Unit. Pulses are emitted from the high-powered solid-state laser through the collimator. Reflected signals are collected by the primary and secondary mirrors and are directed to the photodiode for detection.

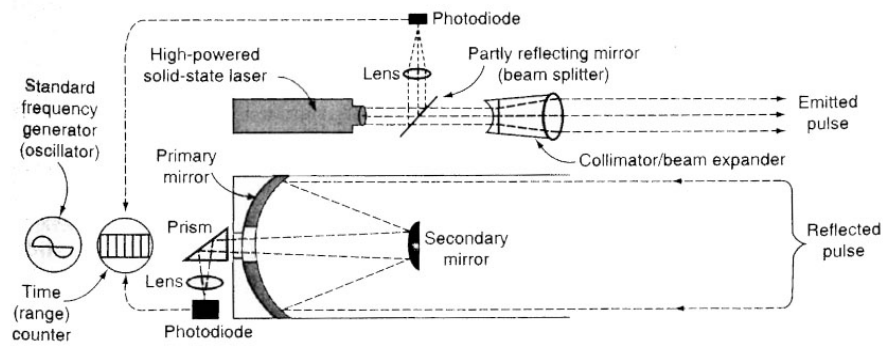


Figure 4. Sample Layout of a Laser Ranging Unit (From Petrie & Toth, 2009)

Advanced laser ranging units are able to detect up to five returns from a single transmitted pulse as illustrated in Figure 5. The return signals are detected when their energies exceed the detection threshold. In forested areas, the first returns correspond to the first leading edge of the detected signal, and may be from the canopy top, from a layer within the canopy or from the ground. The last returns correspond to the leading edge of the latest detected peak, and may be from the ground or from a layer within the vegetation canopy (Harding, 2009).

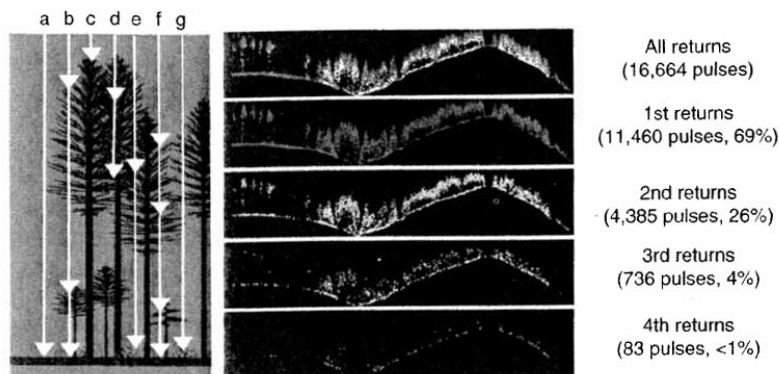


Figure 5. Multiple returns from a forested area (From Harding, 2009)

b. Position and Orientation System (POS)

The position and orientation system (Figure 6) is comprised of a closely integrated Differential Global Positioning System (DGPS) and an Inertial Measurement

Unit (IMU) that provides the ALS's trajectory and attitude (pitch, roll, and yaw). The DGPS requires reference ground stations that must be within 25 km of the ALS to guarantee centimeter level accuracy. The IMU is typically mounted directly on top of the laser ranger scanner (Figure 7) in order to record orientation and aircraft vibrations at the location of the LiDAR. A memory unit or disk is used to store GPS position and IMU data with GPS-time (Wehr, 2009).

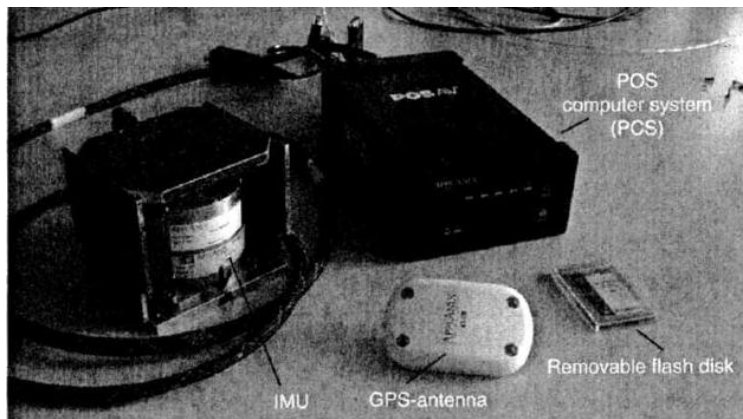


Figure 6. Position and Orientation System Components (From Wehr, 2009)

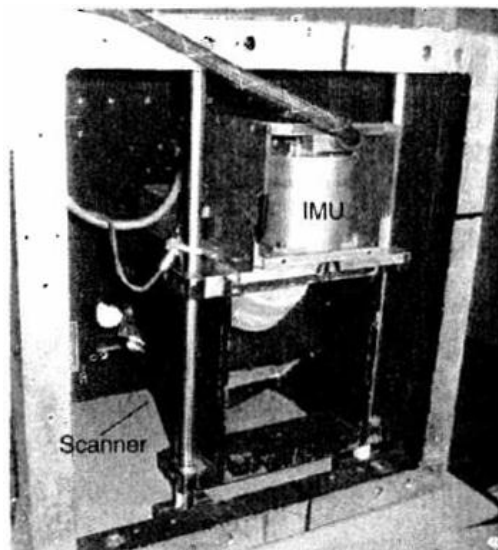


Figure 7. IMU mounted on top of Laser Ranging Unit (From Wehr, 2009)

c. Synchronization

The Laser Ranging Unit and the POS are independent units, which require data or measurement synchronization. The LiDAR Control Unit (LCU) controls and store measurements made by the Laser Ranger Unit while the POS Control Unit (PCU) controls and store measurements made by the GPS and IMU (Figure 8). The LCU timing is defined by its internal computer clock and the PCU timing is related by GPS time. Due to a much higher sampling rate of the Laser Scanner Unit than the POS, a LiDAR file comprises more data lines per time interval than the POS file (Wehr, 2009). Once all data have been synchronized, they are used as inputs for registration.

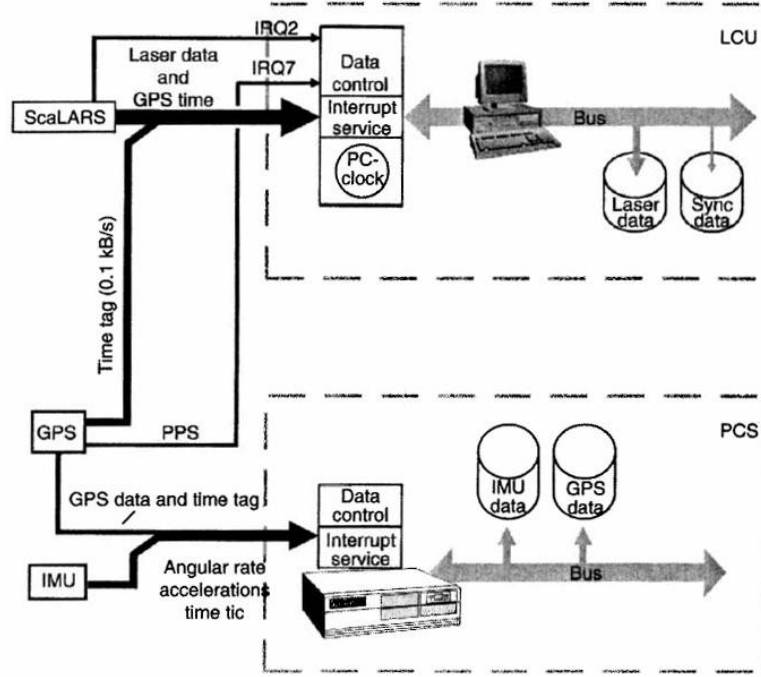


Figure 8. Sample PCU and LCU of an ALS (From Wehr, 2009)

2. Registration

Registration is the process of assigning the location on earth or geocoding of a LiDAR data point acquired in 3D space. It can be described using the simple vector approach illustrated in Figure 9 and expressed in equation 2.3. $\vec{G}$ is the vector from the earth's center to the ground point, $\vec{r}_L$ is the vector from the earth's center to the LiDAR's

point of origin, and $\vec{s}$ is the slant ranging vector. The LiDAR's point of origin defines the origin of the coordinate system L and is the point at which the laser beam originates. The x_L -axis points into the flight direction, the y_L -axis points to the right of the airplane, and the z_L -axis points downwards perpendicular to the plane defined by x_L and y_L axes (Wehr, 2009).

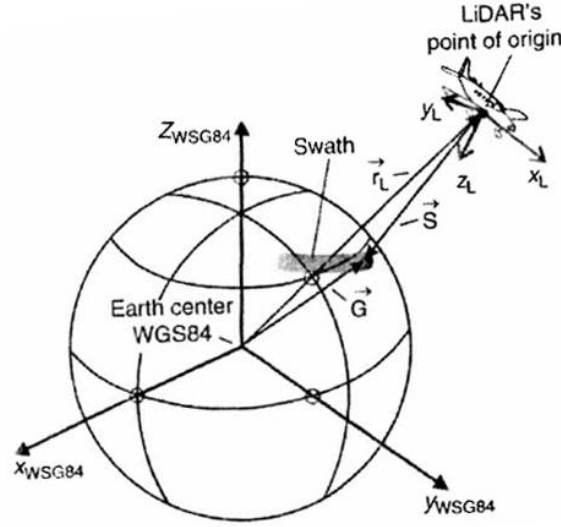


Figure 9. Registration of LiDAR data points (From Wehr, 2009)

$$\vec{G} = \vec{r}_L + \vec{s} \quad (2.3)$$

The IMU, GPS, and LiDAR's point of origin are all in different locations inside an aircraft so a few transformations are necessary to determine the exact $\vec{r}_L$. A sample configuration is shown in Figure 10. The POS data from IMU and GPS need to be transformed to the LiDAR's point of origin using two three-dimensional vectors (called lever arms) to determine the actual location and orientation of the LiDAR: one is from the LiDAR's point of origin to the center of IMU and the other is from the LiDAR's point of origin to the phase center of the GPS antenna. Since GPS systems are in WGS84, the vector $\vec{r}_L$ would be in WGS84 (Wehr, 2009).

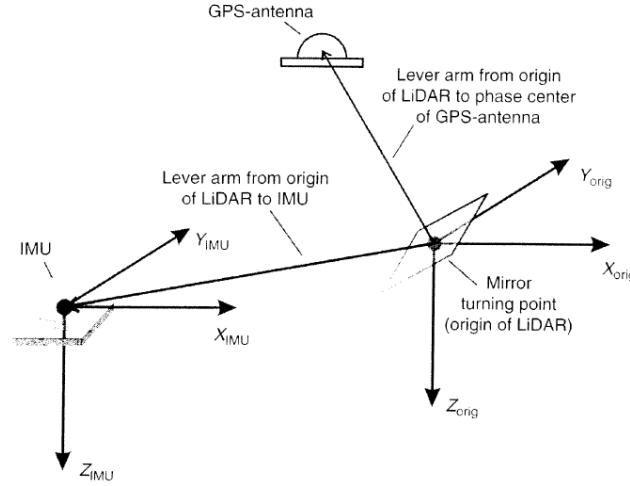


Figure 10. IMU, GPS and LiDAR Configuration (From Wehr, 2009)

The vector $\vec{s}$, measured in the coordinate system L, would require some transformations into WGS84 to determine $\vec{G}$ in WGS84. Thus equation 2.3 has to be modified into equation 2.4:

$$\underline{G}_{WGS84} = \underline{r}_{L_{WGS84}} + (_)_{H}^{WGS84} * (_)_{IMU}^H * (_)_{L}^{IMU} * \underline{s}_L \quad (2.4)$$

The product $(_)_{IMU}^H * (_)_{L}^{IMU}$ describes the orientation of the coordinate system L with respect to the horizontal coordinate system H. The $(_)_{IMU}^H$ matrix describes the orientation of the IMU in relation to the horizontal system H by the roll (ω , rotation about the x_L -axis), pitch (φ , rotation about the y_L -axis), and heading (κ , rotation about the z_L -axis) as shown in Figure 11. If the rotations are carried in the roll, pitch, and heading sequence, the matrix $(_)_{IMU}^H$ can be set up using equation 2.5 where the components are defined in equations 2.6, 2.7, and 2.8 (Wehr, 2009).

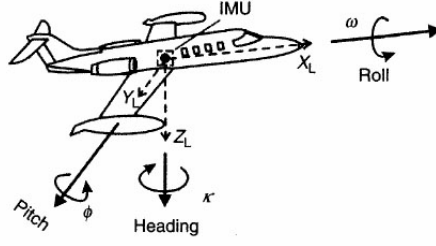


Figure 11. Roll, pitch, and heading of an aircraft carrying an ALS (From Wehr, 2009)

$$(_)_{IMU}^H = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix} \quad (2.5)$$

$$\begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} = \begin{pmatrix} \cos(\kappa) * \cos(\varphi) \\ \sin(\kappa) * \cos(\varphi) \\ -\sin(\varphi) \end{pmatrix} \quad (2.6)$$

$$\begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \end{pmatrix} = \begin{pmatrix} \cos(\kappa) * \sin(\varphi) * \sin(\omega) - \sin(\kappa) * \cos(\omega) \\ \sin(\kappa) * \sin(\varphi) * \sin(\omega) + \cos(\kappa) * \cos(\omega) \\ \cos(\varphi) * \sin(\omega) \end{pmatrix} \quad (2.7)$$

$$\begin{pmatrix} a_{31} \\ a_{32} \\ a_{33} \end{pmatrix} = \begin{pmatrix} \cos(\kappa) * \sin(\varphi) * \cos(\omega) - \sin(\kappa) * \sin(\omega) \\ \sin(\kappa) * \sin(\varphi) * \cos(\omega) + \cos(\kappa) * \sin(\omega) \\ \cos(\varphi) * \cos(\omega) \end{pmatrix} \quad (2.8)$$

The matrix $(_)_{L}^{IMU}$ takes into account a misalignment between the POS and LiDAR. It is similar to the matrix $(_)_{IMU}^H$ except the misalignment angles $\delta\omega$, $\delta\varphi$, and $\delta\kappa$ are used instead of the angles ω , φ , and κ , respectively. If the LiDAR is perfectly aligned with the IMU (the coordinate system L and the coordinate system of the IMU have the same orientation), the matrix $(_)_{L}^{IMU}$ becomes unity (Wehr, 2009).

The matrix $(_)_{H}^{WGS84}$ regards the orientation between the horizontal system H and WGS84. It is defined by the geographical latitude Φ_0 and longitude Λ_0 as shown in equation 2.9 (Wehr, 2009).

$$(_)_{\text{H}}^{\text{WGS84}} = \begin{pmatrix} -\cos(\Lambda_0) * \sin(\Phi_0) & -\sin(\Lambda_0) & -\cos(\Lambda_0) * \cos(\Phi_0) \\ -\sin(\Lambda_0) * \sin(\Phi_0) & \cos(\Lambda_0) & -\sin(\Lambda_0) * \cos(\Phi_0) \\ \cos(\Phi_0) & 0 & \sin(\Phi_0) \end{pmatrix} \quad (2.9)$$

3. Point Density

The point density or laser spots per square meter (ρ) is determined using equation 2.10 where Δx_{along} is the point density in the flight direction and Δx_{across} is the point density across the flight direction. Δx_{along} is dependent upon the speed of the aircraft, v , and the scan rate, f_{sc} , as expressed in equation 2.11 (Wehr, 2009).

$$\rho = \frac{1}{\Delta x_{\text{along}} * \Delta x_{\text{across}}} \quad (2.10)$$

$$\Delta x_{\text{along}} = \frac{v}{f_{\text{sc}}} \quad (2.11)$$

Δx_{across} is calculated using equation 2.12. θ is the swath width expressed in either meters or angular degrees, H is the flying altitude above ground, N is the number of points per scan line, and i is the slope along the scanning line. N is derived from the scan rate and pulse rate f_{pulse} (equation 2.13). Figure 12b illustrates scanning lines over a terrain with slope i (Wehr, 2009).

$$\Delta x_{\text{across}} = \begin{cases} \frac{\theta}{N} * \frac{H}{\cos^2\left(\frac{\theta}{2}\right) * \cos(i) * \left[1 - \tan\left(\frac{\theta}{2}\right) * \tan(i)\right]^2}, & \text{if } i \geq 0 \\ \frac{\theta}{N} * \frac{H}{\cos^2\left(\frac{\theta}{2}\right) * \cos(i) * \left[\tan\left(\frac{\theta}{2}\right) * \tan(i) - 1\right]^2}, & \text{if } i < 0 \end{cases} \quad (2.12)$$

$$N = \frac{f_{\text{pulse}}}{f_{\text{sc}}} \quad (2.13)$$

If a terrain along the scanning line has a flat surface ($i \approx 0$), equation 2.12 reduces down to equation 2.14. Figure 12a shows an aircraft scanning over a flat terrain.

$$\Delta x_{\text{across}} = \left\{ \frac{\theta}{N} * \frac{H}{\cos^2\left(\frac{\theta}{2}\right)} \right\} \quad (2.14)$$

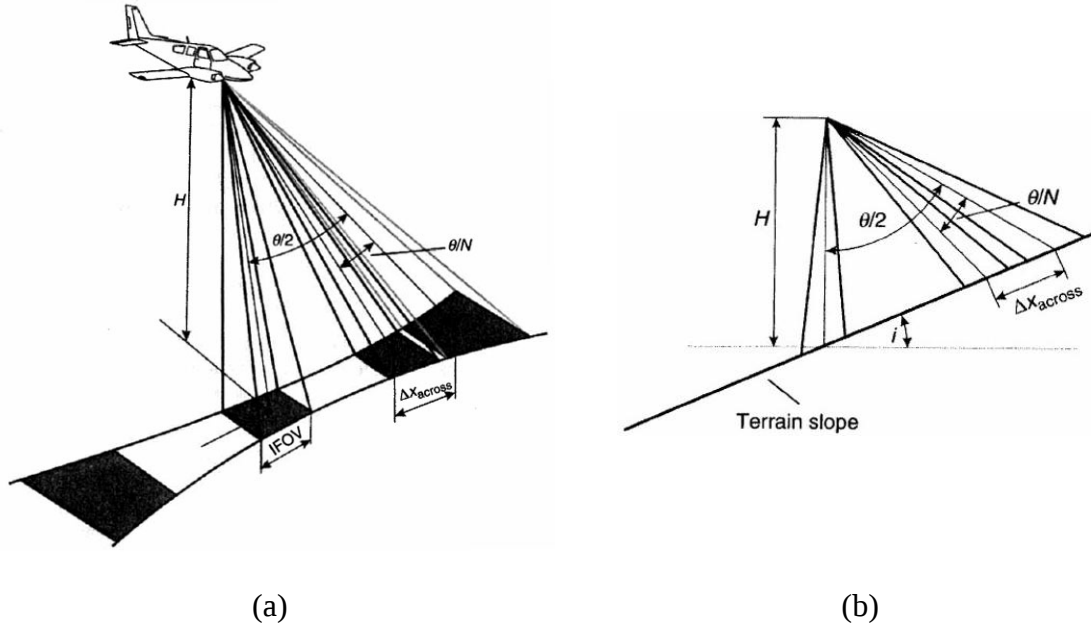


Figure 12. Scanning lines over a flat (a) and sloping (b) terrain (From Wehr, 2009)

4. LAS LiDAR Data Standard

Once the Aircraft has surveyed the area and collected the required data, a post-processing software is used to determine the accurate position, altitude, and attitude of the laser ranger unit and create the LiDAR data in the standard LAS format (Petrie & Toth, 2009b). LAS format is a binary format developed by EnerQuest. It is the format adopted, slightly modified, and approved by the American Society for Photogrammetry and Remote Sensing (ASPRS) as the standard format for LiDAR data exchange. LAS file format does not specify an order to the points in the data file; however, it does require multiple returns be sequentially encoded. For example, a pulse that had three returns will be in sequential order of pulse 1 of three, followed by pulse 2 of three, and then 3 of

three. LAS files contain very rich information of every point acquired during the survey. Table 1 shows some of the LiDAR data attributes contained in LAS file format (Graham, 2009).

Example LiDAR per-Point Data Attributes	
Attribute	Description
X, Y	The planimetric ground location of the point
Z	The elevation of the point
Intensity	The laser pulse return intensity at the sensor
GPS time	The time (in GPS clock time) of the receipt of the return pulse
Number of returns	Number of returns detected for a given transmitted pulse
Return number	The return number of this pulse (e.g., return two of three returns)
Mirror angle	Angle of the scanner mirror at the time of this pulse (only applies to scanning sensors)
Classification	Surface (or other) attribute assigned to this point such as ground, vegetation, and so forth
Point source ID	A unique identifier to reference this point back to a collection source

Table 1. Samples of LiDAR data attributes contained in a standard LAS file (From Graham, 2009)

C. DIGITAL ELEVATION MODEL (DEM) GENERATION PROCESS

The LAS files containing the LiDAR data points require further manipulation to generate a DEM.

1. Filtering

The first process in generating a digital elevation model is filtering. It is one of the critical and difficult steps in DEM generation process involving the separation of LiDAR data into ground (surface) and non-ground (non-terrain) points. Among all the various filtering algorithms developed so far, interpolation-based, slope-based, surface-based, and morphological are the most popular (Liu, 2008). In a study conducted by Sithole and Vosselman, it had been found that most filtering algorithms do well on non-complex landscapes but surface based filters tended to do better on complex landscapes (Sithole & Vosselman, 2003).

2. Model Selection

The remaining points (ground points) are used to generate terrain surfaces. Different model selections have been developed to represent terrain surfaces: regular grid (usually square grid); triangular irregular network (TIN); and contour line model. The regular grid is widely used due to its simplicity and efficient approach in terms of storage and manipulation. However, it introduces discontinuity in representation of the terrain surfaces due to each grid having one elevation value. This effect is minimized by the high density characteristic of LiDAR data (Liu, 2008).

3. DEM Interpolation

Interpolation is the process of predicting the values of certain variable using their neighboring values. It is assumed that a terrain surface is continuous and that a high correlation exists between the neighboring data points. Interpolation methods are classified into deterministic such as Inverse Distance Weighted (IDW) and spline-based and geostatistic such as Kriging. Inverse Distance Weighted (IDW) assumes that each point has a local influence that diminishes with distance and the spline-based fits a minimum-curvature through the sample points. Kriging takes into account both the distance and degree of auto-correlation. A study found there is no single interpolation method that is the most accurate. However, it was pointed out that the IDW method performs well if sampling data density is high, even for complex terrain (Liu, 2008).

D. PREVIOUS DATA ANALYSIS RESULTS

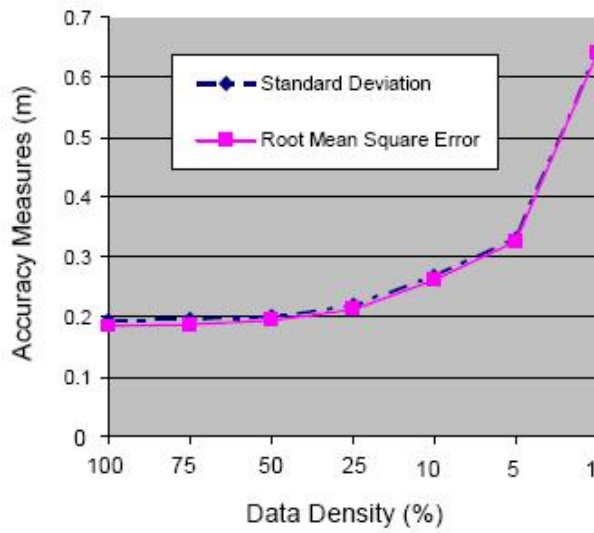
LiDAR data was collected over Corangamite Catchment Management Authority region (south western Victoria, Australia) for an area of 113 square km between 19 July 2003 and 10 August 2003. Using the Geostatistical Analyst extension of ArcGIS 9.1, LiDAR data set was separated into training data set and check point data set by randomly selecting 90% and 10% of the total LiDAR data. The training data set was used for subsequent reduction to produce data sets with varying densities representing 100%, 75%, 50%, 25%, 10%, 5%, 1% of the original training data set (Liu, Zhang, Peterson, Chandra, 2007).

To evaluate the accuracies of each DEM, independent elevation checking of each DEM was conducted against the elevation values of test data using root mean square error (equation 2.15) and standard deviation (equation 2.16) calculations. E_{DEM} is the elevation value from the DEM and E_{REF} is the correspondent reference elevation value from check points. n is the number of check points and $\bar{E}$ is the calculated mean error (equation 2.17). As shown in Figure 13, there is no significant decrease in accuracy for the DEM generated from the 50% (0.018 points per square meter) data set (Liu, et al., 2007).

$$RMSE = \sqrt{\frac{\sum (E_{DEM} - E_{REF})^2}{n}} \quad (2.15)$$

$$\sigma = \sqrt{\frac{\sum (E_{DEM} - \bar{E})^2}{n-1}} \quad (2.16)$$

$$\bar{E} = \sum \frac{(E_{DEM} - E_{REF})}{n} \quad (2.17)$$



(a)

Reduced datasets	Data density (points/per m ²)
100%	0.037
75%	0.028
50%	0.018
25%	0.009
10%	0.004
5%	0.002
1%	less than 0.001

(b)

Figure 13. Accuracy Measurement of Data Reduction using Root Mean Square and Standard Deviation (From Liu, et al., 2007)

THIS PAGE INTENTIONALLY LEFT BLANK

III. OBSERVATIONS

A. LOCATION

This thesis effort used two data sets. One was collected from jungle habitat in Mocoron, Honduras between 11 and 20 February 2008 using a modified Optech 3100 LiDAR system (Anderson, 2008). OSD/RRTO funded and supplied this FOPEN LiDAR collection mission, which they referred to as PENLIGHT. The flight headings were 135 degrees and 315 degrees and the area mapped by the ALS system contained slight overlaps in between flight paths. An area of 1625 meters by 875 meters (Figure 14a) was used covering various terrain types including man-made structures, river waterway, road, jungle foliage, and flat surfaces. A second area of 3005 meters by 844 meters (Figure 14b) was used encompassing the smaller area of 1625 meters by 875 meters previously mentioned to determine the effects of having a larger land area or a larger LiDAR data set.

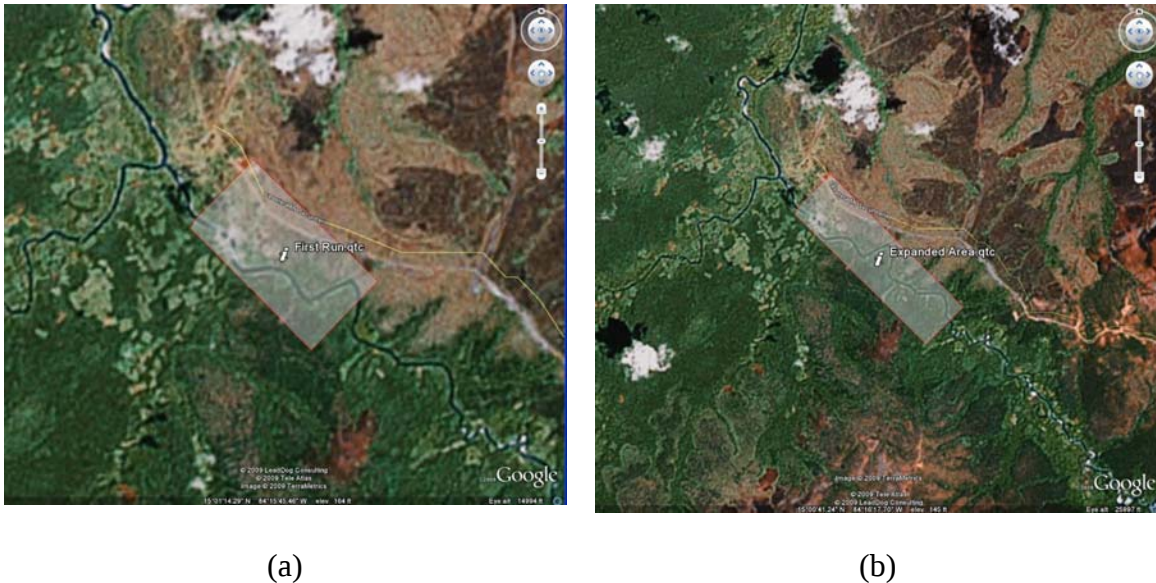


Figure 14. Google earth image of 1625 meters by 875 meters (a) and 3005 meters by 844 meters b) taken from Mocoron, Honduras.

Another set of LiDAR data was collected from Sequoia National Park in California (Figure 15) using Optech 3100 LiDAR system by Airborne1 during the

summer of 2008 (Karatolios & Krougios, 2008). The blue lines show the perimeter of the area covered during data acquisition at a standard resolution and the two inner blue lines show an approximate coverage of a one pass flight of the aircraft at a higher resolution from top right to bottom left.

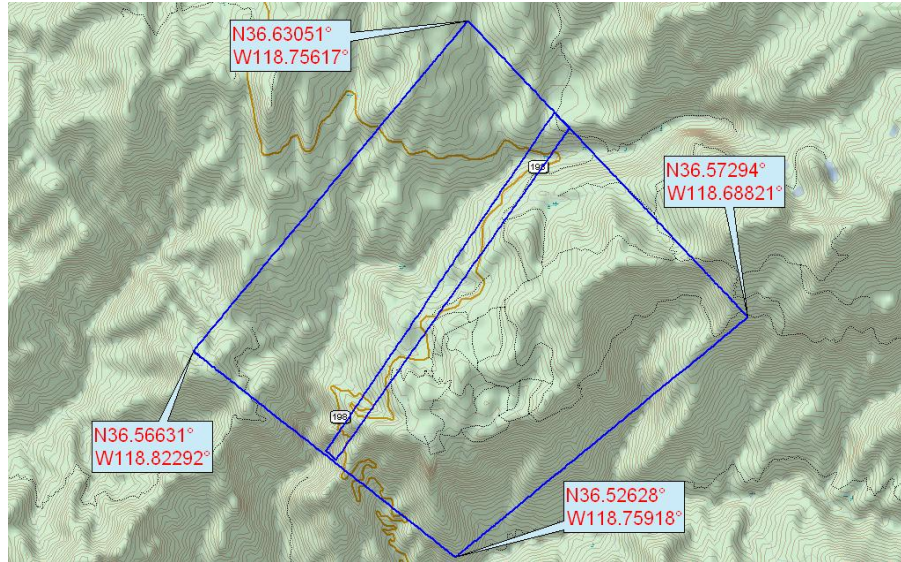


Figure 15. Sequoia National Park LiDAR data coverage

B. POST-PROCESSING AND DATA ANALYSIS SOFTWARE

Post-processing software is required to generate digital elevation models from LAS LiDAR data. Quick Terrain Modeler and its ILAP Bare Earth Extractor plug-in were used to generate the digital elevation models in this study. The Environment for Visualizing Images (ENVI) was used to analyze and compare these digital elevation models.

1. Quick Terrain Modeler (QTM) Version 6.0.6

Johns Hopkins University's Applied Physics Lab developed QTM to visualize large amounts of complex 3-D data. It can view models in various formats, such as QTC and QTT, and can import to and export from files such as GeoTiff DEMs and LAS. QTC or point-cloud files provide a good visualization of the extent of a survey by placing the points exactly where they belong without interpolation or approximation. QTT or surface

model files are used for visualizing terrain by laying out a regular grid across the survey area and placing a height value on all vertices. It builds a solid surface across this grid and the process involves approximation of data values (JHU/APL, 2007).

2. ILAP Bare Earth Extractor Version 1.0

The ILAP Bare Earth Extractor uses ASCII XYZ points representing foliated areas and separates them into surface, cloud, and object files. Surface file represents the estimated bare earth surface, cloud file represents the foliage, and object file represents the points that are non-surface, but whose heights above the estimated ground level fall below a user-specified limit (JHU/APL, 2006).

3. Environment for Visualizing Images (ENVI) Version 4.5

The Environment for Visualizing Images (ENVI), developed by ITT Visual Information Solutions (ITT VIS), is used for visualization, analysis, and presentation of digital imagery. It is utilized in this study to conduct image comparison and analysis of DEMs (in GeoTIFF) with varying density using its warp and mask tools, as well as its statistical computation algorithm.

C. METHODS

1. Generation of the Base Model

In order to visualize and choose the region of study, tiles or files in LAS format had to be converted to QTC and merged together using QTM. Here, the region of choice was selected and other LiDAR data was removed to obtain the reference model. The smaller reference model (1625 meters by 875 meters) in QTC format containing all the LiDAR data of the selected area is depicted in Figure 16.

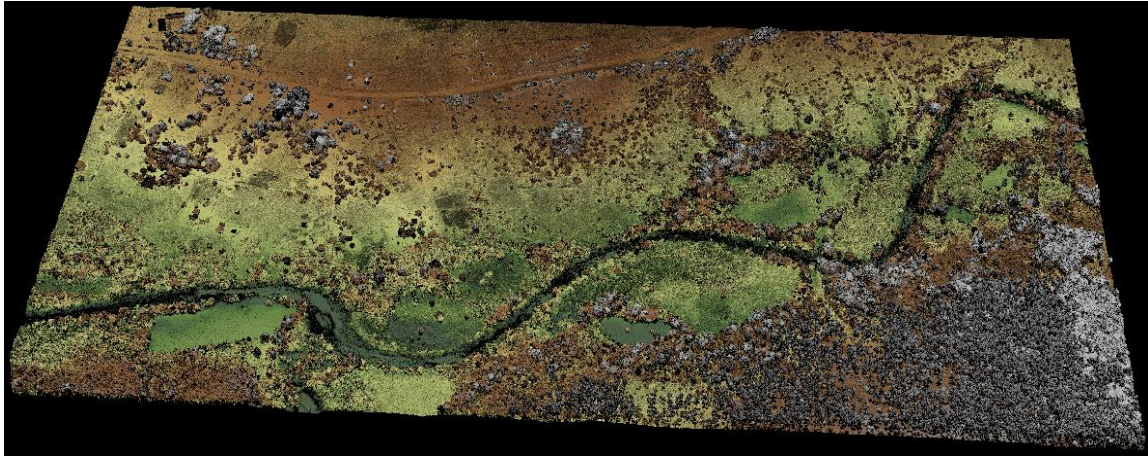


Figure 16. QTC format of the reference model (1625 meters by 875 meters)

2. Generation of Digital Elevation Models of Varying Densities

Once the smaller reference model (1625 meters by 875 meters) had been selected, it was ran through an IDL program written by Angie Puetz to produce a series of data sets containing 90%, 66%, 50%, 30%, 10%, 5%, 3%, 1%, 0.5%, 0.3%, 0.1%, 0.05%, 0.03%, and 0.01% of the smaller reference model. The IDL program randomly selects a number of LiDAR data points corresponding to the density desired and outputs them to a file in ASCII XYZ format. Since the IDL random program was written for ASCII XYZ file format with location x, y, z, and intensity values arranged in columns, the reference model in QTC format was first converted to ASCII XYZ with intensity values of each point using QTM. Each of the points in the QTC point-cloud format were converted into x, y, and z values in the 3-dimensional UTM coordinate. The IDL program is attached in the Appendix under the name `random_pts_fromXYZ_v2.pro`.

Each of the reduced data sets in ASCII XYZ format was used one at a time in QTM's ILAP Bare Earth Extractor plug-in to generate the DEMs.. All default parameters in the parameters box of the ILAP Bare Earth Extractor were used. However, in the Import Options menu, the "Import Surface File as Surface Model" was selected, the "Surface Model Sampling" was set to 1 meter, and the "Above Ground Level (AGL) Upper Limit" was set to "0" for each of the reduced data sets. These options removed all non-surface points and automatically imported the surface points as triangulated surface

or QTT model with a grid spacing of 1 meter (JHU/APL, 2007). Figure 16 shows the DEMs generated from the smaller reference model and from each of the reduced data sets using ILAP Bare Earth Extractor in QTT format. The DEMs below provided visual evidence that the DEMs generated from 0.3% and lower of the reference model distinctly lost most of the features of the DEM generated from the reference model and could be regarded as not useful.

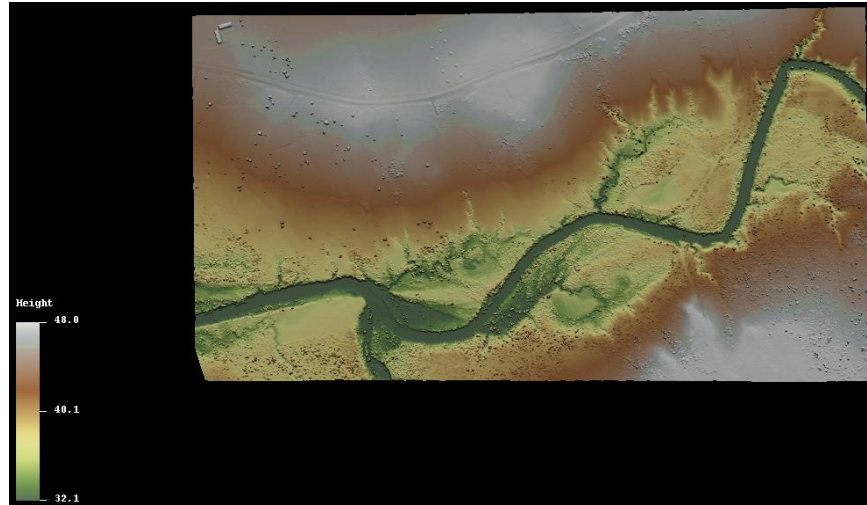
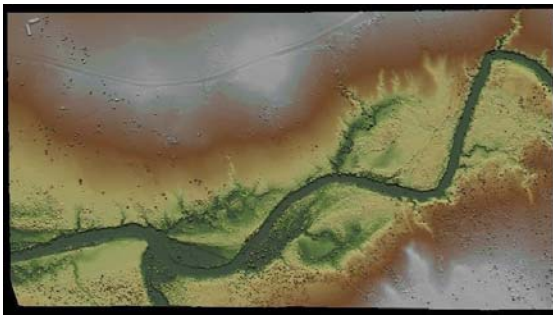
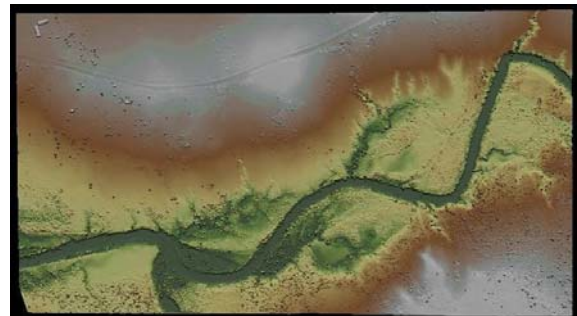


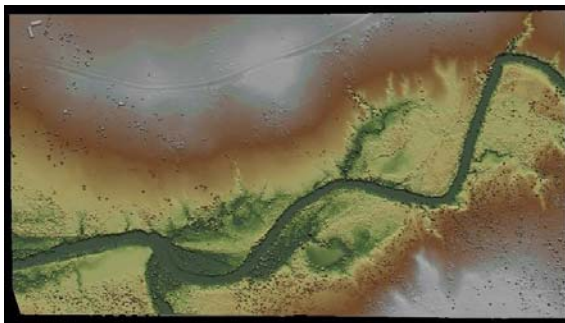
Figure 17. DEM generated from the reference model



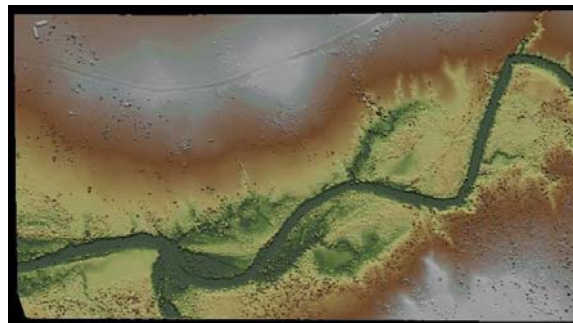
(a) DEM from 90% of the reference model



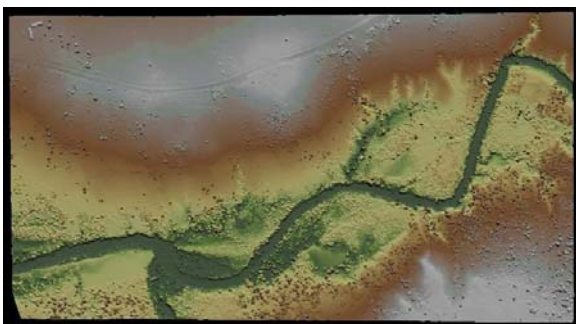
(b) DEM from 66% of the reference model



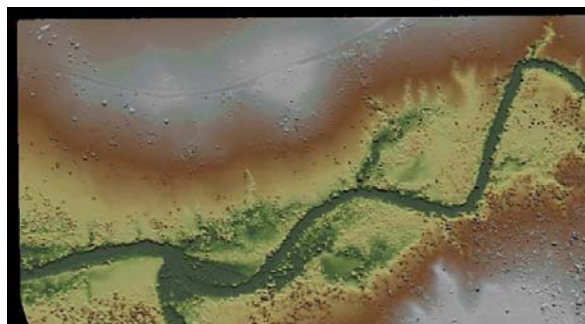
(c) DEM from 50% of the reference model



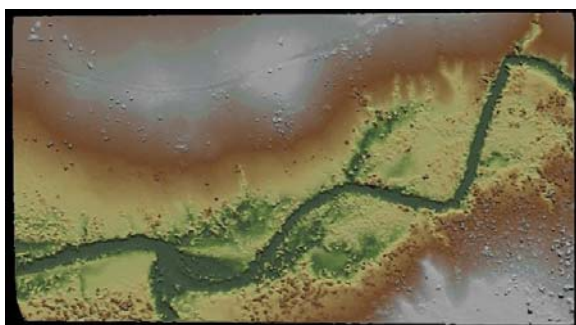
(d) DEM from 30% of the reference model



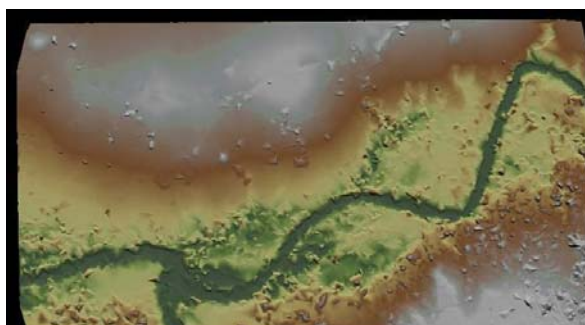
(e) DEM from 10% of the reference model



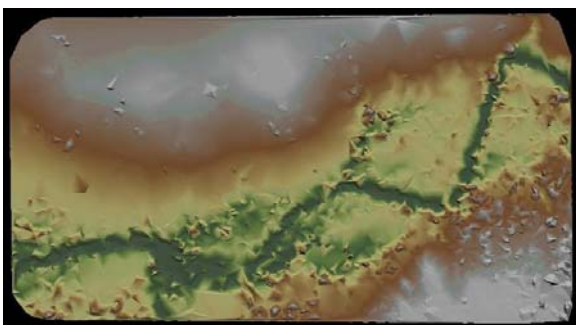
(f) DEM from 5% of the reference model



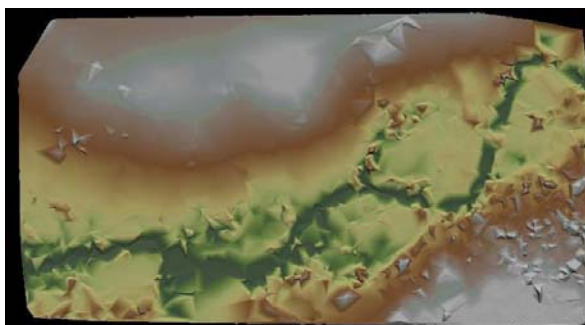
(g) DEM from 3% of the reference model



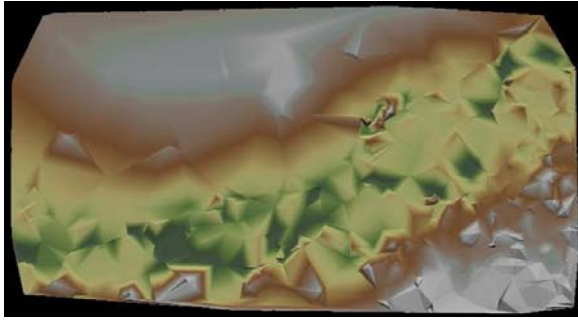
(h) DEM from 1% of the reference model



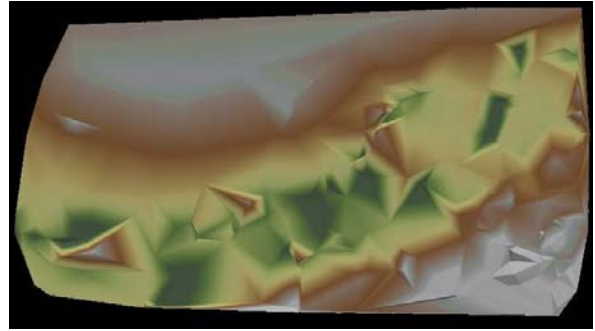
(i) DEM from 0.5% of the reference model



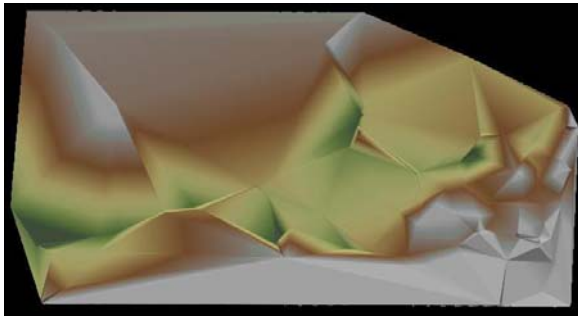
(j) DEM from 0.3% of the reference model



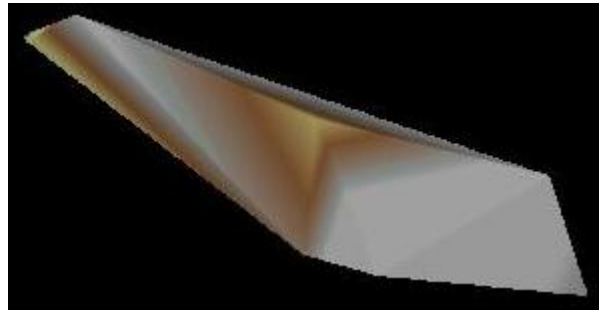
(k) DEM from 0.1% of the reference model



(l) DEM from 0.05% of the reference model



(m) DEM from 0.03% of the reference model



(n) DEM from 0.01% of the reference model

Figure 18. Digital Elevation Models in QTT format visualized using QTM software. DEMs from each of the reduced LiDAR data set of 1625 meters by 875 meters reference model.

The total number of points and densities of the reference model and the reduced data sets are tabulated in Table 2. The “Point Cloud Total Points” are obtained from each of the reduced data files generated using the IDL random_pts_fromXYZ_v2.pro program after converting them to QTC format from ASCII XYZ in QTM. The “Point Cloud density” is calculated by dividing the “Point Cloud Total Points” to the size of the reference model. The “Surface Points” are the points generated by the ILAP Bare Earth Extractor, which were divided by the size of the reference model to obtain the “Surface Point Density.”

Percent Reduction	Point Cloud Total Points	Surface Points	Point Cloud Density [pts/m ²]	Surface Point Density [pts/m ²]
100	8,622,343	4,282,495	6.0641	3.01186
90	7,760,109	3,964,226	5.4577	2.78803
66	5,690,746	2,937,141	4.0023	2.06568
50	4,311,171	2,245,747	3.0320	1.57943
30	2,586,703	1,368,997	1.8192	0.96281
10	862,234	421,183	0.6064	0.29622
5	431,117	158,899	0.3032	0.11175
3	258,670	72,203	0.1819	0.05078
1	86,223	15,123	0.0606	0.01064
0.5	43,111	6,936	0.0303	0.00488
0.3	25,867	3,674	0.0182	0.00258
0.1	8,622	1,002	0.0061	0.00070
0.05	4,311	364	0.0030	0.00026
0.03	2,586	127	0.0018	0.00009
0.01	862	9	0.0006	0.00001

Table 2. Number of Points and Densities of each DEM

IV. ANALYSIS

The data collected over Mocoron, Honduras were used to generate digital elevation models and were analyzed in two segments below, followed by analysis of the data collected over Sequoia National Park.

A. PREPARATION OF DIGITAL ELEVATION MODELS FOR ANALYSIS

1. Warping

Prior to conducting DEM analysis in ENVI, each DEM must match geographically. ENVI's "Rubber Sheet Warp" tool was used to geographically map each DEM to the DEM generated from the reference model. A set of 10 Ground Control Points (GCP's) were selected on the DEM generated from the reference model to be used in the process of warping DEMs generated from 90% to 0.03% of the reference model. Since the DEM generated from 0.01% of the smaller reference model contain a significantly small amount of data and, therefore, did not encompass the GCP's previously selected, a different set of 6 GCP's were used.

The parameters used in creating the warped files are listed in Figure 19. For a polynomial method of warping, the required number of GCP's must be greater than the squared quantity of the degree of polynomial plus one ($\# \text{ of GCPs} > (\text{deg} + 1)^2$) (ENVI help). The background or the areas that did not contain any data was set to -9999.0 to differentiate them from the areas that contain data. All the DEMs generated from 90% to 0.01% of the smaller reference model were warped in the same dimension as the DEM generated from the reference model, which is specified in the "Output Image Extent Option." Figure 20 shows the outcome of using the Warp Tool with the stated parameters. The image on the left is the DEM generated from the smaller reference model and the image on the right is the DEM generated from 0.01% of the reference model. The red circles show the set of 6 GCPs used to warp the image mentioned above.

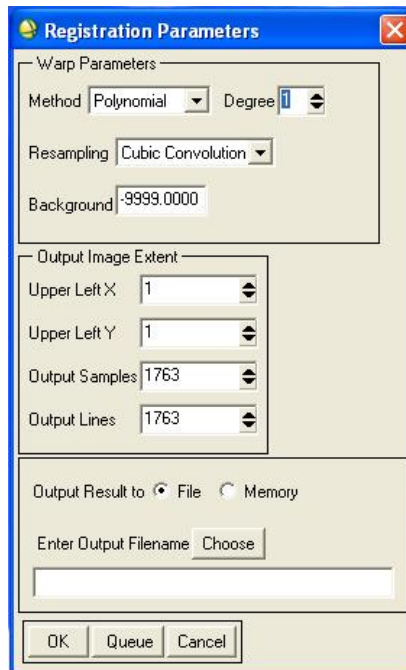


Figure 19. Warp parameters

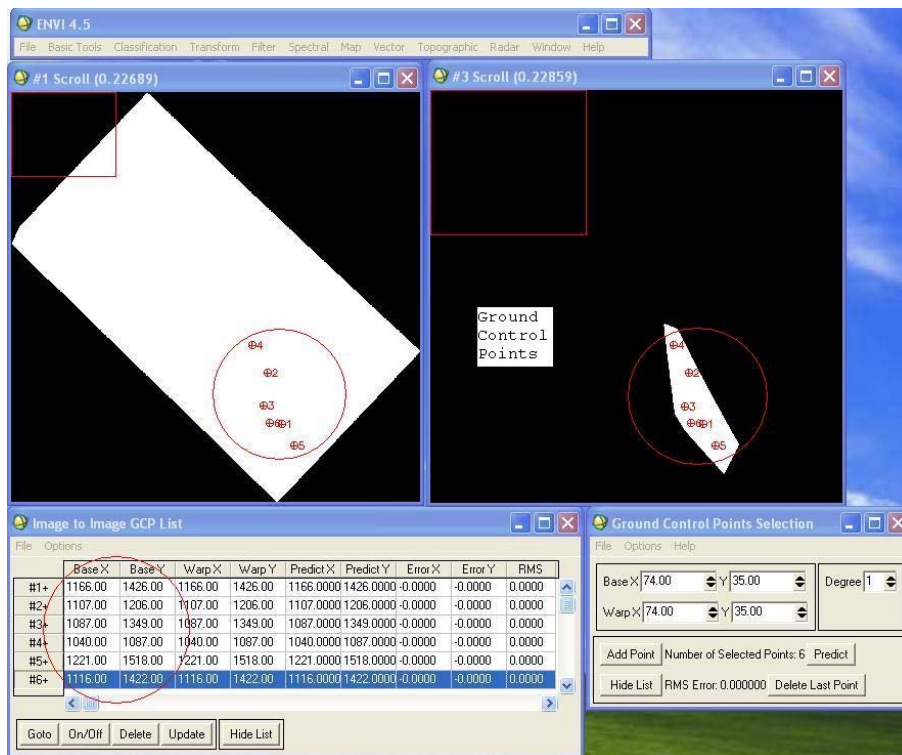


Figure 20. Warped DEM from 0.01% of the smaller reference model (right image)

2. Masking

As Figure 20 shows, the warped DEM generated from the 0.01% of the reference model covers an enormous amount of area (black), which does not contain any data. These areas would skew the DEM accuracy analysis results of the lower resolution DEM, and therefore must not be included in the analysis. The “Mask” tool in ENVI was used to analyze just the data contained in all the warped, lower resolution DEMs. A mask was built based on the warped DEM containing the least amount of data (0.01%) and was applied to the rest of the DEMs including the one generated from the smaller reference model. Figure 21 shows the results after the mask was applied to the DEM generated from the smaller reference model (left) and the warped DEM generated from the 0.01% of the smaller reference model (right).

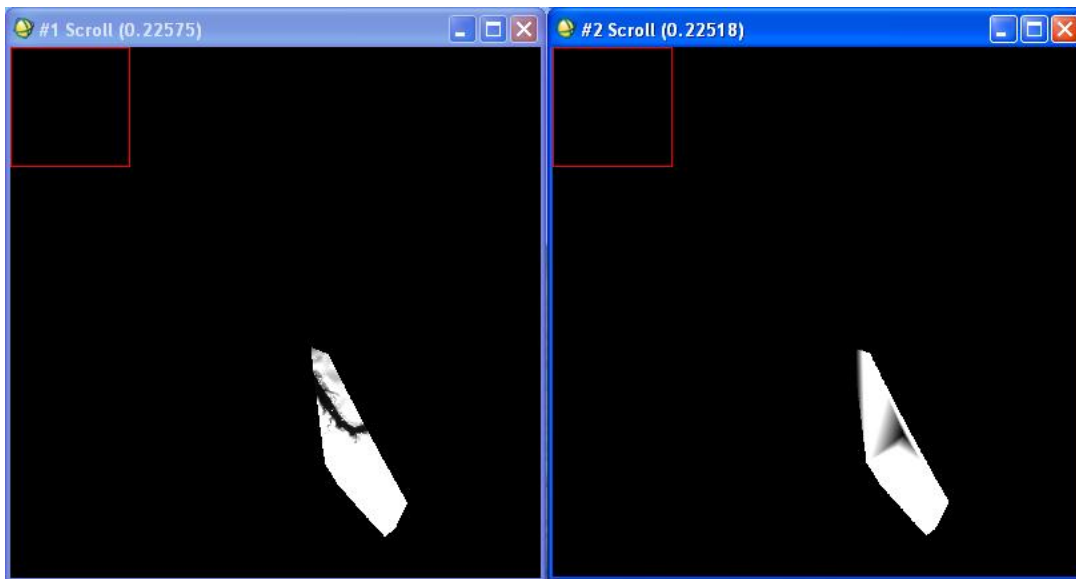


Figure 21. 100% (left) and 0.01% (right) data DEMs after mask has been applied

B. CORRELATION ANALYSIS IN ENVI

The first analysis effort was designed to replicate a previous study conducted by Anderson [2008].

1. Reference Model of Size 1625 Meters by 875 Meters (Smaller Reference Model)

The statistics function in ENVI was used to determine the correlation factors of each of the DEMs generated from 90% to 0.01% of the smaller reference model against the DEM generated from the smaller reference model. A semi log plot of correlation factor versus point density is displayed in Figure 22. It shows that the correlation is non-linear between the DEM generated from the reference model and the rest of the DEMs. The plot shows a slight decrease in correlation until a certain point cloud density, from which the correlation starts to decrease drastically.

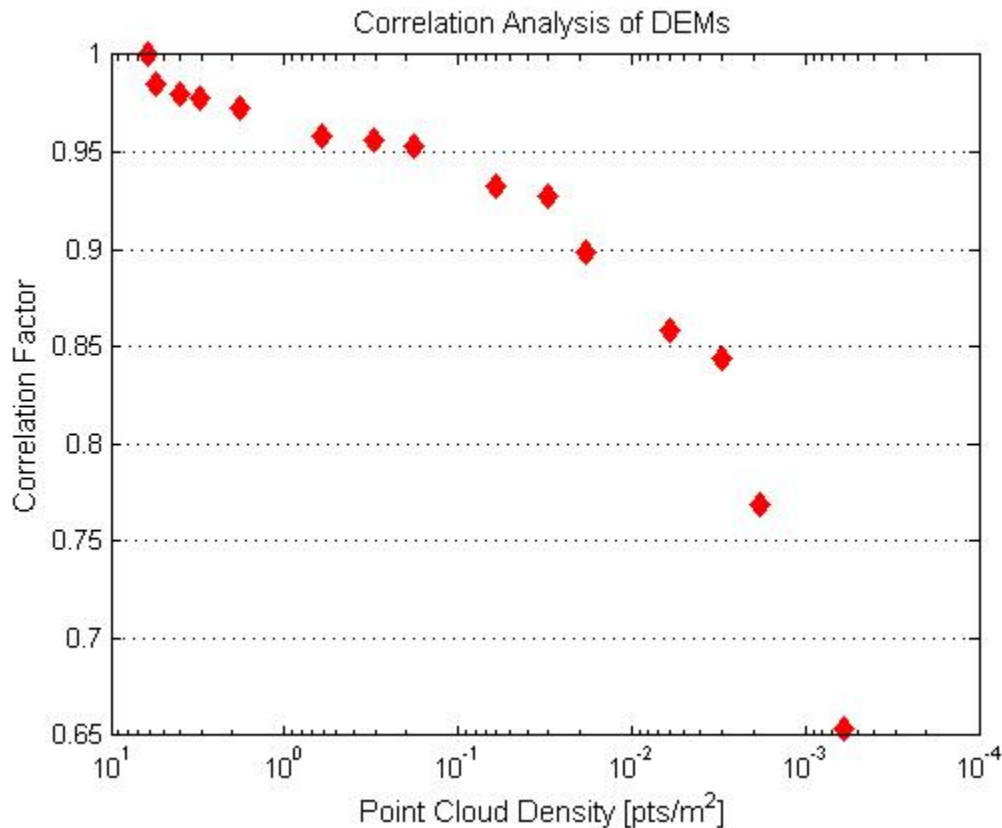
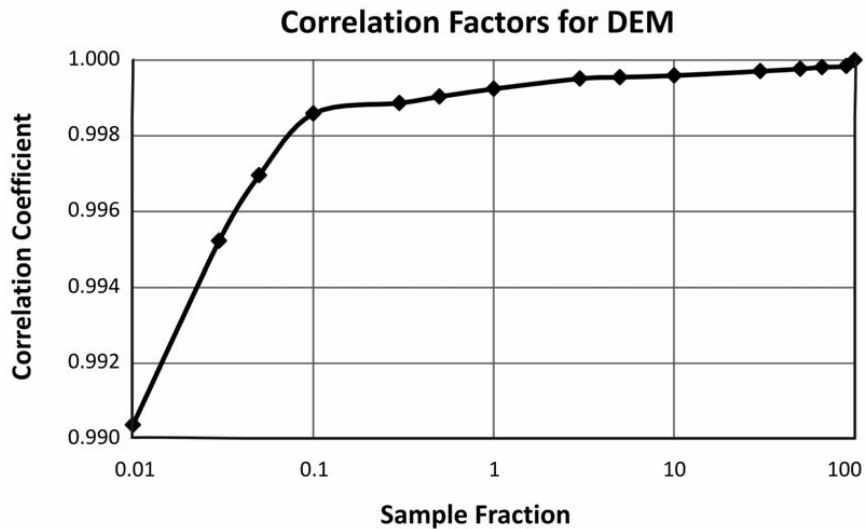


Figure 22. Correlation of each DEM to the DEM generated from the smaller reference model. These are the DEMs depicted in Figures 17 and 18.

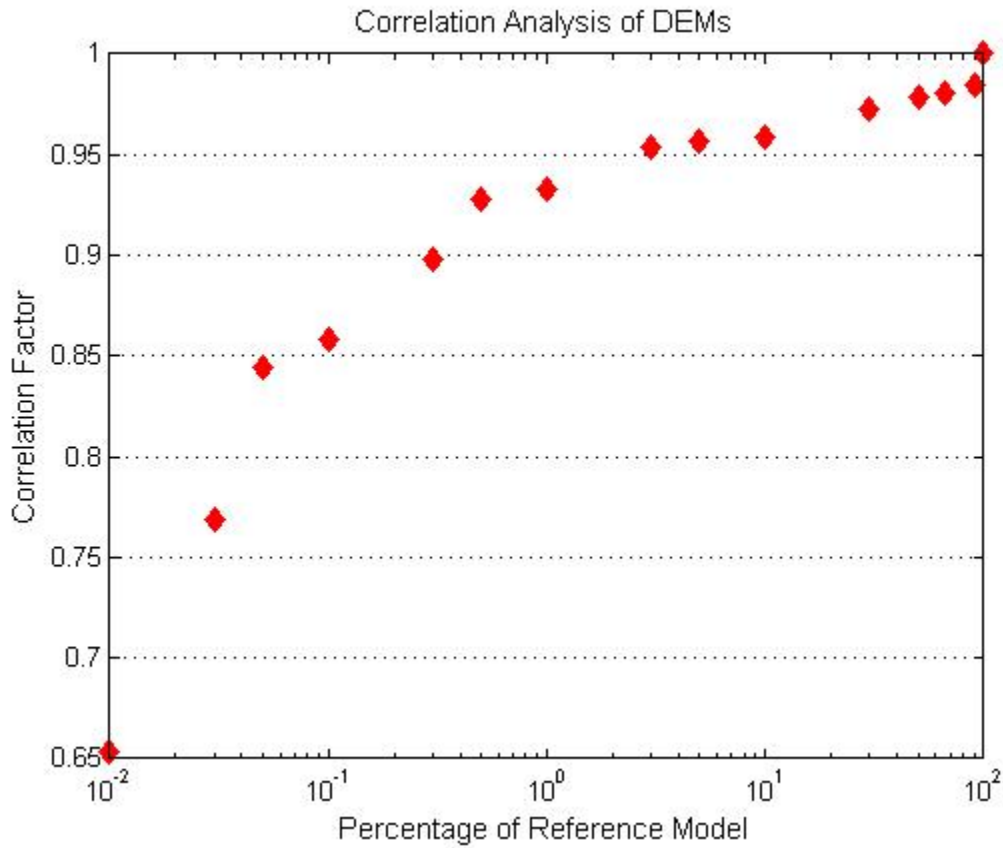
Differences were found between the results obtained from this study and the ones obtained from the study conducted by Anderson [2008]. The region or area selected between this study and the previous study by Anderson [2008] as well as percentage

reductions, are the same. The IDL random selection program to reduce the reference model from 90% to 0.01% has the same algorithm written by the same person, Angie Puetz. The post processing software (QTM and ILAP Bare Earth Extractor) and the software (ENVI) used to determine the correlation factors are the same.

The previous study conducted by Anderson [2008] showed a similar nonlinear decreasing trend in correlation of the DEMs from 90% to 0.01%; however, the lowest correlation factor is 0.99 for the 0.01% as opposed to 0.65 for the 0.01% of this study. The correlation factors of the previous study were very high between the DEM generated from the reference model and the DEMs generated from 90% to 0.01% of the reference model. A plot comparison is illustrated in Figure 23. Figure 23a shows the correlation results of the previous study by Anderson plotted against percent reduction in semi log plot, and Figure 23b shows the plot of the results obtained from this study. The correlation factors obtained from this study were plotted against the corresponding percent reductions of the point densities used in Figure 22 and the x-axis values are now increasing for the purpose of comparing the previous study and this study.



(a)



(b)

Figure 23. Different correlation factors obtained from two similar studies. (a) Results of previous study (Anderson, 2008). (b) Results from this study.

The dissimilarities in the correlation results of the two plots in Figure 23 motivated further investigation. DEMs were generated from additional random subset of the same reference model (1625 meters by 875 meters). The process was identical. More DEMs were generated from the 0.1% data due to the wide distribution of correlation factors compared to the distribution of the correlation factors obtained from the DEMs generated from 3% and 0.3% of the smaller reference model. Figure 24 shows the additional DEMs produced an interesting result and provides a visual representation of the “uncertainties” or variations in DEM generation process. These “uncertainties” become larger as the reference model is reduced from 3% to 0.1% of its original point-cloud density. Looking at the trend of “uncertainties” or range of errors in Figure 24,

these data can be extended and assessed to become smaller as the percentage of the original data increased, and become even larger as the percentage of the original data decreased to 0.01%. A reasonable conclusion is that the differences in Figures 23a and 23b are due to the effect of random subset generation.

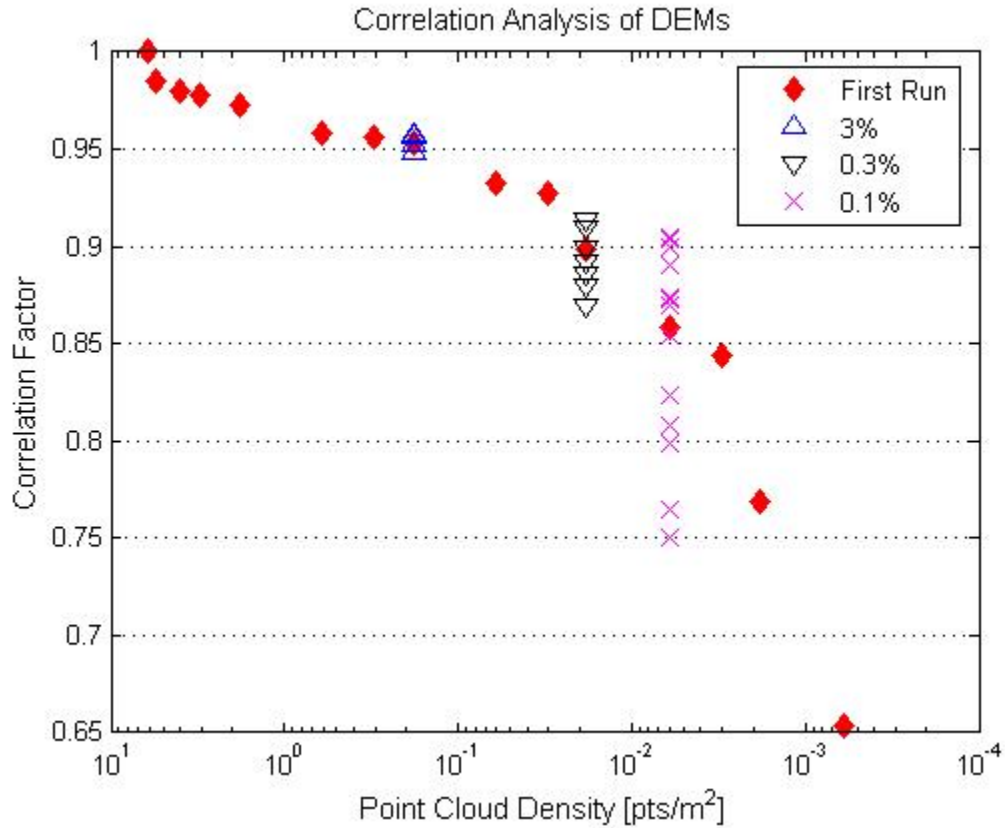


Figure 24. Correlation factors of the first set of DEMs (1625 meters by 875 meters) labeled “First Run” and the additional DEMs generated from 3%, 0.3%, and 0.1% of the smaller reference model

An additional plot (Figure 25) shows the distribution of the correlation factors obtained from the DEMs generated from 0.1% of the smaller reference model. The y-axis shows the sequence of the generated DEMs and the x-axis contains the corresponding correlation factors. The distribution has a mean of 0.84862 with a minimum at 0.75 and a maximum at 0.92 correlation factors.

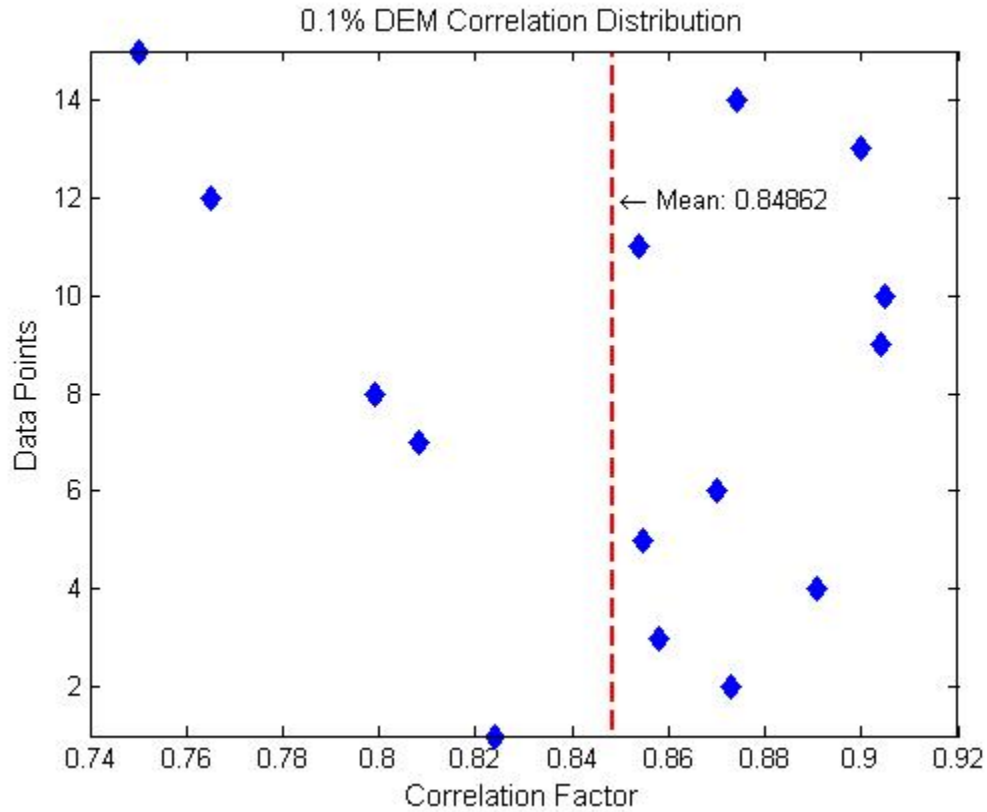


Figure 25. Calculated Mean Correlation Factor of DEM generated from 0.1% of 1625 meters by 875 meters reference model.

Figure 25 also suggests that a DEM generated from at least 0.3% of the reference model (0.0182 pts per square meter point cloud density) may be used as the lowest density of points required for DEM generation using LiDAR data. The 0.3% data has a mean of 0.89213 and a maximum uncertainty of 2.56%. The same result (0.0182 pts per square meter point cloud density) was obtained by Liu as shown in Figure 13. But when the percentage of surface points were plotted against the point-cloud density of each of the reduced data set as illustrated in Figure 26, a DEM generated from at least 0.6064 points per square meter (10% of the reference model) must be used as the lowest density for DEM generation using LiDAR data. Further decrease in density causes the Bare Earth Extractor Plug-in to classify a smaller percentage of points as surface points. Using the Bare Earth Extractor Plug-in, 49.67% of the points of the reference model were classified as surface points. This was fairly consistent (increasing slightly as the density

of the smaller reference model drops from 100% to 30%) as the reference model was reduced to 0.6064 points per square meter or 10% of its original data. The calculated surface point percentages are tabulated in Table 3.

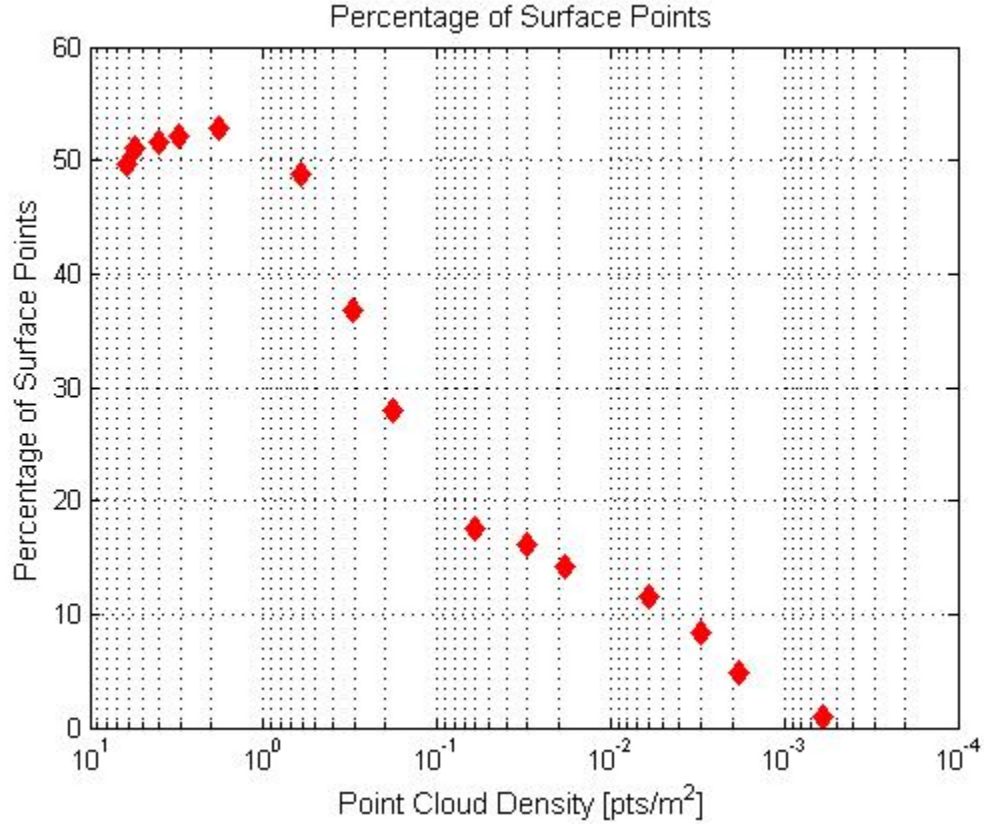


Figure 26. Percentage of Total Point Cloud classified as Surface Points

Percentage of Original Point Cloud Data	Density of Point Cloud Data	Percentage of Point Cloud Data Classified as Surface Points
100	6.0641	49.67
90	5.4577	51.08
66	4.0023	51.61
50	3.0320	52.09
30	1.8192	52.92
10	0.6064	48.85
5	0.3032	36.86
3	0.1819	27.91
1	0.0606	17.54
0.5	0.0303	16.09

0.3	0.0182	14.20
0.1	0.0061	11.62
0.05	0.0030	8.44
0.03	0.0018	4.91
0.01	0.0006	1.04

Table 3. Percentage of Total Point Cloud of the reference model classified as Surface Points

2. Reference Model of Size 3005 Meters by 844 Meters (Larger Reference Model)

Moving beyond the previous study, a larger area of the LiDAR data collected over Honduras was selected using the same method used in the smaller reference model. The larger area contained within it the smaller reference model. Again, using the IDL program `random_pts_fromXYZ_v2.pro`, QTM, and ILAP Bare Earth Extractor, another set of DEMs were generated from 90% to 0.01% of a larger reference model. Figure 27 shows the DEMs generated in QTT format from the larger reference model and Figure 28 shows the ones generated from each of the reduced data sets. Due to a higher number of points compared to the first set of DEMs generated from the smaller reference model, a DEM generated from the 0.01% of the larger reference model produced a larger dimension.

The new set of DEMs below provided visual evidence that the DEMs generated from 0.05% (0.0066 points per square meters) and lower of the larger reference model distinctly lost most of the features of the DEM generated from the 100% of the larger reference model. The number of points and densities of these set of DEMs are summarized in Table 4. The larger reference model contains many more points than the previous reference model (a ratio of about 4 to 1).

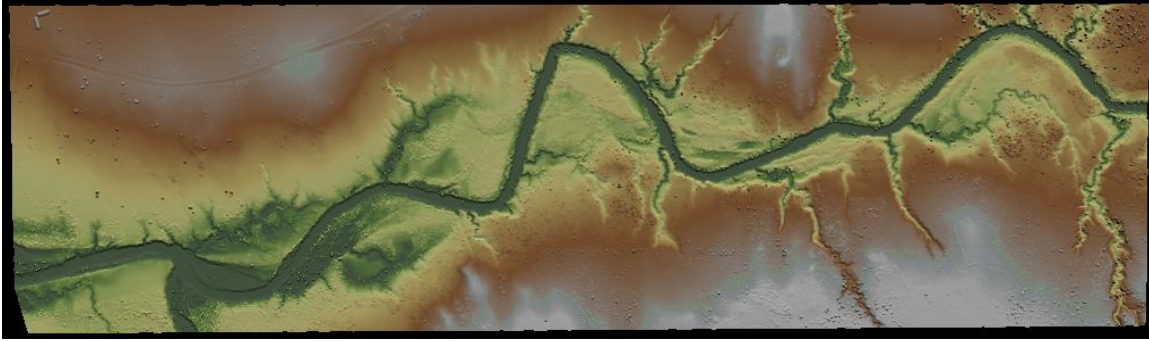
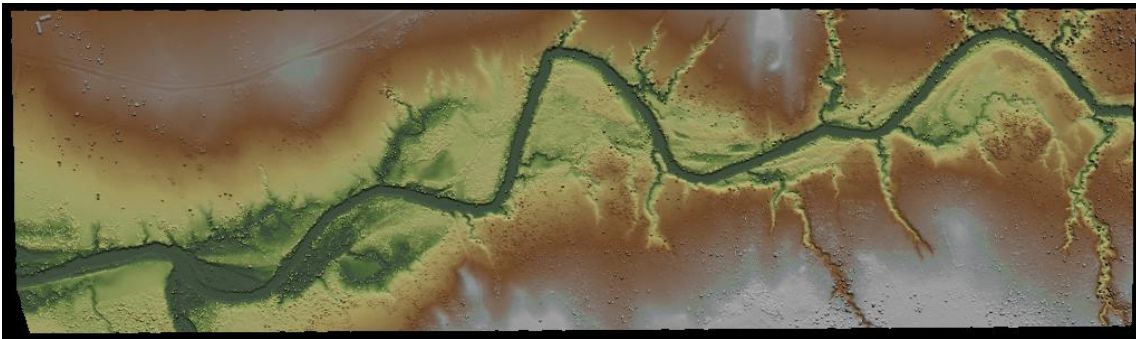
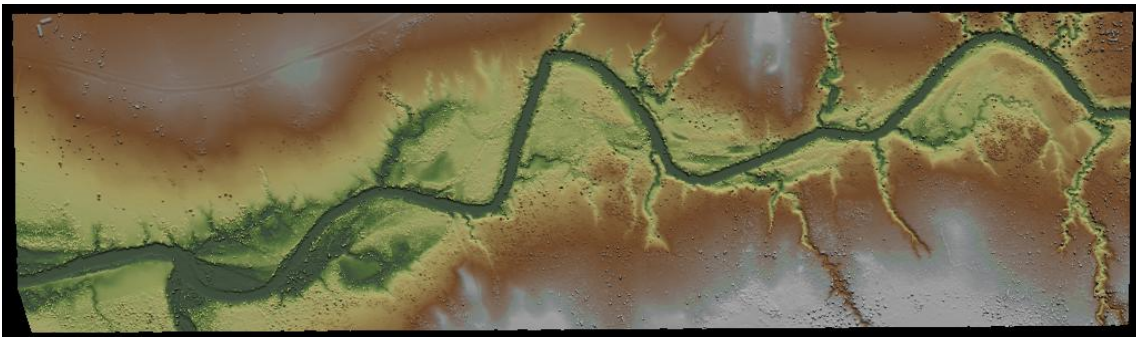


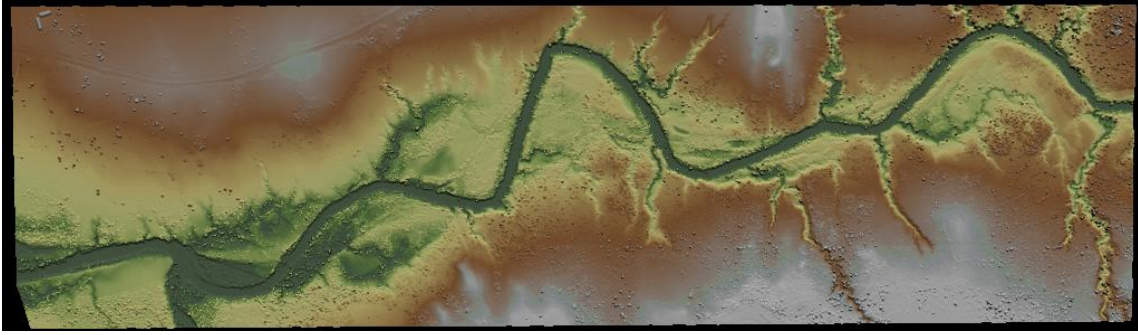
Figure 27. DEM generated from the larger (3005 meters by 844 meters) reference model of Mocoron, Honduras.



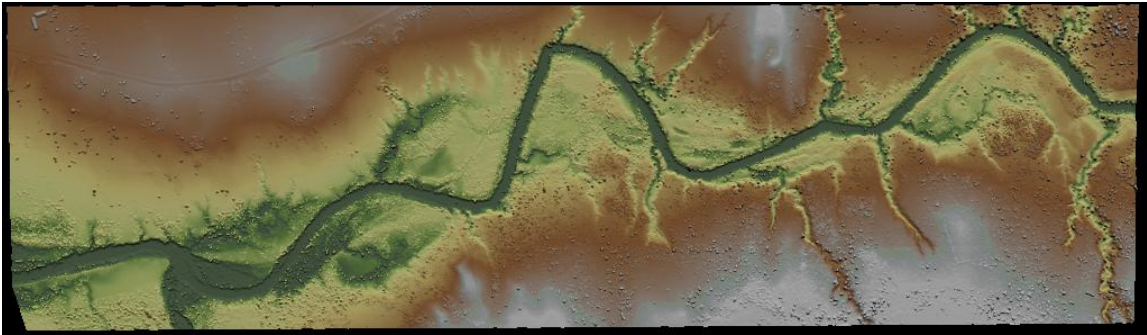
(a) DEM from 90% of the larger reference model



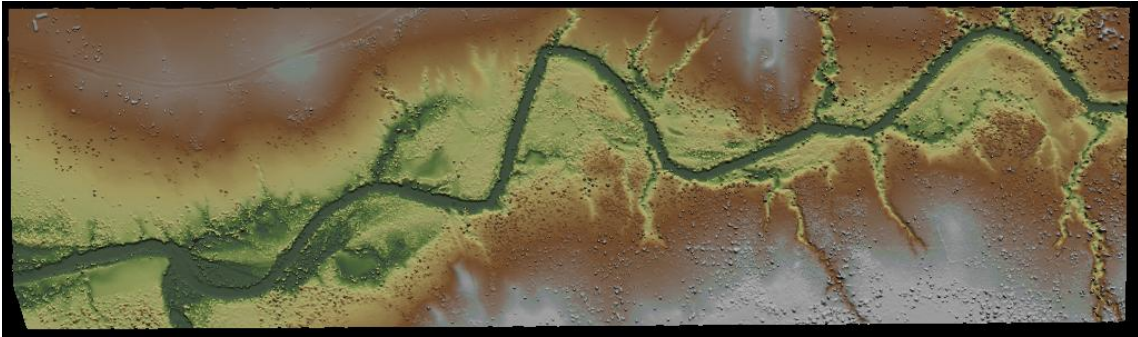
(b) DEM from 66% of the larger reference model



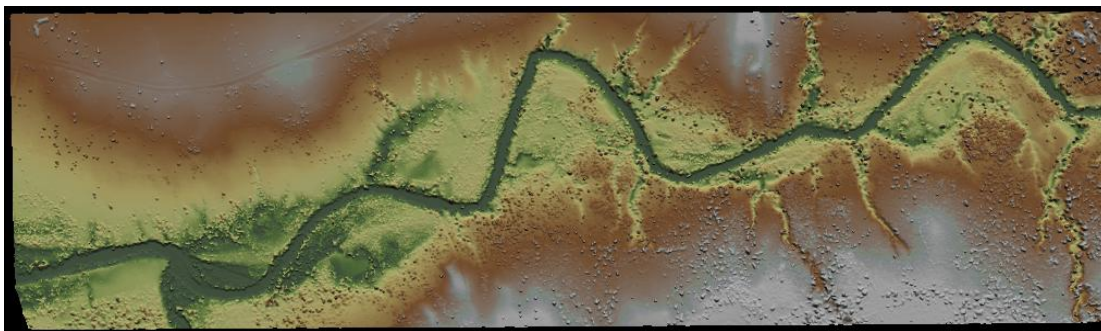
(c) DEM from 50% of the larger reference model



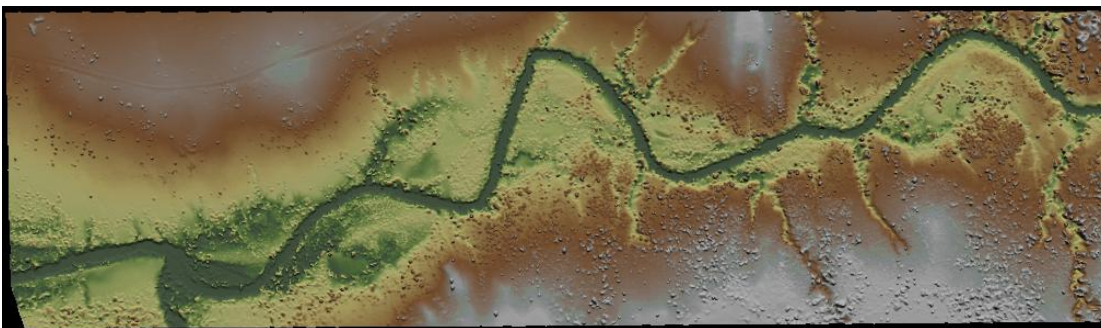
(d) DEM from 30% of the larger reference model



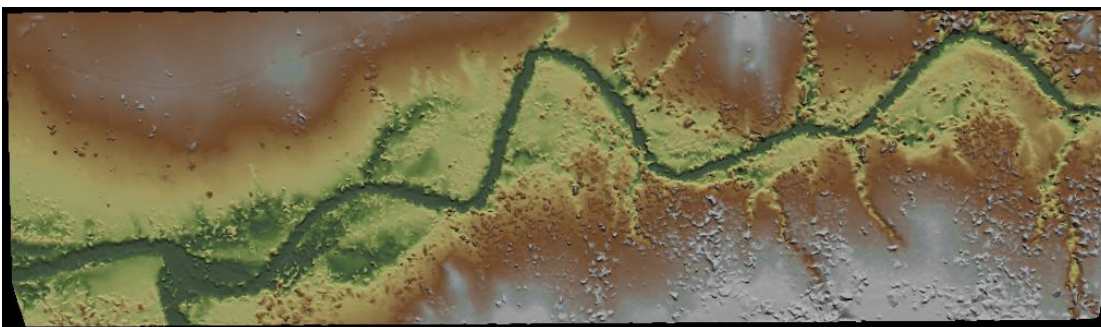
(e) DEM from 10% of the larger reference model



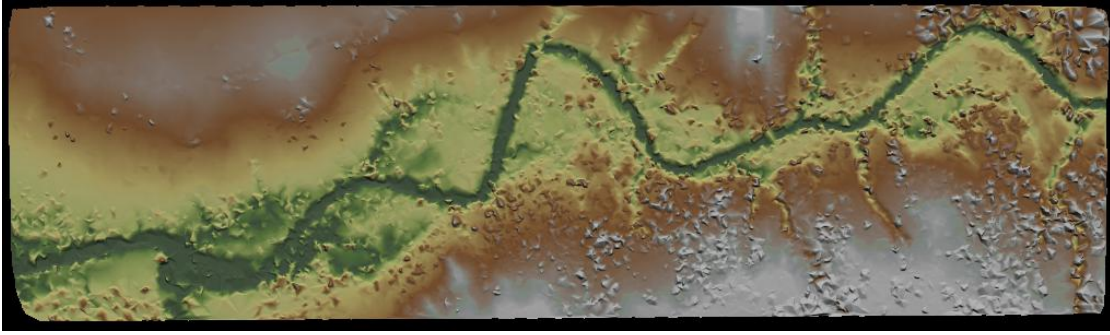
(f) DEM from 5% of the larger reference model



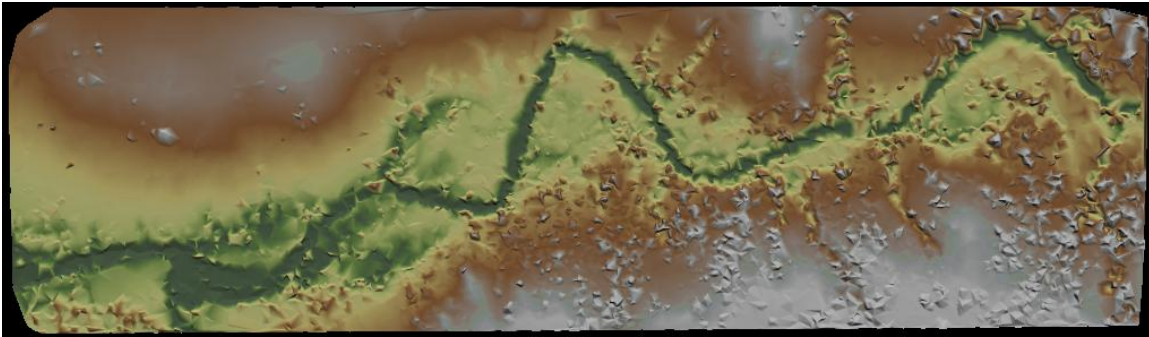
(g) DEM from 3% of the larger reference model



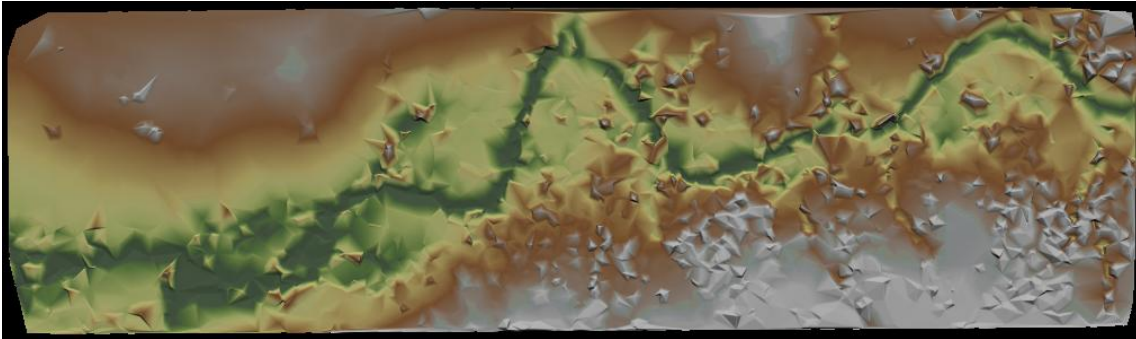
(h) DEM from 1% of the larger reference model



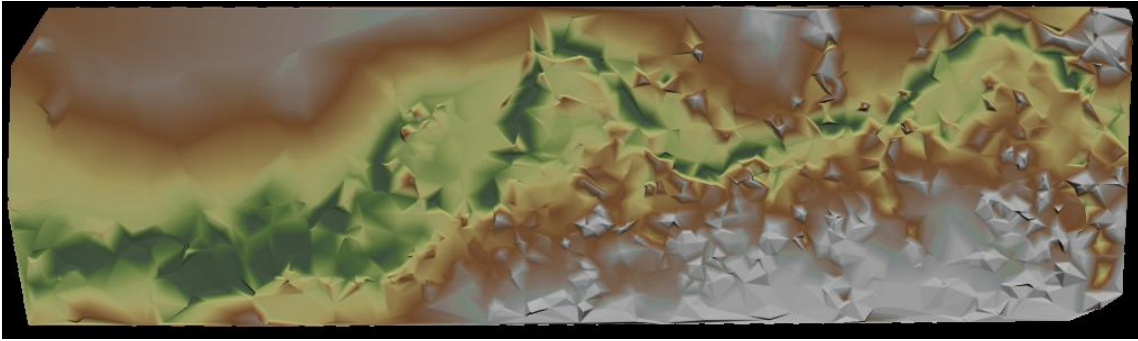
(i) DEM from 0.5% of the larger reference model



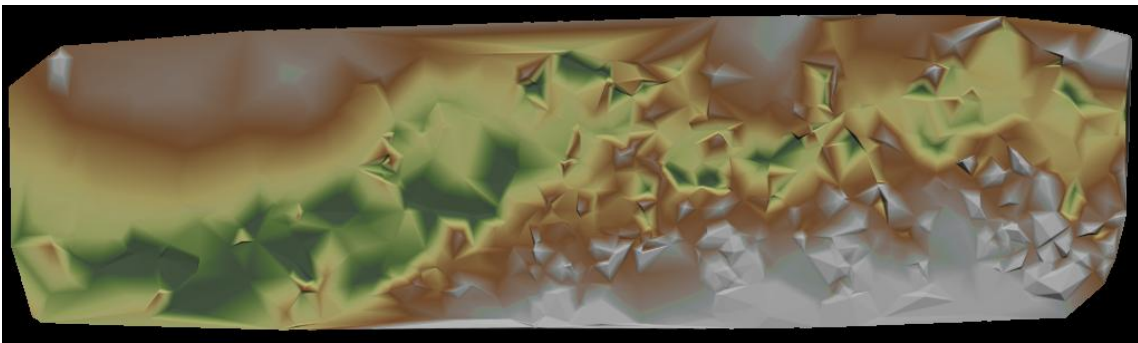
(j) DEM from 0.3% of the larger reference model



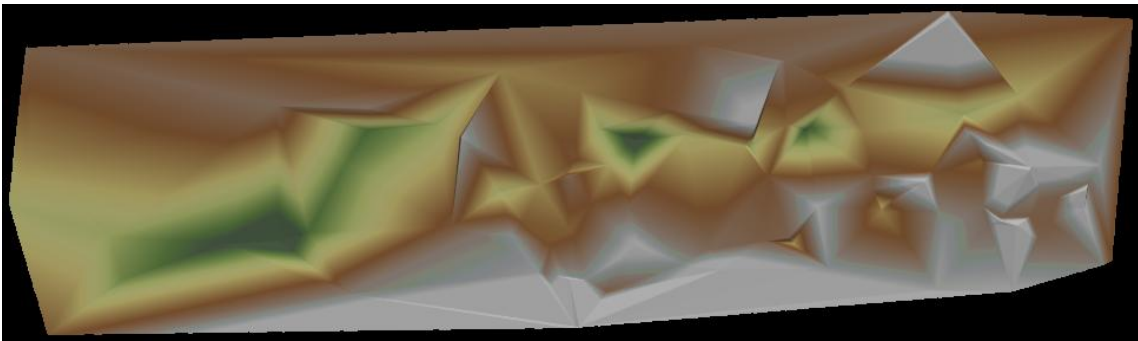
(k) DEM from 0.1% of the larger reference model



(l) DEM from 0.05% of the larger reference model



(m) DEM from 0.03% of the larger reference model



(n) DEM from 0.01% of the larger reference model

Figure 28. Digital Elevation Models in QTT format visualized using QTM software. DEMs from each of the reduced LiDAR data set of 3005 meters by 844 meters reference model.

Percent Reduction	Point Cloud Total Points	Surface Points	Point Cloud Density [pts/m ²]	Surface Point Density [pts/m ²]
100	33,485,312	11,687,678	13.2028	4.6083
90	30,136,780	11,040,747	11.8826	4.3532
66	22,100,306	8,191,138	8.7139	3.2297
50	16,742,656	6,281,736	6.6014	2.4768
30	10,045,594	3,863,111	3.9609	1.5232
10	3,348,531	1,332,758	1.3203	0.5255
5	1,674,265	625,334	0.6601	0.2466
3	1,004,559	319,503	0.3961	0.1260
1	334,853	71,702	0.1320	0.0283
0.5	167,426	29,911	0.0660	0.0118
0.3	100,455	17,851	0.0396	0.0070
0.1	33,485	5,069	0.0132	0.0020
0.05	16,742	2,532	0.0066	0.0010
0.03	10,045	1,216	0.0040	0.0005
0.01	3,348	132	0.0013	0.0001

Table 4. Number of Points and Densities of each DEM generated from a larger reference model.

The same statistical analysis in ENVI as the smaller reference model was applied to these new set of DEMs. The correlation results of these new set of DEMs were plotted with the first set of DEMs (Figure 29). The results were consistent except for the DEM generated from the 0.01% data. The larger reference model shows the familiar decreasing trend of correlation results from DEMs generated from 90% to 0.01% of the smaller reference model. Each of the correlation results from the new set of DEMs are higher compared to the first set of DEMs due to the higher density of points contained in each of the DEMs with the exception of the one generated from 0.01% of the reference models. This can be explained by the uncertainties obtained from the first set of DEMs as shown in Figure 24.

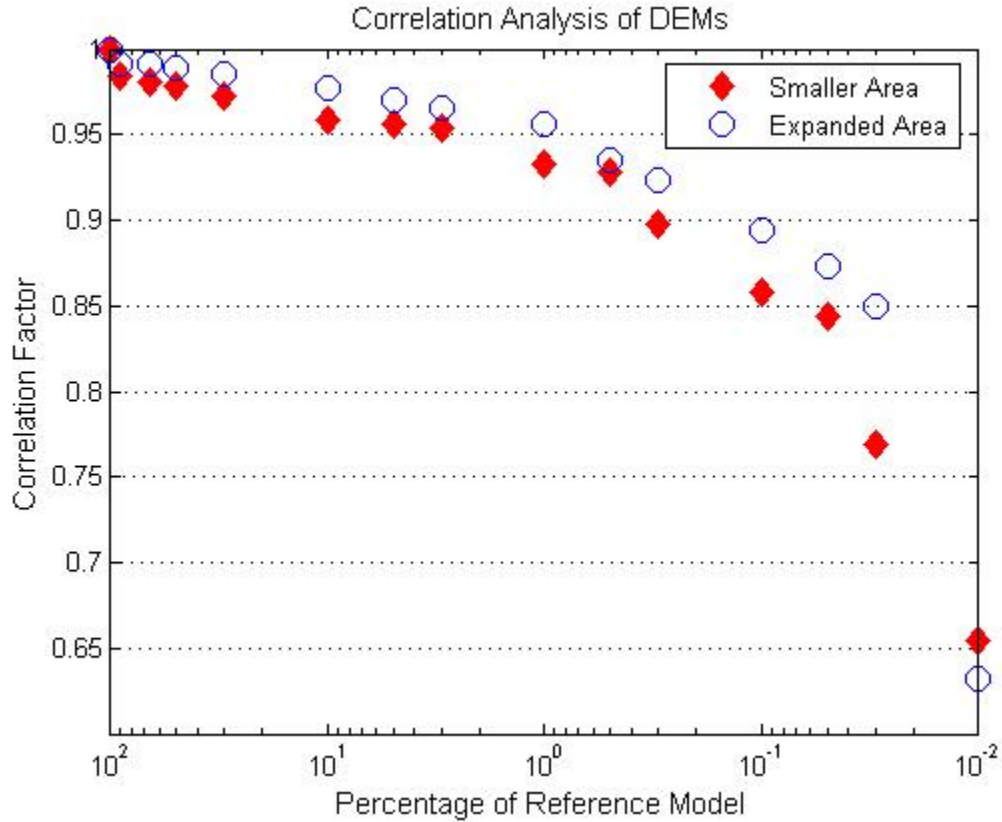


Figure 29. Correlation analysis of DEMs generated from the larger reference model plotted on the same plot with the first set of DEMs generated from the smaller reference model.

The percentage of points classified as surface points for these new set of DEMs were also plotted with the first set of DEMs (Figure 30). The curves of the first set of DEMs and the new set of DEMs are very similar; however, ILAP Bare Earth Extractor only classified 34.90% of the points of the reference model as surface points (vice 49.67% of the smaller reference model). This 34.90% surface point was fairly consistent (increasing slightly as the density of the larger reference model drops from 100% to 10%) as the density of the larger reference model was reduced to 0.6601 points per square meter. The result is similar to the one obtained from the smaller reference model, which was 0.6064 points per square meter. The calculated surface point percentages for these new set of DEMs are tabulated in Table 5.

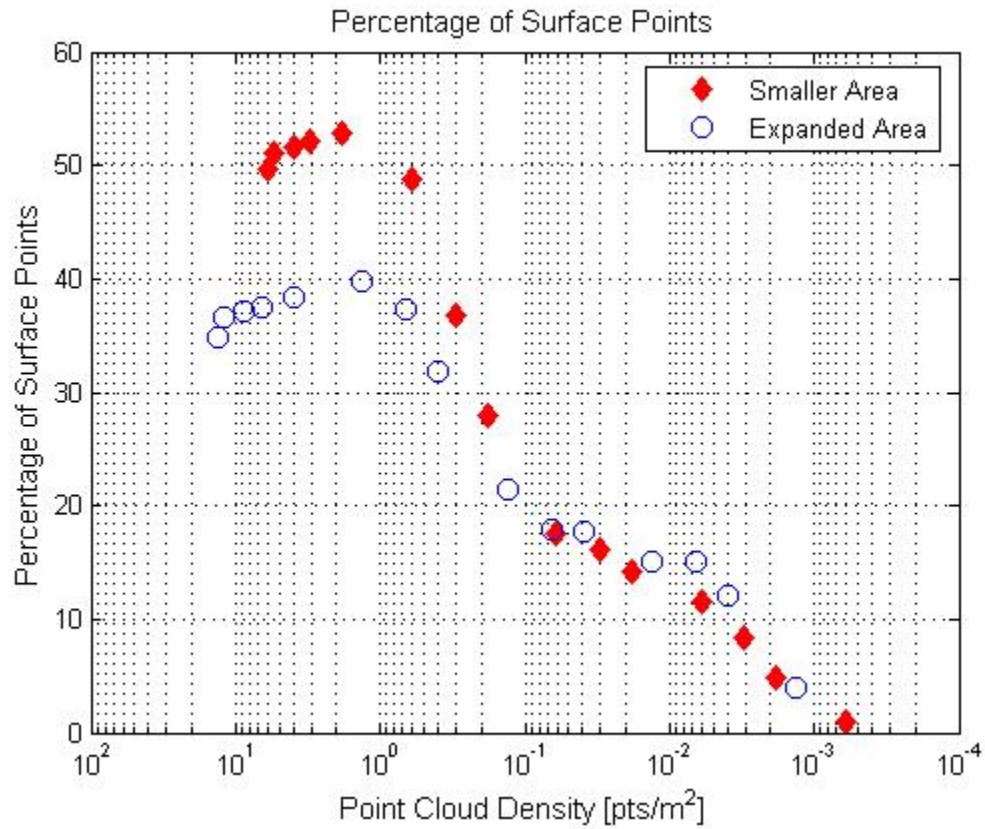


Figure 30. Percentage of Classified Surface Points with Diminishing Point Cloud Density. First set of DEMs (Red diamonds) and new set of DEMs (blue circle)

Percentage of Original Point Cloud Data	Density of Point Cloud Data	Percentage of Point Cloud Data Classified as Surface Points
100	13.2028	34.90
90	11.8826	36.64
66	8.7139	37.06
50	6.6014	37.52
30	3.9609	38.46
10	1.3203	39.80
5	0.6601	37.35
3	0.3961	31.81
1	0.1320	21.41
0.5	0.0660	17.87
0.3	0.0396	17.77
0.1	0.0132	15.14
0.05	0.0066	15.12
0.03	0.0040	12.11
0.01	0.0013	3.94

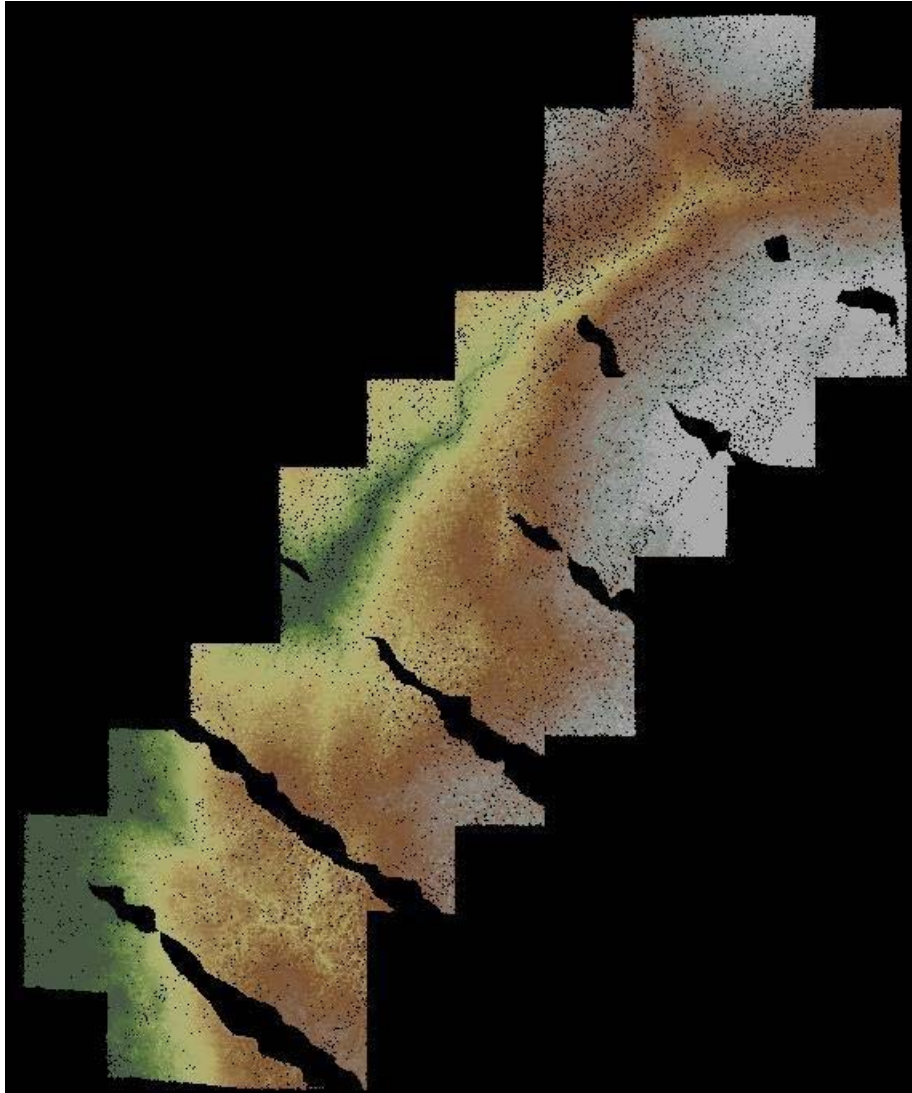
Table 5. Percentage of Total Point Cloud of the larger reference model classified as Surface Points

C. VALIDATION OF DECIMATION APPROACH

The Sequoia data collection effort included a special flight line, with a restricted mirror sweep, low altitude flight line. This resulted in a higher density flight line. Figures 30a and 30b show the high resolution and the standard resolution data acquired in QTC format, respectively.



(a)



(b)

Figure 31. High (a) and standard (b) resolution LiDAR data of Sequoia National Park visualized in QTC format using QTM

An area which overlapped both the high and standard resolution data were selected and used for analysis (Figure 32). The densities of the high and standard resolution data were 1.3984 points per square meter and 1.1516 points per square meter, respectively. The high resolution data contains 13.2% more points than the standard

resolution data. The selected high resolution data set was reduced to 86.8% of its original density to match the density of the standard resolution data set.

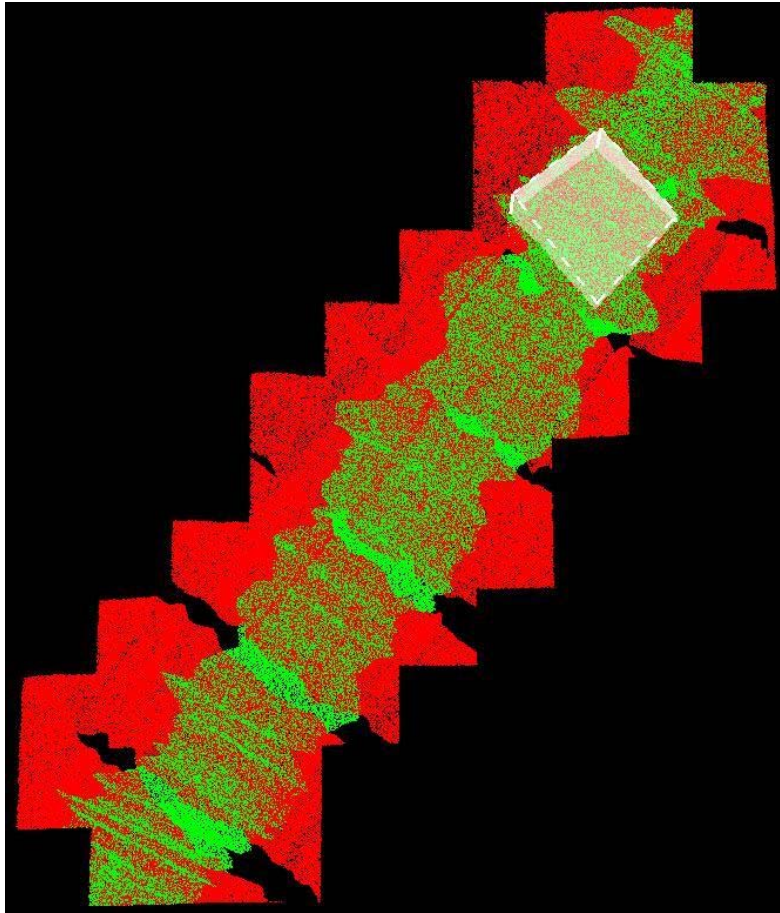
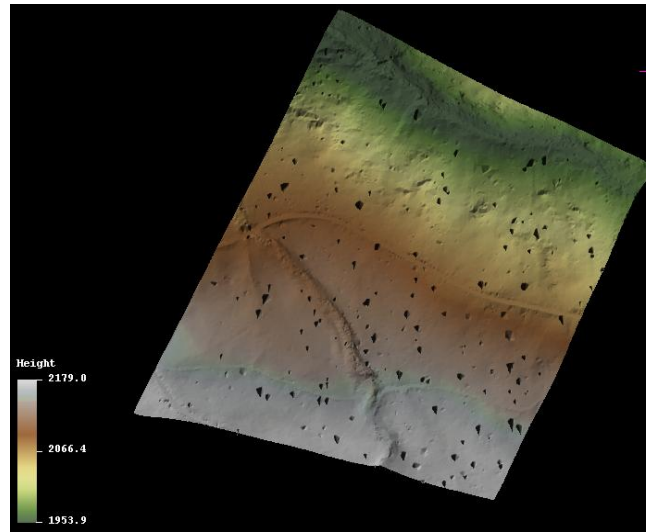
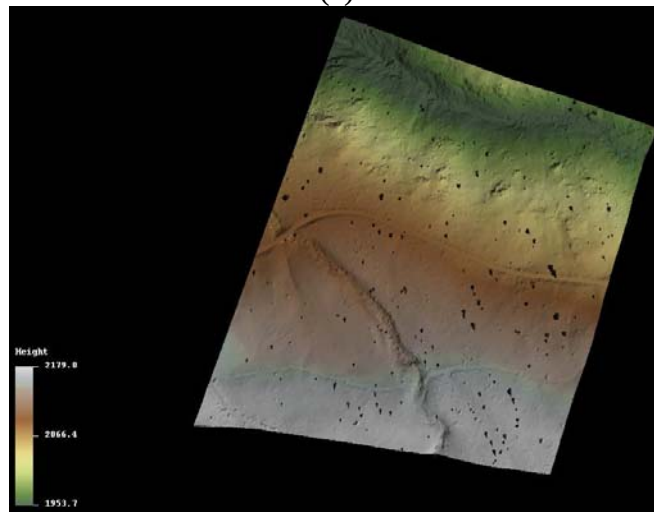


Figure 32. Selected (white box) LiDAR Point Cloud (QTC) of Standard and High Resolution data viewed in QTM

Using the same process stated above to generate the first two sets of DEMs, the reduced high resolution and the standard resolution DEMs were generated and are shown in Figure 34. The two DEMs were analyzed using ENVI and was found that there is a 0.999547 correlation between them, validating the artificial reduction of resolution of DEMs by random selection of points from a higher density data set using a programming code written in IDL.



(a)



(b)

Figure 33. High (a) and Standard (b) Resolution DEMs generated from the selected LiDAR Point Cloud data sets (Figure 33)

THIS PAGE INTENTIONALLY LEFT BLANK

V. CONCLUSION

Digital Elevation Models generated from lower density LiDAR data deviated from the Digital Elevation Model of high-density LiDAR data in a non-linear fashion. As the density of the original LiDAR data covering an area of 1625 meters by 875 meters (8.6 million points) was reduced to 90%, 66%, 50%, 30%, 10%, 5%, 3%, 1%, 0.5%, 0.3%, 0.1%, 0.05%, 0.03%, and 0.01%, the correlation results decreased more significantly at densities less than 0.0182 points per square meter (Figure 22 and Table 2). The same process applied to a larger LiDAR data covering an area of 3005 meters by 844 meters (33.5 million points) produced similar results (Figure 29 and Table 4); however, significant decrease in correlation did not occur until density reached 0.0132 points per square meter.

The first set of Digital Elevation Models having the same LiDAR data density also produced different correlation results (Figure 24). The lower the LiDAR data density the wider the distributions of correlation results become. Three percent (0.182 points per square meter) produced a mean correlation of 0.9535 with uncertainty of 0.56%, 0.3% (0.0182 points per square meter) has a mean correlation of 0.8928 with uncertainty of 2.64%, and 0.1% (0.006 points per square meter) has a mean correlation of 0.8486 with uncertainty of 88.4%. Thus, the Digital Elevation Models created from less than 0.0182 points per square meter LiDAR data density are inadequate.

Further analysis of DEMs indicated that the percentage of surface points contained in the Digital Elevation Models varied as LiDAR data density decreased. The LiDAR data covering an area of 1625 meters by 875 meters identified around 50% of total point cloud data as surface points and the LiDAR data covering an area of 3005 meters by 844 meter identified around 35% as surface points until the LiDAR data density was reduced to less than 0.6064 points per square meter and 0.6601 points per square meter, respectively (Figure 30). Therefore, a minimum of 0.6 points per square meter (or rounding off to 1 point per square meter) is necessary to generate an adequate Digital Elevation Model.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

- Anderson, B. C. (2008). *Assessing Accuracy in Varying LiDAR Data Point Densities in Digital Elevation Models*. Master's thesis, Naval Postgraduate School, Monterey, California.
- Burtch, R. (2002). LiDAR Principles and Application. Retrieved May 28, 2009, from Ferris State University Web site: http://www.ferris.edu/faculty/burtchr/papers/lidar_principles.pdf.
- Graham, L. (2009). Management of LiDAR Data. In J. Shan & C. Toth (Eds.), *Topographic Laser Ranging and Scanning Principles and Processing* (pp. 173-194). Florida: Taylor & Francis Group.
- Harding, D. (2009). Pulsed Laser Altimeter Ranging Techniques and Implications for Terrain Mapping. In J. Shan & C. Toth (Eds.), *Topographic Laser Ranging and Scanning Principles and Processing* (pp. 173-194). Florida: Taylor & Francis Group.
- JHU/APL. (2006). *ILAP Bare Earth Extraction Plug-In Version 1.0, Operator's Manual*.
- JHU/APL. (2007). *Quick Terrain Modeler Version 6 User's Manual*.
- Karatolios, A., & Krougios, P. (2008). *Extracting Hidden Trails and Roads Under Canopy Using LiDAR*. Master's thesis, Naval Postgraduate School, Monterey, California.
- Liu, X. (2008). Airborne LiDAR for DEM Generation: Some Critical Issues. [Electronic version]. *Progress in Physical Geography*, 32(1), 31-49.
- Liu, X., Zhang, Z., Peterson, J., & Chandra, S. (2007). The Effect of LiDAR Data Density on DEM Accuracy. [Electronic version]. In MODSIM07 International Congress on Modeling and Simulation (pp. 1363-1368). Christchurch, New Zealand.
- Petrie, G., & Toth, C. K. (2009a). Introduction to Laser Ranging, Profiling, and Scanning. In J. Shan & C. Toth (Eds.), *Topographic Laser Ranging and Scanning Principles and Processing* (pp. 1-27). Florida: Taylor & Francis Group.
- Petrie, G., & Toth, C. K. (2009b). Terrestrial Laser Scanners. In J. Shan & C. Toth (Eds.), *Topographic Laser Ranging and Scanning Principles and Processing* (pp. 87-128). Florida: Taylor & Francis Group.

Sithole, G. and Vosselman, G. 2003: Report: ISPRS comparison of filters. [Electronic version]. Department of Geodesy, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands.

Wehr, A. (2009). LiDAR Systems and Calibration. In J. Shan & C. Toth (Eds.), Topographic Laser Ranging and Scanning Principles and Processing (pp. 129 172). Florida: Taylor & Francis Group.

APPENDIX: RANDOM_PTS_FROMXYZ_V2.PRO

;This code uses an input in ASCII XYZ format and reduces the input file
;to the desired percentage of the input file. The reduced or output
;file is generated in ASCII XYZ file. The ASCII XYZ file has 3 lines
;of header and 4 columns (x, y, z, ;intensity).

```
*****  
;  
;UPDATE THIS SECTION  
*****  
;  
;Input file directory  
file_dir = 'C:\Documents and Settings\rlduldul\Desktop\  
;Input file  
file = 'name.xyz'  
;Output file directory  
output_dir = 'C:\Documents and Settings\rlduldul\Desktop\  
  
;Percentage to reduce file size to  
pct = [90.0,66.0,50.0,30.0,10.0,5.0,3.0,1.0,0.5,0.3,0.1,0.05,0.03,0.01]  
  
run_number = 'run1'  
*****  
;  
*****  
  
pos = strpos(file, '.')  
tf = strmid(file, 0, pos)  
outfile = output_dir + tf  
  
hdr = strarr(3)  
  
data = double([0, 0, 0, 0])  
temp = double([0, 0, 0])  
temp2 = 7  
  
print, 'input file: ', file_dir+file  
print, 'output file:', outfile  
  
;Open and Count the number of points  
openr, 1, file_dir+file  
readf, 1, hdr  
npts = 01  
hdr1 = ' '  
WHILE ~ EOF(1) DO BEGIN  
    readf, 1, hdr1  
    npts = npts+1  
ENDWHILE  
  
print, 'number of points:', npts  
close, 1  
  
;Open and read in the file  
openr, 1, file_dir+file  
readf, 1, hdr
```

```

data = dblarr(4,npts)
for index = 0L, npts-1 DO BEGIN
    readf, 1, temp, temp2
    data(0:2, index) = temp
    data(3, index) = temp2
    IF index mod 10000 eq 0 then print, 'number of points read:', index
ENDFOR
close, 1
;stop

data = data(*,1:npts-1)

print, 'Starting to extract random subsets'

FOR i=0L, n_elements(pct)-1 DO BEGIN
    nreduced = long(npts*pct(i)/100.)
    if (nreduced lt 2) then nreduced = 2
    print, 'original number of pts:', npts
    print, 'number of reduced pts:', nreduced

    ;Reduce dataset to percentage of original
    reduced_pts = RANDOMU(seed,nreduced)
    index = sort(reduced_pts)
    reduced_pts = reduced_pts(index) * npts
    reduced_pts = long(reduced_pts)
    print, 'Seed for Random number generator:', seed(0)

    data_2 = data(*,reduced_pts)

    ;Output reduced data set is ASCII format
    dir = output_dir

    pcti = fix(pct(i)*100)
    outfile_temp = outfile+'_'+string(pcti, "(I5.5)")+'_pct.xyz'
    print, 'output file:', outfile_temp

    openw, 2, outfile_temp
    printf, 2, 'Points taken from: ', file
    aa = string(pct(i), "(f8.2)")
    bb = string(npts, "(I8)")
    cc = string(nreduced, "(I8)")
    lable = 'Reduced to: '+aa+'%   Original # of pts: '+bb+ ' Reduced #
of pts: '+cc+ ' Seed: '+strtrim(seed(0))
    print, lable
    printf, 2, lable
    printf, 2, 'x, y, z, intensity'
    for j=0L, nreduced-1 do begin
        printf, 2, data_2(0:2,j), fix(data_2(3,j)), format = "( f19.12, 2x,
f20.12, 2x, f18.12, i4  )"
    endfor
    close, 2
ENDFOR
print, 'Finished'
END

```

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California
3. Professor Richard C. Olsen
Naval Postgraduate School
Monterey, California
4. Nancy Ann Budden
Naval Postgraduate School
Monterey, California
5. COL (Ret) Dave Trask
Naval Postgraduate School
Monterey, California
6. Ernie Reith
NRO/NIMA
Bethesda, Maryland