1-1-2015

Improving LiDAR Data Post-Processing Techniques for Archaeological Site Management and Analysis: A Case Study from Canaveral National Seashore Park

Christopher James Griesbach
University of South Florida, cgriesba@mail.usf.edu

Follow this and additional works at: http://scholarcommons.usf.edu/etd
Part of the Anthropology Commons, and the Remote Sensing Commons

Scholar Commons Citation

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Improving LiDAR Data Post-Processing Techniques for Archaeological Site Management and Analysis: A Case Study from Canaveral National Seashore Park

by

Christopher Griesbach

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Arts
Department of Anthropology
College of Arts and Sciences
University of South Florida

Co-Major Professor: Robert H. Tykot, Ph.D.
Co-Major Professor: Lori Collins, Ph.D.
Brent Weisman, Ph.D.

Date of Approval: March 3, 2015

Keywords: Remote Sensing, GIS, Geographic Information Systems, St. Johns

Copyright © 2015, Christopher Griesbach
# Table of Contents

List of Figures....................................................................................................................... ii

Abstract................................................................................................................................. iii

Chapter 1 : Introduction.................................................................................................... 1

Chapter 2 : LiDAR Background ....................................................................................... 7
  2.1 First and Last Returns .................................................................................... 8
  2.2 LiDAR Accuracy............................................................................................ 9
  2.3 Data Point Density........................................................................................ 11
  2.4 Point Clouds................................................................................................. 12
  2.5 Data Point Classification............................................................................... 13
  2.6 Interpolation.................................................................................................. 15
  2.7 Digital Elevation Models ............................................................................... 20

Chapter 3 : LiDAR Applications in Archaeology ............................................................ 22
  3.1 Hillshades and Vertical Exaggeration ........................................................... 22
  3.2 Viewshed Analysis........................................................................................ 25
  3.3 Strike Maps and Ethical Issues .................................................................... 27
  3.4 LiDAR Issues and Data Point Classification ................................................. 29
  3.5 Publicly Available LiDAR Data...................................................................... 34

Chapter 4 : CANA Cultural and Environmental Background ......................................... 41
  4.1 St. Johns II Culture ....................................................................................... 42
  4.2 Previous Archaeological Research ............................................................... 44
  4.3 Cultural Resource Laws ............................................................................... 47

Chapter 5 : Methods and Results .................................................................................. 52
  5.1 Data Acquisition ............................................................................................ 52
  5.2 Metadata ...................................................................................................... 58
  5.3 Data Processing ........................................................................................... 59

Chapter 6 : Discussion and Conclusion ........................................................................ 84

References Cited ................................................................................................................. 92
List of Figures

Figure 1 Map of Florida showing study area highlighted in red. .................................................. 3
Figure 2 NOAA Digital Coast data viewer. ................................................................................ 53
Figure 3 Google Earth screenshot highlighting CANA study sites in green............................. 55
Figure 4 Data Processing Model. .............................................................................................. 63
Figure 5 Comparison of Seminole Rest LiDAR data points....................................................... 66
Figure 6 Comparison of Castle Windy DEMs .......................................................................... 68
Figure 7 Comparison of Ross Hammock DEMs ....................................................................... 69
Figure 8 Comparison of Castle Windy DEMs and aerials.......................................................... 70
Figure 9 3-D DEM model of Castle Windy with data points overlaid. ...................................... 74
Figure 10 Comparison of cleaned and raw Castle Windy LiDAR data points......................... 77
Figure 11 Comparison of cleaned and raw Castle Windy DEMs............................................. 79
Figure 12 Data processing model for combining thinned data points with "ground" classified points. ......................................................................................................................... 81
Figure 13 Comparison of cleaned and raw Ross Hammock mound 1 DEMs......................... 83
Figure 14 Castle Windy slope models. ...................................................................................... 86
Figure 15 DEM of entire Ross Hammock site. .......................................................................... 89
Abstract

Methods used to process raw Light Detection and Ranging (LiDAR) data can sometimes obscure the digital signatures indicative of an archaeological site. This thesis explains the negative effects that certain LiDAR data processing procedures can have on the preservation of an archaeological site. This thesis also presents methods for effectively integrating LiDAR with other forms of mapping data in a Geographic Information Systems (GIS) environment in order to improve LiDAR archaeological signatures by examining several pre-Columbian Native American shell middens located in Canaveral National Seashore Park (CANA).
Chapter 1: Introduction

As technology marches forward, the lens through which we view both the history and prehistory of humanity comes into clearer focus. The relatively rapid pace of technological advancement witnessed during the 20th and into the 21st centuries has provided a wealth of new information and methods within a multitude of industries and fields of study. The field of archaeology is no exception, as technologies such as aerial photography, radiometric dating, isotope analyses, global positioning systems (GPS), and laser total stations have effectively revolutionized the study of ancient peoples and their lifeways. Another such technology is Light Detection and Ranging, or LiDAR as I will refer to it from here on.

LiDAR allows for the visualization of archaeological sites and their associated features in a manner that has never been witnessed before. When combined with an aerial platform, LiDAR can be used to quickly and efficiently produce three-dimensional (3D), digital representations of a given landscape at a level of detail and accuracy that is often not possible with more conventional, ground based methods of measurement and mapping. The advantages that LiDAR mapping offers both the archaeologist and those who manage archaeological sites are manifold, especially when one considers the increasing availability of free, publicly distributed LiDAR data.
It is still important, however, to consider the limitations of LiDAR and to be aware of issues that may arise during the processing and analyses of LiDAR data. This thesis will identify and address these issues by proposing effective, resource efficient methods for the processing and supplementation of LiDAR data that serve the purposes of archaeological site management and analysis. Within this context I specifically focus on prehistoric, south Florida archaeological sites. I also examine the utility of LiDAR data as they relate to archaeological site management regarding the documentation of these sites in a manner that complies with federal laws and standards such as the Native American Graves Protection and Repatriation Act (NAGPRA) and the National Register of Historic Places (NHPA).

The archaeological sites that will serve as the case studies for my investigation are all located within the Canaveral National Seashore Park (CANA) in New Smyrna, Florida. CANA is outlined in red in figure 1 below. I was first introduced to this topic when, from July to August of 2011 and 2012, I had the fortune to work as an Archeological Technician for the National Parks Service Southeastern Archeological Center (SEAC). During this time I assisted Dr. Margo Schwadron, Archeologist and NAGPRA Coordinator for SEAC, with fieldwork at CANA and labwork at the SEAC offices located in Tallahassee, Florida. The aims of this ongoing project are to assess the conditions of pre-Columbian shell middens and burial mounds located at CANA, conduct archaeological data recovery at the sites that face threats to their preservation, and develop a plan that either eliminates or mitigates these threats (Collins et al. 2013; Schwadron 2014).
Figure 1 Map of Florida showing study area highlighted in red
My work at SEAC mainly consisted of assisting with excavations at the CANA sites of Turtle Mound and Castle Windy in Volusia County Florida, as well as conducting laboratory processes such as floating, washing, sorting, and cataloguing of materials recovered from these sites. In conjunction with the National Park Service, the Alliance for Integrated Spatial Technologies (AIST), a research core at the University of South Florida, was also taking part in this project. In addition to my work with SEAC, I assisted AIST with addressing management considerations associated with the archaeological sites at CANA as well. These considerations include erosion processes, habitat restoration, and inaccurate site boundaries and terrain maps (Collins et al. 2013).

In an effort to produce more reliable maps that would give AIST and SEAC a better understanding of the site features and terrain at CANA, I aided AIST in collecting total station data and mapping grade GPS coordinates at the Seminole Rest, Turtle Mound, and Castle Windy sites. However, much of the terrain at CANA is difficult to access and heavy vegetation limits the use of total stations and GPS antennae. Therefore, developing solutions to address the preservation concerns of these CANA sites must include mapping methods that not only allow for a high degree of accuracy, precision, and site representation, but that are also cost feasible and able to document complicated terrain under heavy canopy cover.

Though the total station and GPS data that I was able to collect is in some places limited, these efforts were not in vain. Instead of discarding this data, it can be used to complement LiDAR data that, if processed correctly, is able to see through thick vegetation. Since the objectives of the project being conducted at CANA by SEAC and
AIST include correcting inaccurate site boundaries and terrain maps, LiDAR proves particularly useful in this instance.

Furthermore, high quality LiDAR data is publicly available for free through certain federal government websites, such as the National Oceanic and Atmospheric Administration’s Digital Coast website, which is the organization from which I derive my LiDAR data for this thesis. This allows park managers and archaeologists, who often have to contend with limited resources when considering the management and study of culturally significant sites, to take advantage of the wealth of information that LiDAR data has to offer. However, caution must be exercised when utilizing publicly available LiDAR data, since this data is not collected solely with archaeological interests in mind. Processing techniques employed by those who initially collect the LiDAR data may sometimes obscure the signatures of archaeological features, which can negatively impact the potential utility of LiDAR to archaeologists and park managers.

What other forms of mapping data can be integrated with LiDAR data and what processing methods can be performed to improve the investigation and management of archaeological features? How will these methods help state and federal park managers better document Florida’s archaeological sites in a manner that complies with federal heritage preservation standards? These are the questions that I am addressing in my thesis.

The next chapter of this thesis will provide background regarding the technology of LiDAR itself, followed by a chapter concerning the use of LiDAR within the field of archaeology. Chapter 4 covers the cultural and environmental background of my study
area, Canaveral National Seashore. Chapter 5 discusses where and how I acquired my data, what LiDAR processing methods I chose and why, how I integrated these methods with field data, and any issues I encountered while processing the data. The results of these methods are presented and discussed in chapter 6, followed a summary of my findings and my concluding remarks in chapter 7. Though the GPS and total station data that I collected with AIST at the Turtle Mound and Castle Windy archaeological sites at CANA appears limited when considered independently, these data shows promise when integrated with LiDAR data within a Geographic Information Systems (GIS) environment. This thesis looks at the best techniques for integrating LiDAR data with other forms of mapping and visualization information in order to improve site documentation and determine the best use of products derivable from LiDAR data, such as DEMs, slope, and hillshade maps, for the purposes of restoration, management and interpretation of Florida prehistoric archaeological sites. The results of my research reveal that ground-based, mapping-grade GPS coordinates and total station data can complement LiDAR data by providing multiple sources of additional information regarding the physical characteristics of shell middens.
Chapter 2: LiDAR Background

Before one can appreciate what the advantages and shortcomings of LiDAR are for the study of archaeology, one must first be familiar with the technology itself. Light detection and ranging, or “LiDAR”, describes a method whereby geo-referenced, three-dimensional data points are gathered through the combination of a laser rangefinder and a global positioning unit. This is classified as a remote sensing technique, and it can take either one of two forms, depending on the sort of platform that is used to house the laser scanning device. When a ground-based platform is used, the method is referred to as Terrestrial Laser Scanning (TLS), whereas when an airborne system is employed, it is referred to as Airborne Laser Scanning (ALS). For the purposes of this thesis, whenever I use the term LiDAR, I am referring to airborne systems.

Most LiDAR systems use lasers within the infrared (IR) spectrum and are classified as “active” sensors, meaning that they send out their own beam of light, unlike passive sensors, such as multi-spectral scanner systems, that rely on reflected sunlight. This means that LiDAR scans could theoretically take place at night, though this is not recommended since it is difficult for flight crews to see clouds at night that may negatively affect the quality of the survey (Jones 2010).
2.1 First and Last Returns

Whenever a pulse of infrared light emitted from the LiDAR sensor strikes any sort of object along its path towards the ground, the pulse returns an echo back to the sensor. This echo is recorded as a data point. However, a pulse does not always necessarily make it all the way to the ground, especially in cases where dense vegetation is present. The first echo that a pulse returns is known as the first return. If the pulse is able to pass through any sort of gaps in the first object that it strikes and continue on until it strikes an object that it cannot pass through, this final echo is known as the last return. This last return may represent the ground or it could also be a tree trunk, building, car, or simply areas of vegetation so dense that the pulse could not pass through.

Early-generation LiDAR sensors were only able to collect a small number of echoes from each pulse, usually just the first and last returns. Nowadays, LiDAR systems have developed to the point where the sensors can record what is known as the “full waveform”, which means that every echo that a pulse generates before producing the last return can be recorded.

Having access to this full waveform data allows whoever is processing the raw data points after the scan is complete to use different algorithms in order to separate the data in different ways depending on the needs of the user (e.g. Pluckhahn and Thompson 2012). Though it is often the first and last returns that are considered the most important, being able to see every echo along the path of the pulse is especially important for users who simply want those points associated with the ground level and not any other “false” last returns that represent dense vegetation, man-made structures, etc.
2.2 **LiDAR Accuracy**

Though LiDAR is today mainly utilized for acquiring detailed topographic data, the first application of LiDAR is seen in the 1960s in Airborne Laser Bathymetry (ALB) in an attempt to detect submarines (Jones 2010). At the same time that ALB was being developed, a similar concept was also extended to topographic LiDAR for the purposes of measuring surfaces on land. However, the core components of the LiDAR systems can only measure the position of an object or surface relative to the laser scanner based on the time taken for a single laser pulse to be emitted from a sensor array, strike a surface below, and be detected as a reflected signal. Therefore, these collected data points are simply floating out in space and do not have any sort of real-world coordinates attached to them.

While Global Positioning Systems, or GPS, could potentially alleviate this problem by acquiring real-world coordinates for the collected data points while a scan is taking place, a couple of issues initially presented themselves. A GPS platform capable of achieving a positional accuracy suitable for the purposes of LiDAR did not exist until the Navstar System became operational in 1994 (Jones 2010). This development coincided with the improvement of Inertial Measurement Units (IMUs) that are used to compensate for the motion of the aircraft while a LiDAR scan is taking place. The combination of these technologies introduced a degree of both relative and positional accuracy that make LiDAR a practical reality.

How accurate a LiDAR scan actually is, however, depends upon a number of factors. The most fundamental of these elements is the intended purpose behind LiDAR data acquisition. Even though LiDAR data are often publicly distributed by federal and
state agencies, these data were not originally collected with the sole purpose of public consumption. There are several issues with this sort of publicly available LiDAR data as they relate to archaeological utility, some of which I will highlight and attempt to address in later chapters. First and foremost, however, is that LiDAR projects commissioned by government agencies are generally focused on the large-scale mapping of environmental features. This can result in the use of a LiDAR system that generates data of a lower resolution than data produced for projects that focus on smaller areas or where greater precision is required, such as modern construction or archaeological sites (e.g. Pluckhahn and Thompson 2012).

Regardless of what sort of LiDAR scanner is used for a project, certain variables are inherent to every LiDAR system that can help control the intended accuracy of a scan. These variables are (Young 2011:5-6):

- The flying altitude of the LiDAR platform (lower altitude will concentrate the laser pulses, producing higher resolution but it will also limit the area covered by each flyover);
- The flight line spacing (overlapping flight lines means more data points are collected for a given area, though this will also limit the overall area covered by a scan);
- The repetition rate at which the laser is pulsing (more pulses will produce a denser collection of data points);
- The scan angle or field of view (the distance that the scanner moves from one side to the other. For instance, a scan angle of 30 degrees is tighter, results in
higher pulse densities, but also covers less area than a scan angle of 90 degrees);

- The nominal point spacing (the average point spacing that the LiDAR system is set to try to achieve).

2.3 Data Point Density

As one can gather, many of these variables relate to the density of data points collected for a given area. A higher density of points signifies more data relating to whatever features of the landscape the laser pulses struck, and additional data results in more accurate models. The variable that has the greatest influence regarding scan accuracy is the repetition rate at which the LiDAR system pulses. More laser pulses every second means more data points collected. A higher pulse volume also increases the probability that some of those pulses will weave their way through areas of dense vegetation and provide data points that represent the ground surface.

Newer LiDAR systems increase the laser pulse repetition rate. On average, however, when mapping large areas, as is the case for many government LiDAR projects, point spacing is usually between 1 and 2 meters (Young 2011:7). At this spacing, horizontal accuracy is about 0.5 m and vertical accuracy is around 15-20 cm (Young 2011:7). What features are recorded by the LiDAR scan, however, depend largely on the nature of the landscape in question.

Though LiDAR is capable of collecting millions of highly precise and accurate three-dimensional measurements in a short time, the key element of LiDAR is light, and
as such the system is limited to measuring places where light can reach. Therefore, in areas where dense vegetation canopy is present, LiDAR cannot penetrate “through” canopy coverage per se, though it can in some circumstances exploit gaps in the canopy that make it possible to record the ground surface. Nonetheless, thick vegetation can still partially or completely block light pulses, hence affecting the accuracy of a scan. This is especially true in areas of Florida where dense, low vegetation is present.

Conventional LiDAR scanners typically use a threshold of 1.5 m to discriminate between two consecutive return echoes, which is usually enough to differentiate mature trees from the ground surface (Pluckhahn and Thompson 2012:2). However, 1.5 m may not be enough to distinguish smaller trees and understory from the ground. This sort of low-lying vegetation is prevalent in many parts of southern Florida, CANA being no exception. Therefore, it is generally recommended that LiDAR surveys take place during the winter months in order to minimize vegetation interference.

2.4 Point Clouds

As stated earlier, the primary data collected by the LiDAR scanner are simply a series of points in space, yet these data points are only useful if they have been placed in a common coordinate system through the use of GPS. Once the data have been registered within this coordinate system, it is then necessary to align the grids of individual survey swaths to ensure that there are no discrepancies between scans that could lead to interference patterns (Jones 2010). If viewed as a text file, the point data appear as strings of numbers with different columns for X, Y, and Z values and each row represents
data from a single laser pulse. However, when imported into a GIS package and viewed as a three-dimensional model, the data at this stage produce what is referred to as a “point cloud”. A point cloud can be defined as “a collection of XYZ coordinates in a common coordinate system that portrays to the viewer an understanding of the spatial distribution of a subject” (Jones 2010:5).

The easiest way I have come across to visualize what this point cloud represents is to liken the data points to snowflakes settling on every surface that they contact. After a snowfall, some of the flakes will be scattered over trees, bushes, cars, etc., and some will also reach the ground. If you then remove everything on which the ‘snow’ has settled, you are left with a cloud of flakes floating in three-dimensional space (Jones 2010:9).

2.5 Data Point Classification

The next step in data processing is to use algorithms to automatically classify and filter data points. Once the data points are divided into distinct classes, the user can then remove any classes of points from the point cloud that are not wanted in the final product. Which points are removed based on the parameters established by the selected algorithm depends upon what type of surface the LiDAR data is meant to represent. For instance, if the user of the data wants a digital representation of the natural topography of the landscape, say for construction or drainage purposes, then any points that the filtering algorithm deems as related to vegetation and man-made structures will be eliminated. On the other hand, if the user simply wants to see the degree of vegetation canopy coverage for some sort of environmental purposes, then those points that correspond to vegetation
will be preserved. What sort of filtering algorithms are used in this process has a tremendous effect on the sort of data that are either retained or eliminated.

Filtering algorithms can be categorized into four general groups based on the structure of bare-earth points in a local neighborhood (Sithole and Vosselman 2004). The first of these are slope-based algorithms, which measure the difference in slope (or height) between two points and simply assume that the highest point belongs to an off-ground object if the slope is higher than a given threshold. The next set of classification methods are referred to as Block-minimum algorithms, whereby bare-Earth points reside on a horizontal plane, with a corresponding buffer zone. Any data points that fall outside of this zone are considered off-ground.

Clustering/segmentation filtering techniques assume that clusters of data points belong to off ground objects if they are higher than their surrounding neighborhood. Lastly, surface-based classification algorithms use a parametric surface with a corresponding buffer that defines a region in three-dimensional space where ground points are expected to reside. Though similar in concept to the Block-minimum algorithms, surface based algorithms are more flexible and appear to provide better results in classifying ground surface points, as they take into account the natural curvature of the Earth and any undulating topography (Sithole and Vosselman 2004).

Much of the time, these data point classification and filtering steps are completed by the company that collects the LiDAR data, though it is sometimes carried out by the customer that ordered the data in the first place. What effects these algorithms can have on those data points related to archaeological features and whether or not the
archaeologist has access to the raw, unfiltered data points are topics that shall be addressed in the next chapter.

2.6 Interpolation

Regardless of what algorithms are chosen to eliminate certain data points and who makes that decision, the resulting, “filtered” point cloud still does not present the data points in a format that is very informative to the user. While the points may be dense enough in some areas to help define certain features of the landscape, another processing step must be taken to fill in the gaps between the data points and create a three-dimensional surface. This step is referred to as “interpolating”.

Interpolation uses the information provided by the collected sample points to predict the unknown values for any geographic data points, such as rainfall, noise levels, chemical concentration, etc., that were not recorded during data collection. In the case of LiDAR data points, the relevant information is elevation, or the z-value as it is also known. This interpolation process produces a raster surface made up of cells. This raster surface can then be used to create a three-dimensional model of the landscape topography.

How the unknown elevation information is derived from the collected data points depends on the method of interpolation employed. Just like the previously-mentioned algorithms that are used to process the raw point data, there are also different methods of interpolation. Each of these methods interpolate the data in various manners, and each possess their own strengths and weaknesses depending on the sort of data present and what the researcher is hoping to extract from it. I will provide a brief overview of two of
these interpolation techniques, as well as the circumstances that are most appropriate for their respective uses. Each of these interpolation methods are available as tools within the ArcMap software by Environmental Systems Research Institute, which I will refer to as ESRI from here on.

The first of these interpolation methods is known as Inverse Distance Weighting (IDW). The IDW interpolation method determines unknown point values in a relatively straightforward manner, using a “linearly weighted combination of a set of sample points…the weight is a function of inverse distance” (ESRI website, accessed August 14, 2013). Basically, an unknown point is assigned a value based on the average of values of surrounding sample points.

How strong of an influence the respective values of these sample points have on determining the average value of the unknown point, however, depends on the distance of the sample point from the unknown point. The smaller the distance of the sample point from the unknown point, the greater its influence in determining the unknown point’s value. How many surrounding sample points are used in the determination of the unknown point, how much influence nearer sample points have as opposed to farther, and how far from the unknown point these sample points are allowed to be are all parameters that can be set by the user within the IDW tool in the ArcMap program that I use for this thesis.

In addition to being rather straightforward conceptually, IDW is also relatively less intensive in terms of processing power than other interpolation methods. This makes it well suited for larger datasets. The best results are obtained from IDW interpolation when
collected data sampling is adequately dense with regard to the local variation that the user is attempting to simulate. Otherwise, in areas where data coverage is sparse or uneven, as may be the case in certain areas of dense vegetation where LiDAR is concerned, the resulting raster surface may not sufficiently represent the desired output (Philip and Watson 1985).

Another issue to keep in mind when using IDW interpolation is the fact that the output value for an unknown point is based on the average of surrounding sample point values. Therefore, being an average, this value cannot be greater than the highest or less than the lowest input sample values. This means that any extreme discrepancies in values that are not represented by sample points will not show up in the final raster surface. When applied to topography, IDW interpolation will not retain in the output any drastic changes in elevation, such as cliffs, ridges, valleys, etc., if these features are not adequately represented in the collected data points (Philip and Watson 1985). This effect can be negated to a degree if one uses barrier features in order to specify the location of linear features known to interrupt the surface continuity during IDW interpolation, though using barriers will significantly extend the processing time (ESRI website, accessed August 14, 2013).

IDW is referred to as a deterministic interpolation method because it is based directly on the surrounding measured values. A second set of interpolation techniques consist of geostatistical methods, which utilize autocorrelation in order to examine the statistical relationships among the measure points in order to predict unknown points and produce a surface (ESRI). Because of this, geostatistical methods also provide some quantifiable measure of certainty or accuracy of these predictions, unlike the deterministic
methods. One such geostatistical method that I will highlight is commonly known as “Kriging”, named after Daniel G. Krige, a mining engineer who pioneered the field of geostatistics.

Like IDW, the Kriging interpolation process is available as a tool within the Spatial Analyst toolbox within the ArcGIS software package. The ArcGIS Help website explains that Kriging “assumes that the distance or direction between sample points reflects a spatial correlation that can be used to explain variation in the surface” (ESRI website, accessed August 14, 2013). In order to discover this correlation and derive predictions, the Kriging method goes through a two-step process.

The first step includes creating *semivariograms* and covariance functions in order to estimate the statistical dependence values that are based on the model of spatial autocorrelation. This involves the creation of a graph of the empirical semivariogram, which is computed through the following equation for all pairs of sample points separated by distance $h$ (ESRI website, accessed August 14, 2013):

$$\text{Semivariogram}(\text{distance}_h) = 0.5 \times \text{average}((\text{value}_i - \text{value}_j)^2)$$

The above formula involves calculating the difference squared between the values of the paired sample data points.

The last step is to fit a model to the points forming the empirical semivariogram so that unsampled points may then be predicted. This model is based on one of five functions defined by the user. These functions are: circular, spherical, exponential, Gaussian, and linear. It is important for the user to try different models and select the correct one, since the selected model influences the prediction, especially when the shape of the curve near
the origin of the empirical semivariogram differs significantly. The steeper the curves of
the model near the origin, the more influence the closest neighbors to the predicted point
will have on the prediction, resulting in a less smooth output surface (ESRI website,
accessed August 14, 2013). Each model listed above is designed to fit different types of
phenomena more accurately.

The general formula used by both the Kriging and IDW interpolation methods is
formed as a weighted sum of the data. This formula is shown in the equation below,
where:

- \( Z(s_i) \) = the measured values at the \( i \)th location
- \( \lambda_i \) = an unknown weight for the measured value at the \( i \)th location
- \( s_0 \) = the prediction location
- \( N \) = the number of measured values

\[
\hat{Z}(s_0) = \sum_{i=1}^{N} \lambda_i Z(s_i)
\]

In IDW, the weight, \( \lambda_i \), depends solely on the distance to the prediction location. On the
other hand, with the Kriging method, the weights are based not only on the distance
between the measured points and the prediction location, but also on the overall spatial
arrangement of the measured points.

Because of the two step process that the Kriging interpolation requires, it is said
that kriging uses the data twice (ESRI website, accessed August 14, 2013). This
dramatically increases the processing time and can lead to issues when dealing with large
LiDAR datasets. The user also has to be familiar with the study area in question and know
if there is a spatially correlated distance or directional bias in the LiDAR data. My personal experience with processing LiDAR data revolves mostly around the IDW method mentioned earlier, though I examine both interpolation techniques when I process LiDAR data for CANA.

2.7 Digital Elevation Models

The three-dimensional model of the terrain that is produced from the interpolation process can depict different features of the terrain depending on the sorts of data-points that are either retained or discarded during the raw data processing phase. If just the first return data-points are included in the interpolation, then the resulting 3D model will depict whatever features of the landscape that the LiDAR pulses first struck during data collection. Therefore, these first returns often represent vegetation or man-made structures in addition to the ground surface. The 3-dimensional model created from interpolating these first returns is sometimes referred to as a Digital Surface Model, or DSM. DSMs are often used for environmental or urban planning purposes, as they provide detailed and measurable information regarding vegetation coverage and/or building distribution (Jones 2010).

On the other hand, if the last returns are retained during raw data processing instead of the first returns, the resulting 3-dimensional model will represent the “bare-earth” surface. This model can be referred to as either a Digital Terrain Model (DTM) or Digital Elevation Model (DEM). From here on, however, I only use the term DEM to refer to this bare-earth 3D model. DEMs provide a relatively accurate visualization of a given
landscapes’ topography without any obscuring above-ground features. Because of this reason, DEMs are perhaps the most useful of the 3D surface models extractable from LiDAR data with regards to the study of archaeology.

The possibilities for manipulating LiDAR data, however, do not end with the DEM. Vertical exaggeration, hillshades, strike maps, and slope models are just a few examples of the sort of information that can be gathered from LiDAR data, all of which is of great use to the archaeologist. What these LiDAR derivable products are and how they can be such a valuable source of archaeological information is a subject that I address in the next chapter.
Chapter 3: LiDAR Applications in Archaeology

LiDAR data can provide a wealth of information to archaeologists with respect to the prospection, analysis, and management of archaeological sites. However, the DEMs mentioned in the previous chapter are just the tip of the iceberg with regards to the types of information and visualizations that can be extracted from LiDAR data. In this chapter I highlight the value of DEMs and other sorts of 3-dimensional models derivable from LiDAR data to archaeology through the examination of several case studies where LiDAR has proved to be an invaluable tool for the purposes of archaeological inquiry. At the same time, I also draw attention to issues with LiDAR data processing methods as they relate to the study of archaeology and what these problems can mean in terms of the preservation of archaeological sites and their associated information.

3.1 Hillshades and Vertical Exaggeration

One of the first governmental institutions in the world to embrace the possibilities of LiDAR information was the Environmental Agency of the United Kingdom, which in 1999 commissioned a LiDAR survey for the purposes of monitoring river corridors and coastal areas of England and Wales (Brown 2008). Since that time, these datasets have been continuously updated. Though collected ostensibly for environmental reasons, these Environmental Agency LiDAR datasets have been used for archaeological
purposes since 2002, when archaeologists Holden et al. (2002) utilized DEMs derived from this data to identify and record the subtle earthwork traces of a Roman period fort. Though this site is located on active farmland, thereby eliminating the need to remove any obscuring canopy vegetation signatures from the collected data points, the fact that this land has been ploughed for decades means that the earthworks which characterize the fort have been reduced to less than 1 m in height (Holden et al. 2002). Aerial photography surveys had previously missed this site. However, by using a DEM of the area combined with different visualization techniques that included hill shading and vertical exaggerations, Holden et al. (2002) were some of the first researchers to demonstrate the utility of LiDAR data for detecting previously unknown archaeological sites that would normally be missed through traditional remote sensing methods.

The hillshading and vertical exaggeration techniques that Holden et al. use beg further explanation, as they are methods that I employ with regard to the CANA LiDAR datasets. Hill shading is a DEM processing method that simulates a traditional aerial photography technique whereby photos of a given landscape are taken at a time of day when the sun is at a low angle relative to the ground surface. Therefore, due to the low angle of the light source, any changes in topography will throw shadows that are visible in the photos. These changes in elevation are sometimes indicative of a buried archaeological feature and may be noted for further investigation.

This aerial photography technique, however, is severely limited by a number of factors, the first of which is that it requires a bare landscape devoid of any vegetation that would obscure or altogether block the shadows made by buried archaeological features. Therefore, this method is usually reserved for agricultural areas that have recently been
harvested and above ground vegetation is at a minimum. Even in such a best case scenario, however, the static nature of the light source, i.e. the sun, used in low-angle aerial photography raises another issue. Linear features, such as trenches, canals, or sometimes mounds that lie perpendicular to the light source will often not show up in the low angle aerial photos. Such occurrences can negatively affect the study of known archaeological sites as well as the uncovering of unknown sites in cases where this data is used for prospecting purposes.

Hillshade models produced from LiDAR data, on the other hand, have the advantage of a movable light source that can illuminate the three-dimensional model of a landscape from virtually any conceivable angle. This gives the researcher the ability to view an area of interest from multiple lighting directions, thereby eliminating the issue of linear archaeological features seemingly disappearing when lined up perpendicular to a low-angle light source. In addition, the fact that the DEM from which a hillshade model is produced is a “bare-earth” representation of a landscape means that hillshade modeling, unlike aerial photography, is not restricted to areas with limited vegetation.

The other technique employed by Holden et al. (2002) to reveal subtle Roman period earthworks is known as “vertical exaggeration”. Vertical exaggeration takes advantage of one of three values associated with every data point featured within any given point cloud. As mentioned prior, the data points within a point cloud represent a laser pulse return. The three values recorded in this return are the georeferenced X, Y, and Z coordinates of the respective point. The Z value pertains to the elevation of the point and is what makes the point cloud and subsequent DEM three-dimensional.
Though the vertical accuracy of LiDAR is relatively high, the topographic traces of some archaeological features are so subtle that their elevation signatures require some degree of amplification in order to facilitate the discovery and/or study of these features through the use of DEMs, hillshades, or any other models that may be generated from the point cloud data. This outcome is achieved through a process known as vertical exaggeration, whereby the Z values for all of the points within a point cloud are multiplied by a factor specified by the user. This has the effect of “exaggerating” so to speak the topography of the area of interest. Ergo, what may have been barely distinguishable differences in elevation prior to the vertical exaggeration, any slight rises or depressions should stand in a relatively starker contrast to the surrounding topography after this operation is completed. How slight the initial differences in elevation are determines how large of a factor the user will need to multiply the Z values by.

3.2 Viewshed Analysis

Due in part to the success of the study conducted by Holden et al. (2002), the English Heritage department in the UK commissioned a LiDAR survey in 2001 specifically for the archaeological investigation of the Stonehenge World Heritage Site. The results of this survey are reported by Bewley et al. (2005). The objective of this study was to determine how many known archaeological sites around the immediate vicinity of Stonehenge could be detected using LiDAR survey. Though much of the landscape surrounding Stonehenge is relatively clear, this lack of vegetation is due to the fact that the areas bordering Stonehenge have been used as farmland for hundreds of years. Consequently, centuries of ploughing and other agricultural activities have significantly
reduced the topographic signatures of the archaeological features in the vicinity of Stonehenge (Bewley et al. 2005).

As we have already witnessed, LiDAR data and the digital models derivable from it are particularly suited to this sort investigation, and the results of Bewley et al.’s study (2005) demonstrate this. Through a combination of DEMs and hillshade modeling, Bewley et al. were able to not only contribute new information regarding known archaeological sites around Stonehenge, but also uncover previously unknown Neolithic sites in this area.

While hillshade models can be visually interpreted in order to identify archaeological features, Bewley et al. also use LiDAR data to explore the spatial relationships between Stonehenge and other Neolithic sites in the area through a process known as *viewshed* analysis. Viewshed analysis allows the researcher to examine what sightlines are present between certain areas. In order to identify these sightlines, however, researchers have to determine how elements of the surrounding landscape influence the visual perspectives from and between archaeologically significant features within said landscape. LiDAR data, and DEMs specifically, can greatly assist in this endeavor, as they represent a three-dimensional, digital model of a given area’s topography, whereby one can project sightlines from any given point on the landscape and identify what these sightlines intersect.

In the case of Bewley et al. (2005), viewshed analysis is utilized in to order to study the “inter-visibility” of the known and previously unknown Neolithic sites with respect to each other and Stonehenge proper. Therefore, LiDAR allows the archaeologist to not only examine archaeological features within the overall context of the site; LiDAR also
provides the means by which a researcher can investigate an archaeological site within the context of the surrounding landscape.

3.3 Strike Maps and Ethical Issues

Another example of how utilizing LiDAR data can help shed new light upon old archaeological research questions and even change long-held perspectives regarding ancient peoples and the way they lived comes from an article entitled “Airborne LiDAR archaeology, and the ancient Maya landscape at Caracol, Belize” (Chase et al. 2011:387). In this article, the authors use aerial LiDAR data to map a relatively large area in and around the Classic Period (A.D. 250-900) Maya population center of Caracol, located on the Vaca Plateau in western Belize.

Researchers have attempted to map the site of Caracol for nearly 60 years (Chase et al. 2011). Unfortunately, however, the dense tropical vegetation present in this area makes ground-based mapping methods very difficult and labor-intensive, and hence quite expensive as well. The thick jungle canopy also renders traditional aerial mapping techniques, such as aerial photography, fruitless. Due to these issues, researchers have concentrated primarily on the monumental public architecture that is present within most site epicenters (Chase et al. 2011).

Only limited portions of the Mayan settlements and agricultural terracing surrounding epicenters such as Caracol have been mapped, resulting in largely incomplete spatial layouts of such sites. As a consequence, “…how the ancient Maya distributed and organized themselves over the landscape and how they supported large
populations continue to be debated” (Chase et al. 2011:387). Though agricultural terracing has been documented for Caracol (Healy et al. 1983; Chase and Chase 1998), the full extent of the modified landscape in this area has been difficult to demonstrate.

After only several days of airborne LiDAR flyovers and three weeks of post-field data processing, however, Chase et al. (2011:388) were able to produce results that “far surpassed over two and half decades of on-the-ground mapping by revealing images of a massive, modified landscape that ties settlement, roadways, and agricultural terraces together into a complete settlement system”. The LiDAR-derived imagery generated by this research reveals both previously mapped and hitherto undiscovered archaeological features. The clear, unobstructed perspective of Caracol and the surrounding landscape, with clear evidence of extensive terracing, that the LiDAR derived DEM affords is both striking and extremely informative for the purposes of archaeology. Whereas before only 3.5 sq km of terracing and 23 sq km of settlement around Caracol were previously recorded (Chase and Chase 1998; 2001), the LiDAR data analyses performed by Chase et al. (2011) have made it possible to identify archaeological features throughout the entire 200 sq km area of the Vaca Plateau.

While the new information concerning Caracol and ancient Mayan settlement patterns provided by LiDAR analyses is archaeologically significant, Chase et al. (2011:393) also make it clear that this data should not just stand on its own. More traditional archaeological methods, such as ground-mapping and excavation, can add details, functional information, and dating to what is shown through remote sensing, especially in areas where LiDAR data point coverage is thin or nonexistent due to factors such as excessively heavy foliage, cloud cover, or modern development.
One tool that is useful for identifying such areas is known as a strike map. Strike maps are a form of metadata that reveal the degree of LiDAR data point coverage over any given region within a study area. By identifying areas of interest within the LiDAR data that feature minimal data point coverage, researchers can pinpoint certain sectors for additional data collection that can be used to supplement the LiDAR data. This helps limit the resources expended on ground-based investigative methods.

Other concerns that Chase et al. (2011) raise are the ethical issues associated with LiDAR data and its associated imagery. Possession and distribution of such highly detailed representations of archaeological sites must be handled with the upmost care, as it is possible that the open sharing of such information could lead looters directly to new targets of opportunity (Chase et al. 2011:397). On the other hand, the LiDAR data analyses conducted for Caracol by Chase et al. (2011:396) do demonstrate “…the large-scale integration of a Maya metropolis; the high density of its dispersed ancient inhabitants, and the intensive development of the Maya landscape for a sustainable base.” Therefore, the benefits of integrating LiDAR data into archaeological investigations far outweigh any negatives, as this inquiry into ancient Mayan social organization demonstrates.

### 3.4 LiDAR Issues and Data Point Classification

Though it is clear that LiDAR can contribute a wealth of information regarding archaeological inquiry and site management, one must also be aware of factors that can limit the archaeological utility of this data. The next case study that I will present raises
these sorts of issues, though it also presents a set of methods designed to address these issues. In an article entitled “Flights into the past: full-waveform airborne laser scanning data for archaeological investigation”, authors R. Lasaponara, R. Coluzzi, and N. Masini (2011) present a LiDAR data processing chain and threshold-based algorithm that they devised in an effort to detect archaeological features at two historic sites in Southern Italy.

The first of these sites is the medieval village of Monte Serico, which is located on a hill with an elevation of around 590 m that faces over a hilly landscape that is crossed by a river (Lasaponara et al. 2011). This village was abandoned by the first half of the 15th century, and the only buildings that remain on the surface of the hill today are a castle and a church (Lasaponara et al. 2011). The second area that the authors examine is another medieval village known as Monte Irsi. Monte Irsi is surrounded by a forest, which is also investigated by the authors, and there are no standing medieval structures located at this location. Aside from their medieval origin, these two sites also share similarities in terms of their surrounding landscapes, which are characterized by close canopy, dense under-storey, low vegetation on steep slopes, and terrain with abrupt changes (Lasaponara et al. 2011).

It is these very same features of the landscape that, according to the authors, present issues regarding LiDAR analyses for the study areas (Lasaponara et al. 2011). As has been mentioned in previous sections, thick vegetation, especially low-lying vegetation, can make it difficult for LiDAR pulses to penetrate foliage and strike the ground surface. In addition, the authors also state that geomorphologic discontinuities, such as the steep slopes and sharp ridges present at both study sites, are also factors of
complexity that have to be considered when processing the LiDAR data for an area (Lasaponara et al. 2011).

Of particular concern for the study of archaeology is the ability of LiDAR processing techniques to distinguish ground from non-ground data points in LiDAR datasets. As was shown in the previous case study (Chase et al. 2011), the reliability of DEMs and any products derivable from them is directly related to the successful identification and removal of non-ground data points, while at the same time preserving those points that represent archaeological features. Therefore, in this article, the authors present a LiDAR data processing chain and a threshold-based algorithm devised specifically for the purposes of archaeology. The effectiveness of these methods is assessed through their application to the two test sites mentioned above.

The LiDAR surveys for both sites were carried out over September and October of 2008 using a full-waveform scanner mounted on a helicopter. A relatively high resolution (20 points/m2) and accuracy (25 cm horizontally and 10 cm vertically) were achieved for the data points (Lasaponara et al. 2011). The algorithm that the authors use to classify ground points is derived from what is known as the Axelsson TIN model. This algorithm is named after its creator, Peter Axelsson (2000), and is based on a progressive Triangulation Irregular Network (TIN) densification.

Basically, this method starts with a coarse surface composed of triangles whose vertices are obtained from reference data points that meet neighborhood minima among (Axelsson 2000). The triangles multiply and progressively get smaller as new points are added in an iterative manner so long as they meet criteria based on distances to TIN
facets and angles to the vertices of the triangle (Axelsson 2000). This has the effect of gradually refining the surface and adding higher detail to landscape features.

The Axelsson algorithm has been independently tested and shown to perform relatively well, with low total errors (Sithole and Vosselman 2004), and it is included in Terrasolid’s Terrascan commercial software for ground classification. The key input parameters that must be assigned by the user, such as maximum building size, terrain angle, maximum edge length, etc., all pertain to the proper identification of threshold-values that can cope with different surface characteristics and cover types (Lasaponara et al. 2011). This requires some degree of background knowledge regarding the landscape in question and may even require fieldwork in order to ascertain these threshold values.

However, the authors are also looking to classify all of the data points within their collected dataset, not just ground surface points. To this end, the Axelsson algorithm is applied several times to the data using a strategy based on a set of “filtrations of the filtrate” (Lasaponara et al. 2011:4). Using this method, “appropriate criteria for the classification and filtering were set to gradually refine the intermediate results” while the workflow for these refinements can be summarized as follows (Lasaponara et al. 2011:4):

i. Low point Classification

ii. Isolated points classification

iii. Air points

iv. Ground Classification

v. Classification of points below surface
vi. Classification of points by class

vii. Classification of points by height from ground for different heights

The authors determine that LiDAR derived DEMs are powerful instruments for detecting archaeologically related subtle elevation changes and provide a “detailed spatial characterization of urban fabric of the two medieval villages, even in the presence of significant vegetation cover, and with intense erosion process” (Lasaponara 2011:9). Indeed, the ability to identify areas either damaged by or susceptible to certain geomorphological processes, such as erosion or flooding, is an added bonus in this instance. Not only is LiDAR useful for finding and studying archaeological sites, it can also aid in the understanding of geomorphological processes in the area of concern. Knowledge of how such processes interact with a landscape can help archaeologists and government land managers determine any threats to the preservation of archaeological remains that said landscape contains.

Though the research conducted by Lasaponara et al. (2011) provides a valuable, as well as adaptable, processing model for LiDAR data, there are a few elements of their methodology that may make it too costly to undertake for many archaeological research and/or management projects. Terrascan is specialized, proprietary software designed solely for the processing of LiDAR data. According to the Terrasolid website, a full version of Terrascan alone costs approximately $6,600 (Terrasolid Ltd. 2012). This is a steep price for a program that pertains only to LiDAR data, and could prove prohibitive to individuals or agencies with limited resources hoping to use LiDAR data for archaeological research or site management.

Yet, the high cost associated with collecting and processing LiDAR data is usually offset by the value of the information that it provides. As we have seen in the case studies
exhibited thus far, mapping an area of interest with LiDAR is often much less time and resource intensive than alternative, ground base methods, such as total station mapping. In addition, LiDAR provides data in such a format that unique and varied analyses can be performed, such as viewshed analyses, hillshades, and DEMs.

Nevertheless, due to the recent economic climate, budgets for archaeological research and government land management are frequently being cut, forcing archaeologists and land managers to spread financial resources more and more thinly. However, this does not mean that LiDAR data is off-limits to such interested parties. Because LiDAR data is of value to many different industries, it is often collected for purposes other than archaeological inquiry. If the organization collecting the LiDAR data happens to be a state or federal agency, this data is usually made available to the public for free. Publicly available LiDAR data can alleviate many of the issues associated with collection costs mentioned above, though use of these data does not come without cautions.

3.5 Publicly Available LiDAR Data

LiDAR analyses provide a wealth of information to numerous industries aside from archaeology, and it is this versatile value that makes LiDAR data an important resource for archaeologists to exploit. For example, the detailed and highly accurate elevation information that LiDAR supplies can help determine flood zones in a given area, which in turn can aid insurance agencies in evaluating flood risk or may assist government agencies in establishing evacuation routes and emergency response protocols.
One must also be aware of the problems that the sharing of LiDAR data between very different fields of study presents. Collecting airborne LiDAR data can be an expensive proposition, as one must consider the costs not only of the LiDAR device itself, but also the price of the aerial platform that will bear the laser scanner and any specialized personnel and computer programs required to collect and process the data, such as pilots to fly the aerial platform and algorithms to process the raw data. Therefore, due to the high costs associated with acquiring LiDAR data, its collection is often restricted to industries and government agencies with sufficient financial resources that can cover the prodigious price tag.

The case studies that I have mentioned thus far have utilized LiDAR data that were acquired specifically for the archaeological project at hand. However, despite the wealth of information that LiDAR can provide archaeological investigations, paying for the collection of these data can be prohibitively expensive. Fortunately, free, publicly available LiDAR data is becoming more widespread. The only costs associated with this LiDAR data are related to the programs used by the researcher to process and visualize the data. Though the limited project budgets that many archaeologists have to work with may restrict them to acquiring publicly available LiDAR data, being a secondary consumer of LiDAR data can have both positive and negative effects in terms of the archaeological utility of LiDAR technology. The last case study that I present below highlights these advantages and disadvantages, as well as illustrating the use of LiDAR within the field of Florida archaeology.

While the use of LiDAR within archaeology in general is becoming more and more common, finding an example of its application within Florida archaeology specifically was,
admittedly, difficult. One of the few published investigations I was able to find in this regard is the relatively recent study conducted by archaeologists Thomas J. Pluckhahn and Victor D. Thompson (2012), which is detailed in their article entitled “Integrating LiDAR data and conventional mapping of the Fort Center Site in south-central Florida: A comparative approach”. This article proved invaluable for my own research, as the authors not only face much of the same issues with LiDAR data processing that I am attempting to tackle, but they also address these issues in a manner similar to my own, and all within the context of Florida archaeology. Therefore, I will both summarize the findings of this study here and use the methods employed by Pluckhahn and Thompson as a point of comparison for my own methods and results throughout the remainder of my thesis.

The Fort Center site, the study area for the Pluckhahn and Thompson investigation, is located in Glades County, Florida, west of Lake Okeechobee. The archaeological features within the site consist of circular ditches, circular and oval mounds less than 1 m in height, and a mortuary complex composed of one large burial mound and one smaller burial platform (Pluckhahn and Thompson 2012). These features are spread over a relatively wide area extending a minimum of 1.5 km east-west and 1 km north-south.

Prior to the work conducted by Pluckhahn and Thompson, the most intensive investigations of the Fort Center site were those by William H. Sears, who began work at the site in 1967 and completed seven field seasons there (Sears 1982, in Pluckhahn and Thompson 2012). Despite such extensive excavation at Fort Center, due to the sprawling nature of the site, Sears only mapped a portion of the archaeological features.
topographically with a transit. The rest of the features appear as line drawings. In addition, misalignment of the grid and altering of the topography, whether due to backdirt piles or leveling of some features, during Sears’ excavations have resulted in both incomplete and misinformed mapping of the Fort Center site.

The overall project conducted by Pluckhahn and Thompson at Fort Center was designed to test the hypothesis that maize was grown at the site during Woodland or earlier periods (2012). However, even though Pluckhahn and Thompson (2012:5) refer to their mapping of the site as ancillary to this project, they also state that “any archaeologist who has read Sears’ report [on Fort Center] will recognize the need for a new map”. Ergo, Pluckhahn and Thompson use their mapping of the Fort Center site as a case study for a method they propose that integrates LiDAR and total station data. Through this article, the authors also intend to present “a practical guide for archaeologists who wish to use [LiDAR] data for site mapping, but whose knowledge of the topic is limited” (Pluckhahn and Thompson 2012:2).

Using LiDAR data not only to complement, but also to correct site maps produced through traditional methods is another valuable application of LiDAR within archaeology. However, contrary to some of the other studies that I have highlighted in this chapter, the LiDAR data used by Pluckhahn and Thompson in this investigation was not acquired specifically for archaeological purposes. Instead, Pluckhahn and Thompson (2012:6) employ LiDAR data that was originally collected by a private firm in 2007 at the request of several federal and state agencies for the purposes of hydraulic modeling of the Herbert Hoover Dike that surrounds Lake Okeechobee.
As has been discussed previously, using publicly available LiDAR data for archaeological purposes is much cheaper than acquiring specifically for the project. On the other hand, utilizing public LiDAR data does remove any control over how this data is collected. In the case of the Fort Center study, Pluckhahn and Thompson (2012:6) mention that the metadata for their LiDAR states that the data was retrieved on flights at lower elevations than had originally been planned due to cloud cover. However, lower flight elevations mean that the LiDAR scan covers less area, which in turn results in more flyovers needed to cover the intended area and higher costs. Concordantly, the scan angle was widened from 30 degrees to 90 degrees in order to cover the same amount of area and keep the price of data collection down (Pluckhahn and Thompson 2012). Pluckhahn and Thompson (2012:6) mention that this wider angle may have affected the quality of the LiDAR data for their area of interest. In addition, public LiDAR data are usually processed with some form of proprietary filtering algorithms before being distributed. As a result, Pluckhahn and Thompson (2012:6) state that they are “unable to describe the classification of the data in detail, nor to quantify this processing as a potential source of error”.

Due to these potential sources of error in their LiDAR data, as well as an observed lack of LiDAR coverage in certain locations within their study area due to the wider scan angle and heavy vegetation, Pluckhahn and Thompson (2012) decided to supplement their LiDAR data with a targeted total station survey. Over the course of three to four weeks of fieldwork, Pluckhahn and Thompson collected several thousand additional elevation data points within Fort Center's major archaeological features where LiDAR coverage was lacking. After combining and interpolating the total station and LiDAR data
points within ArcGIS, Pluckhahn and Thompson were able to create a DEM of the Fort Center site featuring more accurate digital representations of the major archaeological features present at the site than a DEM composed of LiDAR points alone.

Because of the large area which the site covers, Pluckhahn and Thompson state that the DEM of Fort Center is not well suited for publication. Consequently, the authors instead rely on contour maps generated from their LiDAR DEMs in order to represent the topography of the site. As Pluckhahn and Thompson (2012) point out, the contour map produced by combining LiDAR and total station data dramatically improves the representation of this feature when compared to the map generated from only LiDAR data and the map redrawn from Sears’ (1982) original report.

Pluckhahn and Thompson (2012:11) eventually conclude that, given the size of the Fort Center site and the density of vegetation in the area, it would have been impractical to map the site using the total station alone. Combining total station data with LiDAR, on the other hand, allowed the authors to make a relatively more accurate map of Fort Center than previous efforts in both a shorter amount of time and at little cost (2012:12). However, Pluckhahn and Thompson also raise cautions with the use of LiDAR data that are very similar, if not the same, as those that I have already mentioned concerning the archaeological sites at CANA.

The authors state that the Fort Center case study clearly illustrates “the need for archaeologists to understand the manner in which ALS [LiDAR] data are collected and processed, both in general and specifically for the data they employ” (Pluckhahn and Thompson 2012:12). Pluckhahn and Thompson stress that publicly available LiDAR data
be field checked and even supplemented with elevation data gathered through conventional mapping techniques, such as total station survey, in areas where LiDAR coverage is low or nonexistent. Utilizing metadata, such as strike maps and field notes, is one of the techniques that can help target such areas, thereby preserving project resources.

As I have noted both in earlier chapters and in this chapter, local topography, vegetation, and archaeological features can vary immensely from one study area to another. Therefore, any processes developed to modify and/or supplement LiDAR data must be specific or adapted to the archaeological and environmental realities of a given location. Unlike other LiDAR data processing projects that I have come across during my research (Lasaponara 2011), the mapping work completed at Fort Center by Pluckhahn and Thompson (2012) is especially applicable to the field of Florida archaeology since the site being investigate is, after all, a Florida site.

Pluckhahn and Thompson (2012) mention throughout their article that dense, low lying vegetation, relatively flat topography, and archaeological features characterized by both slight and dramatic changes in elevation were all factors that had to be considered in the development of their data processing methods. These circumstances also ring true for many other archaeological sites throughout southern Florida, and CANA is no exception. The archaeological sites within my study area present issues similar to those faced by Pluckhahn and Thompson, though there are some difference as well. The environmental and archaeological factors that I must consider in developing my own LiDAR data processing methods become apparent in the next chapter, as I present both the cultural and environmental backgrounds for CANA.
Chapter 4: CANA Cultural and Environmental Background

One of the keys to correctly tailoring LiDAR data processing methods for archaeological purposes is an intimate knowledge of the study area and the historic or prehistoric culture/s in question. Therefore, if one hopes to use LiDAR as a tool for the purposes of preservation and management of archaeological resources within CANA, it is imperative that one be familiar with the landscape of CANA as well as with the major prehistoric culture present within its archaeological record. In this chapter, I will offer just such a background.

CANA is located within Volusia County, which is situated along the East-Central archaeological region of Florida as designated by archaeologists Jerald Milanich and Charles Fairbanks (1980:22). Figure 1 in chapter one shows the location of CANA in red. This region encompasses the St. Mary’s River to the north and extends to Vero Beach along the Atlantic Coast, including the St. Johns River drainage and most of the coastal lagoon associated with it. The southern boundary of this East-Central region is more ambiguous, however, and features a significant transitional zone with the circum-Glades archaeological region (Milanich and Fairbanks 1980:28).
4.1 St. Johns II Culture

As mentioned in a previous chapter, the predominant prehistoric culture witnessed in the archaeological record of CANA, and the culture to which my research sites of Turtle Mound, Ross Hammock, and Castle Windy are attributed, is known as the St. Johns II culture. In order to place the St. Johns II culture within a temporal context, I will follow the cultural chronology of prehistoric Florida as presented by Milanich and Fairbanks (1980: Table 1 and Table 4), who in turn synthesize the earlier work of other archaeologists in east Florida such as John Goggin (1947, 1949, 1952), Irving Rouse (1951), and Ripley Bullen (1972).

Changes in pottery and technology in Florida during the late Orange phase (1200-500 B.C.) mark the beginning of the Formative Stage, whereby fiber-tempered ceramics are replaced by limestone-tempered, sand-tempered, and temperless wares (Austin and Layman 1989:15). Milanich and Fairbanks state that this Formative Stage denotes “...a beginning of formal, settled communities, with the gradual development of more complex forms of political and religious community organization” (1980:20), and they go on to state that this stage is marked by a greater amount of regional diversity than previous stages (1980:20). This regional diversification, primarily attributed to local adaptation to varied ecological conditions within Florida, is conventionally described in terms of cultural periods that are based on variations in ceramic types (Austin and Layman 1989:15).

The ceramic tradition associated with Central and East Florida is known as the St. Johns cultural tradition and is divided into the following periods: St. Johns I (500 B.C. – 100 A.D.), St. Johns Ia (A.D. 100-500), St. Johns Ib (A.D. 500 -800), St. Johns IIa (A.D. 800 – 1300), St. Johns IIb (A.D. 1300-1513), and St. Johns IIc (A.D. 1513-1565) (Austin
and Layman 1989:15). The St. Johns I period is characterized by temperless pottery, both plain and incised, constructed through a coiling method. It is also during this period that low burial mounds first appear.

These burial mounds become more common during the St. Johns IIa period, though it is the emergence of check-stamped pottery that serves as the true diagnostic marker for this stage. St. Johns IIa pottery is either undecorated or check stamped, and is characterized by its chalky tactile quality, which is traditionally attributed to the natural occurrence of microscopic sponge spicules within the raw clays used by St. Johns potters, though it has been argued that freshwater sponges were purposefully added to St. Johns pottery as a temper (Rolland and Bond 2000).

The most extensive occupation of coastal lagoons occurs during the St. Johns IIa period, and archaeological sites associated with this culture include small resource gathering camps as well larger, more permanently occupied settlements. These larger settlements are marked by shell middens and a sand burial mound, typically 1 to 3m in height (Moore 1894, 1896; Russo 1992). My CANA study sites (Turtle Mound, Castle Windy, and Seminole Rest) are attributed to the St. Johns IIa period. Therefore, it is this stage, along with the cultural traits that characterize it, which will serve as the primary focus for my research.

St Johns IIa sites are generally concentrated around estuarine environments, and as such it has been suggested that their subsistence practices consisted of a year round coastal-exploitation focused on the capture of small estuarine fish, shellfish, and other aquatic fauna, both estuarine and freshwater (Russo 1992; Russo et al. 1993). Although
faunal remains of deer, raccoon, opossum, and aquatic reptiles such as turtles are found in limited quantities at St. Johns Ilia sites, the largest contributor of vertebrate biomass to the diet of St Johns II peoples were fish (Russo et al. 1993) However, there is no evidence suggesting the practice of ocean or deep-water fishing or sea mammal hunting (Ashley 2002). Of the shellfish species collected by St. Johns Ilia peoples, oyster was by far the most intensively utilized, although Atlantic ribbed mussel and whelk were also gathered from marshes, and coquina was procured from the ocean surf zone (Ashley 2002).

Flaked stone artifacts are typically infrequent at St. Johns Ilia sites; though when they do occur include small triangular points, bifacial and unifacial tools, and debitage, all of which consists of poor to moderate quality chert (Ashley 2002). The fact that there are no naturally occurring stone sources in Florida, in addition to the expedient nature of these lithic artifacts, suggest that they were made on site out of imported stone (Ashley 2002).

4.2 Previous Archaeological Research

The history of archaeological inquiry in Volusia County begins in 1605 with the journey of Spanish explorer Mexia down the Halifax, Mosquito, and Indian Rivers (Goggin 1952). During his travels, Mexia makes note of and even maps mounds of oyster shells that he considers to be simply areas where Indians lived. One such mound is Castle Windy, which had been abandoned before the arrival of Columbus (Austin and Layman 1989).

Though Mexia believes Indians may have once lived at these mound locations, archaeological remains are not uncovered at these sites until John Bartram and his son
William explore the St. Johns River in 1766 (Goggin 1952). On their journey, the Bartrams identify 24 archaeological sites many of which are shell middens. Despite the observations of the Bartrams, however, the actual construction of these shell mounds is not attributed to humans until James Pierce notes in 1824 that natural shell deposits contain shells with valves together (Goggin 1952) as opposed to the shells used for sustenance found in middens, which have been separated in order to access the meat inside.

During the second half of the 19th century, several researchers contribute to the study of prehistoric archaeological sites in Volusia County, such as Daniel G. Brinton, Jeffries Wyman, Andrew Douglass, and M. Francis LeBaron (Austin and Layman 1989). Yet, the most extensive archaeological excavations performed during this period are carried out by Clarence Bloomfield Moore, better known as C.B. Moore. Starting in 1891, C.B. Moore first investigates sites along the St. Johns River until 1894 and then works on the Indian River mounds until 1896 (Rouse 1951). Though C.B. Moore’s excavation techniques would not stand up to today’s standards, he carefully recorded his findings and published detailed reports of his archaeological inquiries (Moore 1892a, 1892b, 1894, 1896).

The works of C.B. Moore took place amidst a time when archaeology was just becoming an established vocation, and studies conducted by people such as Moore were a part of an emerging perception of archaeology as a scientific discipline during the second half of the 19th and early part of the 20th centuries (Austin and Layman 1989). The information collected by Moore and other such archaeologists helps to inform later archaeological investigations into the St. Johns and other prehistoric cultures within
Volusia County. One such inquiry is performed by N.C. Nelson, who, using data from the Oak Hill Site in southern Volusia County, produces the first stratigraphic sequence in Florida by tracing recognizable ceramic changes within the various strata of a shell midden (Nelson 1918).

Additional archaeological work in Volusia County was limited during the 1920s. However, in 1931 Gene M. Stirling explored 11 sites in the Cape Canaveral area by means of a grant from the Harvard Peabody Museum (Austin and Layman 1989). Gene Stirling was followed by his brother, Matthew W. Stirling, who continued this work at Cape Canaveral through his position as supervisor of the Civil Works Administration (CWA) program for archaeology in Florida (Austin and Layman 1989).

The beginning of the current period of modern archaeological study in Florida is credited by some (Goggin 1952) to James B. Griffin, who provides an overview of the St. Johns River area’s late Archaic ceramic producing cultures (J.B. Griffin 1945). In addition, J.B. Griffin also incorporates his data concerning the St. Johns region into the broader research questions of Southeastern U.S. archaeology (J.B. Griffin 1945).

The first chronological sequence constructed specifically for Volusia County was produced by John W. Griffin (1948). In composing this sequence, J.W. Griffin uses Nelson’s general Florida chronological data from Oak Hill mentioned previously. However, J.W. Griffin also includes data collected from surface scatters, site surveys, and excavations at three Volusia County sites conducted by himself and his Florida Park Service colleague, Hale G. Smith (J.W. Griffin 1948). Though J.W. Griffin’s sequence has...
undergone some adjustments since first being published, it remains essentially valid even today (Austin and Layman 1989).

Another work from around the same time in Florida archaeology that still serves as a primary reference for research regarding the prehistory of Volusia County is John M. Goggin’s *Space and Time Perspective in Northern St. Johns Archaeology, Florida* (1952). In this work, Goggin provides a synthesis of prehistoric cultural development in Northeast Florida. The portions of Goggin’s book that are devoted to Volusia County are informed by several archaeological surveys that he performed in this area.

Perhaps the most important figure in Florida archaeology during the 1950s and 1960s was Ripley P. Bullen (Austin and Layman 1989). Though the scope of his studies encompassed many other regions in Florida, as well as the Southeastern U.S., Bullen’s work in Volusia County included investigations at sites such as Castle Windy and Ross Hammock, among others. Through the data gathered by these investigations, in addition to analyses of materials collected from previous excavations by archaeologists at Volusia County sites, Bullen added new information on these sites’ stratigraphy as well evidence regarding subsistence, technology, and environment (Austin and Layman 1989).

### 4.3 Cultural Resource Laws

Around the time of Bullen’s contributions to Florida archaeology, American archaeology as a discipline entered a new period around 1960 that saw “an increased emphasis on anthropological theory, the use of more rigorous scientific methods, the application of computers for sophisticated analyses of archaeological data, the
development and testing of models, and interdisciplinary research efforts…” (Austin and Layman 1989:10). The impact of this “New Archaeology” is reflected in archaeological research conducted in Florida since the late 1970s. This new approach is contemporary with another prominent influence on archaeology during this period, the advent of Cultural Resource Management, or CRM as it is commonly known.

CRM is a blanket term that refers to any federal and state laws that require archaeological site assessment surveys prior to construction or development. Whether this construction takes place on federal, state, or private land, however, determines which laws apply and can vary from state to state. Regardless, Florida archaeology has witnessed a resurgence since the introduction of these laws and the implementation of CRM has resulted in more archaeologists being employed, archaeological surveys and excavations being conducted at faster rates, and the discovery of numerous archaeological sites (Austin and Layman 1989).

These impacts are as dramatic for Volusia County as they are for the rest of the state. CRM related assessment surveys have been conducted for private developments (e.g. Austin and Ballo 1989; Johnson and Ste. Claire 1988), public utilities and transportation (e.g. Griffin and Benton 1980; Hardin and Piper 1983), and federal lands such as CANA (e.g. Ehrenhard 1976; Taylor 1982). Many counties and municipalities, including Volusia county (e.g. Austin 1989), have proactively invested site inventories and developed management plans, archaeological sensitivity maps and predictive models in an effort to provide development planners with information concerning the number, kind and significance of sites in their areas (e.g. Piatek 1986). Even CRM excavations performed in order to salvage archaeological data prior to site destruction due to
development have resulted in significant new information about the region’s prehistory and history (e.g. Russo et al. 1989a, 1989b).

As mentioned previously, the cultural resource laws that apply regarding the development and/or management of land containing culturally significant sites depends on whose jurisdiction the land falls under. Different laws apply depending upon the nature of the culturally significant site and whether this resource is located on federal, state, or private land. Since CANA is federal land, which contains Native American burial sites and can be considered a cultural landscape overall since habitation sites are so pervasive, several laws federal laws apply. These include the National Environmental Policy Act (NEPA), National Historic Preservation Act (NHPA), Native American Graves Protection and Repatriation Act (NAGPRA), Archaeological and Historic Preservation Act (AHPA), Archaeological Resources Protection Act (ARPA) and the Historic American Landscapes Survey (HALS) (King 2008).

CANA is a national seashore, and as such it falls under management of the National Parks Service. Therefore, the NPS is responsible for the preservation of the cultural resources that CANA contains in a manner that complies with the federal laws and regulations stated above. However, the guidelines for preservation and documentation set forth by these federal laws can also serve as a set of “best practices” so to speak to help inform any general cultural resource management plans undertaken by managers of state and local lands as well.

For instance, of particular note in this case is the Historic American Landscapes Survey, or HALS. In addition to formally designed landscapes such as gardens, parks,
and campuses, HALS also attends to what are usually called “cultural” landscapes (King 2008). These cultural landscapes are those that reflect human activities like fishing or hunting in some more or less unplanned manner. The shell middens indicative of Native American habitation along much of the Florida coastline certainly fall under this designation. Furthermore, and of particular relevance to my thesis, the Cultural Resources Geographic Information System (CRGIS) initiative – which promotes GIS in the national parks and State Historic Preservation Officer (SHPO) offices- is included with HALS under the “Heritage Documentation Programs” designation by the NPS (King 2008).

The utility of GIS and remote sensing methods, of which LiDAR is one, is mentioned throughout the guidelines for HALS (Robinson et al. 2005; The Jaeger Company 2005). For example, the HALS History Guidelines specify current site features of a landscape that should be documented are spatial organization, visual relationships, natural systems, topography, vegetation, water/hydrology, and buildings/structures (Robinson et al. 2005). As was demonstrated in several case studies in the previous chapter, LiDAR is a proven, invaluable tool for recording and analyzing these features of a landscape. The guidelines acknowledge this as well, stating that “techniques such as GIS, GPS and remote sensing …prove useful in the site documentation” (Robinson et al. 2005:8).

While the advantages of using LiDAR for archaeological site analysis and management are evident, the methods by which to process LiDAR data in a manner that benefits the user for these purposes are not abundantly clear. Concordantly, the next chapter proposes techniques that will allow archaeologists and park managers to quickly
and efficiently produce topographic models from publicly available LiDAR data that accurately document Native American shell middens and burial mounds.
Chapter 5: Methods and Results

Nowadays, LiDAR data is becoming increasingly available to the public in the United States. This publicly accessible data is usually generated by state or federal agencies for environmental management and emergency preparedness purposes. The LiDAR data that I will be using for my research originates from just such a source. The Florida Division of Emergency Management (FDEM) has been collecting LiDAR data for Florida coastal areas for several years, and has made this data available on multiple websites. One such site is hosted by the International Hurricane Research Center at the Florida International University (Pluckhahn and Thompson 2013), and another is the National Oceanic and Atmospheric Administration’s (NOAA) Digital Coast website. For the purposes of my project I chose to acquire my LiDAR data from the NOAA Digital Coast website simply because I am more familiar with its interface.

5.1 Data Acquisition

Though the FDEM datasets are not the only LiDAR data available for my study area through NOAA, I chose the FDEM source because it is in the LAS format. I will address the benefits of the LAS format later in this chapter. First of all, however, in order to obtain the appropriate data for my sites of interest within CANA it is necessary to be familiar with their geographic location relative to the park boundaries. This is especially
true considering the nature of the data viewing and downloading tools featured on many government funded websites such as on the U.S. Geological Survey (USGS) National Map Viewer (TNM) and the NOAA Digital Coast Data Access Viewer. These sites require the user to place a border around their area of interest user either a topographic map or aerial image provided by the data viewer as a base map.

The data viewer for the NOAA Digital Coast website is shown in Figure 2 below. The base map shown is an aerial image, and the yellow box in the image marks the area that I wish to acquire data for. The area in this case happens to encompass the site boundary for Castle Windy. As one can see, unless the user is very familiar with the defined boundaries of the archaeological site in question, it can be difficult to delineate an area on the data viewer that fully covers the boundaries of the site.

![Figure 2 NOAA Digital Coast data viewer](image-url)
This is especially true if the area of interest features thick and relatively unbroken vegetation cover, as is the case for all of my study sites within CANA. Therefore, to assist in the delineation of my areas of interest within the NOAA data viewer, I created a guide through the Google Earth program that clearly shows the site boundaries in question. I acquired a shapefile produced by the National Parks Service from AIST that features the site boundaries of all known archaeological sites within NPS parks. In order to eliminate superfluous information related to other sites within both CANA and other national parks, as well as convert the data into a file format that can be displayed within Google Earth, I first imported the NPS shapefile into the ArcMap program. I then extracted the information regarding my study sites from the NPS archaeological sites data by selecting the relevant data using the select tool and exporting my selection to a new shapefile that I labeled “Griesbach_NPS_CANA_Study_Sites”.

Using the “Convert to KML” tool, this shapefile was converted to a kml file called “CANA_FMSF” that could then be imported into Google Earth. The site boundaries were changed to a neon green color so as to stand out against the aerial imagery provided by Google Earth, as shown in Figure 3 on the next page. With the help of this Google Earth guide, I was able to identify, select and download from the NOAA Digital Coast website the relevant data for my areas of interest within CANA. Though LiDAR data is also available from the United States Geological Survey website, I chose to attain my data from NOAA because it is the only government site I know of that features LiDAR data available for download in the LASer (LAS) format. All of the LiDAR data for my study sites was acquired in the LAS format from the 2006 Volusia County LiDAR dataset.
The LAS format is defined as “…a public file format for the interchange of three-dimensional point cloud data between data users” (ASPRS 2012, accessed October 15, 2013). LAS is a binary file format released by the American Society for Photogrammetry and Remote Sensing (ASPRS) that offers an alternative to generic ASCII file interchange and proprietary LiDAR data formats (ASPRS 2012). The obvious issue in using
proprietary formats is that they cannot be easily transferred from one processing program to another. In addition, ASCII elevation data file sizes are often quite large even for small datasets. This can result in data storage issues and very slow computer processing speeds when attempting to read or interpolate LiDAR data in this format. Raw data within ASCII format elevation files relevant to the collection process of the LiDAR data is also lost when it is processed, making it difficult to troubleshoot processing issues.

The LAS format addresses many of these issues, as it retains information specific to the LiDAR nature of the data while at the same time not being inordinately complex. LiDAR data points in the LAS format are organized through a classification field, which can be used to separate the data according to the following categories and codes (ASPRS 2012): unclassified (1); ground (2); low vegetation (3); medium vegetation (4); high vegetation (5); building (6); low point (noise) (7); and water (9). These categories and codes are established through conventions set forth by ASPRS, which helps keep data point classification consistent across projects that use the same LiDAR data. Because of this uniformity, efficiency, and detail, LAS is considered the LiDAR industry standard data format (ASPRS 2012) and is therefore the data format that I will be using for my analyses.

From the 2006 Volusia County LiDAR dataset, I acquired data for the Seminole Rest, Ross Hammock, and Castle Windy sites and surrounding areas. Another advantage of the NOAA website is that it allows the user to specify what sort of preprocessing he or she wishes to be performed before the user receives the data. One of the primary subjects of my thesis is to investigate how the pre-processing of LiDAR data can affect LiDAR related investigations with regard to the identification, management, and analysis of archaeological sites in Florida. Comparing the final products of LiDAR data that have
undergone different sorts of pre-processing will help identify these effects. Therefore, I downloaded two distinct LAS datasets for each of my respective CANA sites. Each of these datasets underwent a different degree of preprocessing.

The first dataset collected for each site underwent the least amount of preprocessing. On the NOAA Digital Coast data downloader tool, this option is described as “Unclassified”, and the return type is designated as “Any”. “Unclassified” means that the dataset will contain almost all of the data points available for my designated study area. These data points are not classified into categories such as “Ground” or “High Vegetation” based on algorithms used by NOAA.

A return type of “Any” signifies that all of the data points will be included within the dataset, regardless of whether they are designated as “First” returns, “Last” returns, or anything in between. As was discussed in chapter 2, each infrared laser pulse sent out by the LiDAR receiver during data collection generates one or several echoes depending on what surfaces that pulse strikes before coming to a surface that the pulse cannot pass through. The first echo that a pulse generates is referred to as the first return, and the last echo generated is the last return.

By specifying “Unclassified” for the data points and “Any” returns on the data downloading tool, the user is stating that he or she wants to receive all of the echoes generated by all of the pulses for the given area of interest. Having datasets such as these for all of my study sites provides me with LiDAR data that is as close to “raw” as I can get. With this data, I can perform my own pre-processing methods that, in some cases, provide a better end product that will be of better use for archaeological purposes.
However, I will not know if my pre-processing methods are better than those provided by NOAA unless I have a dataset to compare them to. Therefore, the next datasets that I obtained for each study site were ones featuring data points classified as “Ground”, with return types set to “Last”. Because a laser pulse cannot pass through the ground, the last echo is usually associated with the ground, a man-made object, or sturdy vegetation like the trunk of a tree. Ergo, by setting the return type to “Last” and the data point classification to “Ground”, I can achieve as close an approximation of bare-earth topographic models for each of my study sites as can be attained through the use of the pre-processing algorithms offered through NOAA.

5.2 Metadata

Though they may contain different data points, all of the datasets for each study site were projected in State Plane 1983 Coordinate System. In addition, the State Plane zone for each dataset was set to Florida East, the Horizontal Datum to NAD83, the vertical datum to NAVD88, and both horizontal and vertical units were set to U.S. feet. I also use this coordinate system for any products derived from the LiDAR data. I chose the State Plane Coordinate System (SPCS) as the geographic projection for my data because it is one of the most accurate coordinate systems considering the scope of my study area.

The metadata PDF file (O’Neill 2007) coupled with the LAS files that I downloaded for the CANA sites indicates that this data is part of a larger LiDAR dataset collected by Woolpert, Inc. at the request of the Volusia County Public Works Department for the purpose of supporting the 2006 Volusia Countywide Digital Orthophoto Imagery Project.
The area scanned covers approximately 1,432 square miles and is, of course, centered on Volusia County. A total of 143 flight lines of LiDAR data were acquired in eleven sessions from March 2-8, 2006 (O’Neill 2007). The data were collected at an average point spacing of 1 meter (3.3 feet) in order to develop a DEM suitable for generating 1 foot contours that meet 2 foot contour accuracy specifications (O’Neill 2007).

The metadata states that this LiDAR dataset is intended for orthophoto rectification purposes only, and that it is not intended for “engineering, design, water management, or any other purposes other than as stated within this document” (O’Neill 2007:1). It is clear that this dataset conforms to the generalizations concerning LiDAR data acquired for government projects (Pluckhahn and Thompson 2013). Orthophoto rectification is a relatively larger scale project when compared to other sorts of endeavors, such as engineering and, notably, archaeology. The data point sampling distance of 3.3 feet supports this assumption, as a higher resolution is usually preferred for the documentation of archaeological features (Coluzzi et al. 2010; Chase et al. 2011; Masini et al. 2011). Though as previous case studies have shown (Pluckhahn and Thompson 2013), and as I intend to demonstrate through my thesis, this should not deter the use of publicly available data for archaeological purposes.

5.3 Data Processing

After downloading the appropriate data from NOAA, my next step is to begin processing the unfiltered LiDAR data. Typically, if one were trying to create a bare-earth DEM of a given area they would download LAS data already classified as representing
ground data points. As has been stated previously, this classification is a result of an algorithm automatically reclassifying the LAS LiDAR data based on slope in order to group certain data points within a discrete ground class. This is an issue with regard to Native American middens and mounds in Florida, since the relatively dramatic increase in elevation indicative of these archaeological features is often mistakenly identified by slope-based algorithms as modern man-made structures or vegetation. Hence, the data points for these features are eliminated from the collected LiDAR dataset before being given to the researcher. My intention is to assume control of this reclass process, and the tool that I use to do so is known as “LASthin”, which is a script within the LAStools ArcMap tollbox.

LAStools is a set of scripts executable within ArcMap that allow the user to modify and process raw LAS data. LAStools was developed by rapidlasso GmbH, a company based out of the University of North Corolina. LAStools is free to download from the rapidlasso website, and is described on said website as a “software suite [that] has deep market penetration and is heavily used in the commercial sector, government agencies, research labs, and educational institutions alike” (http://rapidlasso.com/).

Though there are several useful scripts within LAStools, LASthin is the one that features the most utility for my research. As the name implies, LASthin is a thinning algorithm for LAS, LAZ, and ASCII LiDAR data formats. The official description for this tool on the rapidlasso website sates that LASthin “places a uniform grid over the points and within each grid cell keeps only the point with the lowest…Z coordinate” (http://rapidlasso.com/). Essentially, what this allow user to do is filter out the LiDAR data points within a grid cell, leaving only the point with the lowest z (elevation) value. All of
the lowest points within these grid cells are then selected out of the LiDAR and used to create a new data point layer that is hopefully more representative of the ground surface. This also preserves the original data point layer, allowing the user to easily run the operation multiple times with different parameters.

One of the by-products created through running LASthin are error points. Error points come about when the lowest data point in a grid cell does not actually represent the ground surface. This may occur in areas of dense vegetation where none of the laser pulses were able to reach the ground surface during data collection. Usually, these error points stand out from the surrounding data points and hence are easy to visually identify and manually eliminate from the dataset after the LASthin process has taken place.

The advantage of the LASthin method is that it circumvents the issues experienced when basing ground classification off of slope. As has been mentioned by myself and others (Pluckhahn and Thompson 2013), South Florida presents a unique problem with slope based LiDAR ground classification because of the relatively flat topography in this region. Archaeological features, such as mounds and middens, often exhibit drastic changes in slope from the surrounding landscape, and because of this they are sometime classified as off-ground by slope-based classification algorithms and hence eliminated. LASthin, on the other hand, simply takes the point within each grid cell that has the lowest z value and assumes this represents the ground.

The key element in this operation is determining a suitable grid cell size. The official description of the grid size parameter that is given in the LASthin script is as follows:
“Specifies the granularity of the grid that the LiDAR points are thinned with. If the grid size is set to 1 then maximally 1 point per unit squared will survive. If the grid size is set to 0.5 then maximally 4 points per unit squared will survive. If the grid size is set to 2 then maximally 1 point will survive for every 4 square units” (rapidlasso 2014).

The larger the grid size used for this operation, the greater the loss in resolution of the LiDAR data. If too large of a grid size is specified, then the landscape may be excessively smoothed out, effectively erasing any fine topographic details. On the other hand, though smaller grid cell sizes better preserve the resolution of the original dataset, small cells also take less data points into consideration when determining the point with the lowest z value. This increases the likelihood that the lowest point does not actually represent the ground surface, which in turn can create more work for the user with regard to manually cleaning up the error points after the LASthin operation has completed. Generally, one should start with a small grid size and work their way up until mainly the desired data points are being returned with only a few errors.

Following this principle, the first step in my data processing chain was to run the unclassified LiDAR points for each of my study areas through the LASthin process multiple times at various grid cell sizes so as to determine which cell size produces the best representation of the ground surface for each site. This step, along with my entire data processing model, is visible in Figure 4. Fortunately, each of my study areas are of a small enough size that the LAS data did not have to be broken into multiple tiles in order to facilitate downloading from the NOAA website. However, for larger study areas where the LAS data has been tiled, LAStools features another script known as LASmerge, which allows the user to merge multiple LAS data files into a single file.
Starting with a grid cell size of 2, I ran the LAS data for each of my study sites through the LASthin tool multiple times, each time increasing the grid cell size by 2. I then converted each of the resulting datasets into a multipoint featureclass using the “LAS to Multipoint” tool within the 3-D Analyst toolbox. 3-D Analyst, along with all of the other tools that I use in my model aside from LASthin, comes preloaded as part of the ArcMap package.

This step is necessary in order to view the point layers within ArcGIS and compare the different thinning grid cell sizes. However, though the LASthin grid cell size that I eventually use and some of the parameters that I choose for the LAS to Multipoint conversion, as well as several other steps in my data processing methods, vary depending on the CANA study site, the overall processes that I use and their order do not change.
I was able to apply this model to Castle Windy and Ross Hammock with varying degrees of success, each step of which I will discuss later in this chapter. For Seminole Rest, however, I immediately ran into an issue after downloading the data that prevented further analysis. While this was disappointing for my research, I believe that it does provide a cautionary tale of sorts for archaeologists and park managers hoping to utilize publicly available LiDAR data for archaeological purposes.

As stated earlier in this chapter, I acquired two sets of LAS data for each of my study sites; a “raw” set of unclassified points, as well as a set of points classified as “ground” by NOAA for comparison. Since the unclassified points represent the entirety of the LiDAR data points collected for a given area, they should then cover not only those areas covered by the ground classified points, but also any areas that the ground points do not extend to, such as vegetation and buildings. In the case of Seminole Rest, however, it appears that the “unclassified” data points are incorrectly labeled within the NOAA dataviewer. As one can see in Figure 5 on page 66, there appear to be gaps in the coverage of the red, “unclassified” LiDAR data points, particularly in the northern portion of the Seminole Rest site. Furthermore, these same gaps do not appear in the green, “ground” classified points. Concordantly, since unclassified points should represent all of the LiDAR points collected for a given area, it can be assumed that the unclassified points are incorrectly labeled.

In order to ensure that this was not simply due to user error, I re-downloaded the “unclassified” and “ground” LiDAR points for Seminole Rest from the NOAA Digital Coast website on 01/25/14. Once again, the same gaps appeared in the “unclassified” LiDAR data point coverage as previously observed. Since obtaining truly unclassified data points
is crucial for my methods, I was forced to abandon Seminole Rest as one of my study sites. My other CANA study sites, however, do not display this same discrepancy, despite the fact that the LiDAR data for all of Volusia County was collected for the same project in 2006. This is because LiDAR data collection and processing is often split into separate “missions” for projects that cover large areas, such as counties. Therefore, it is likely that the data points collected for the areas that include Castle Windy and Ross Hammock were part of different collection missions than those for Seminole Rest, and hence may not be subject to the same error.

Nevertheless, this experience serves as a cautionary tale not only for park managers and archaeologists, but to anyone who wishes to use publicly available LiDAR for any purpose. As we have seen, it is important to become familiar with LiDAR data and to maintain a critical eye when receiving it from a public source. This includes learning what potential errors look like when evaluating and double checking the data.

Fortunately for my purposes though, I still have two more study sites at CANA that do not appear to exhibit any such errors in their respective LiDAR datasets. However, even though the Castle Windy and Ross Hammock sites reside within the same national park and are only roughly 4.5 miles apart, differences in the degree of vegetation cover present at each, as well as the nature of the archaeological features themselves, necessitates certain adjustments in my data processing methods for their respective LiDAR points. Therefore, I will highlight any differences in the application of the data processing model shown in Figure 4 to the sites of Castle Windy and Ross Hammock throughout the remainder of this chapter.
Figure 5 A) Seminole Rest LiDAR data points that appear to be mistakenly labelled as “unclassified”. B) Seminole Rest LiDAR data points automatically classified as “ground” through slope-based algorithms.
The relatively large areas of both the Castle Windy and Ross Hammock site extents, 653,000 square feet and 356,000 square feet respectively, featured in the NPS shapefile posed a problem early on while processing their LAS data. As was mentioned previously, when using the LASthin tool, one wants to strike a balance between grid cell size and the amount of error points featured in the final product. In the case of Castle Windy and Ross Hammock, I did not reach this balance until I thinned the LAS points for each by a grid cell size factor of 8. This was surprising, as I had initially thought that I would not need to go above a grid cell size of 4 at the most, and that anything above this would sacrifice too much resolution to be of any use.

However, once IDW interpolated and viewed as a DEM for quick comparison, thinned datasets featuring a grid cell size of less than 8 featured a high degree of noise. This noise is readily apparent in Figure 6 on the next page, which compares the Castle Windy DEMs produced from LAS datasets featuring thinning grid cell size factors of 4 and 6, with a DEM derived from the NOAA ground classified points. The numerous points of high elevation viewable in the thinned datasets are most likely contributable to tall, dense vegetation. Vegetation interference is to be expected, considering that CANA is in a subtropical environment and that the LiDAR dataset that I am using was collected during the month of March (O’Neill 2007), a time of the year in which foliage cover is increasing.

Though thinning the LAS datapoints by a factor of 8 helped clarify the ground signature in some areas, and it is at this stage that the forms of the Castle Windy midden and Ross Hammock mounds become clearer, Figures 7 and 8 reveal that vegetation still obscures the bare earth returns throughout the Castle Windy and Ross Hammock site areas.
Figure 6 A) DEM of Castle Windy derived from “Ground” classified NOAA LAS data points. B) DEM derived from unclassified, any return LAS data points. C) DEM derived from unclassified LAS data points thinned by a factor of 4 using LASthin tool. D) DEM derived from unclassified LAS data points thinned by a factor of 6.
Figure 7 DEM of Ross Hammock derived from “Ground” classified NOAA LAS data points. B) DEM derived from unclassified, any return LAS data points. C) DEM derived from unclassified LAS data points thinned by a factor of 8 using LASthin tool. D) Aerial photo of Ross Hammock.
Figure 8 A) DEM of Castle Windy derived from “Ground” classified NOAA LAS data points. B) DEM derived from unclassified LAS data points thinned by a factor of 8 using LASthin tool. C) Aerial photograph of Castle Windy depicting total station and RTK GPS points collected by AIST. D) Aerial photograph depicting Castle Windy midden extent.
However, based on my personal experience of excavating at Castle Windy, I know that the Castle Windy midden itself does not occupy the entire site area as defined by the NPS CANA site layer that I am using. Therefore, in an effort to more clearly define the bare-earth signatures associated with the midden, the most topographically prominent feature of the Castle Windy site, I decided to extract from the grid size 8 thinned dataset only those data points that are in the immediate area of the Castle Windy midden proper.

In order to accomplish this, I first defined the Castle Windy midden area by creating a shapefile that I could use to extract the relevant data points with. In creating this shapefile, I drew upon my own personal excavation experience as well as robotic total station and real time kinematic global positioning system (RTK GPS) points gathered at Castle Windy in August of 2011 by AIST. These points are shown on top of an aerial image of the Castle Windy midden in Figure 8 above. Many of these points are labeled with short descriptions of what sort of features they represent, such as the foundation corners of an old bait shop that once resided next to Castle Windy.

The points relating to the midden extent are highlighted in red within Figure 8. These points, as well as the DEMs produced from the ground points and factor 8 thinned LAS points, provided a guide when creating the shapefile depicting the Castle Windy midden area. This process was not necessary for Ross Hammock, however, since this site area includes several features of archaeological interest, such as St. Johns sand burial mounds in the southern and central regions of the site, as well as a shell midden ridge to the north.
I then converted the thinned LAS data points for Ross Hammock and the Castle Windy midden into a format that I can use to extract the data points I want and which would facilitate DEM interpolation later. The “LAS to Multipoint” tool within the ArcMap “3D Analyst” toolbox accomplishes just this, as it allows the user to convert the LAS file data point into a point shapefile. However, before I can select the individual points that are relevant to my study, another step is necessary. The LAS to Multipoint function produces a multipart shapefile, which essentially groups the data points together. This means that I would not be able to select just those points that fall within the site boundaries. Instead, I would only be able to select all of the points within the dataset. In order to circumvent this issue, I used the “Multipart to Singlepart” tool within the “Data Management” toolbox in ArcMap, which basically breaks up this group and creates a duplicate point shapefile that is made up of individual data points.

Once this was accomplished, I used the “Select by Location” tool within ArcMap to select all the data points within the grid size 8 point layer that fell within the confines of the Ross Hammock site boundary as defined by SEAC, as well as those within the Castle Windy midden extent shapefile that I created earlier. I also did the same for the data points derived from the ground classified LAS data for comparisons sake.

Yet, even after limiting the thinned data points to just those within the respective site boundaries, it is apparent that some of these points still do not represent the ground surface of the study area. Based on the elevation readings provided by the total station data collected by AIST at Castle Windy, I know that the maximum elevation of the Castle Windy midden is approximately 15 feet. This height is further corroborated in a report by Ripley P. Bullen and Frederisk W. Sleight (1959), in which the maximum height for the
Castle Windy mound is given as 17 feet. Unfortunately, I do not have any total station or RTK GPS data points for the Ross Hammock site, though there is another report produced for Ross Hammock, this time by Ripley P. Bullen and Adelaide K. Bullen (1967), which states that the maximum height for the taller, southern mound is 21 feet. This information will help me eliminate those points that fall above these elevation thresholds and give a clearer outline of the sites' topographies in the subsequent DEMs.

However, in order to select out the relevant data points, it is first necessary to assign elevation information to these points. Though the data points already contain z coordinates, this data is not visible within the attribute table for the data point layer within ArcMap, therefore it is not initially possible to select data points on the basis of elevation values. I use a tool within the ArcMap 3D Analyst toolbox, known as “Add Z Information”, to extract the elevation data from the points and add a z-coordinate field within the attribute table for my data point layer. After this, I was able to select those data points that have an elevation of 17 feet or less for Castle Windy and 21 feet or less for Ross Hammock, and then export these points as a new shapefile.

Though this new data point layer is getting closer to a bare-earth representation of the Castle Windy and Ross Hammock sites, there are still a few obvious error points here and there that most likely represent some form of high vegetation. In order to facilitate the manual selection and deletion of these final error points, I used the ArcScene program to achieve a three dimensional view of the data points. Viewing the points within a 3-D space helps provide perspective when evaluating the elevations of the data points relative to each other. This is evident in Figure 9 on the next page, which presents a screenshot of the Castle Windy midden points as viewed from an oblique angle within ArcScene.
In the case of Castle Windy, I used the total station and RTK GPS points collected by AIST and the thinned LiDAR data points themselves as guides. Although less than a dozen total station and RTK GPS points were collected for Castle Windy, this data proved invaluable during this process, particularly regarding the extent of the shoreline.

Yet, since no total station or GPS points were collected at Ross Hammock, I instead in this instance had to rely solely on comparing the relative elevations of the data points themselves. This, in combination with the large area that the archaeological features within the Ross Hammock site cover, made the manual elimination of error data points much more difficult for this site than Castle Windy. Initially, I tried downloading the “high vegetation” classified points from NOAA in order to possibly identify and evaluate any points within the thinned dataset that may also be within this vegetation set. In this manner, I had hoped to quickly eliminate any error points. However, none of the thinned data points fell within the vegetation dataset. This may be because the vegetation dataset from NOAA is actually considered “high” vegetation. Therefore, I may have already
eliminated the high vegetation points from the thinned dataset by only extracting those points below the maximum height of the Ross Hammock mounds.

For both study sites, I also color coded the thinned data points based on their z values and produced preliminary digital elevation models so as to identify and eliminate from the dataset those points that represent a dramatic increase in slope. Though this may sound similar to the slope based algorithm classification process that I am trying to avoid, the difference in this instance is that I am looking for much more dramatic increases in slope over a smaller section of the landscape. For example, considering that Castle Windy has a maximum elevation of 17 feet, if a data point has an elevation of 15 feet and is located directly adjacent to a data point with an elevation of 3 feet, or if the higher point is not anywhere near the center of the midden, then this point is most likely an error.

This process is rather subjective and requires a great deal of familiarity with the site in question, the site’s surrounding topography, and with interpreting digital elevation models. The field work I participated in at CANA in conjunction with the NPS and AIST, as well as the GIS and LiDAR experience I have gained while working at AIST over the last few years, greatly aided my efforts to remove any error points. Using this method, I was able to visually locate and delete data points that do not appear to follow the natural curvature of the landscape’s topography. All in all, I eliminated approximately 200 error points from 750 overall thinned data points for Castle Windy and 2,500 error points from 5,488 thinned data points for Ross Hammock.

However, despite having been thinned considerably through the LASthin tool and manual deletion of error points, Figure 10 on page 77 illustrates that the resulting data
points for Castle Windy are still sufficiently dense relative to the data points classified as "ground" points by NOAA, especially around the areas of highest elevation near the apex of the midden. Considering the error points that were manually deleted, along with those points removed because they fell above the 17 foot maximum height, a total of 386 remain in the thinned LiDAR data set, compared to 237 in the "ground" classified points. Therefore, the point density actually increased using the thinning methods for Castle Windy.

On the other hand, the opposite is actually the case for Ross Hammock. Unfortunately, considering the overall size of the Ross Hammock site and number of points within this area, it is difficult to appreciate any differences in point density between the thinned and "ground" data points on an 8.5 x 11 inch map. Looking at the point count, however, the 4,490 points present in the "ground" dataset is considerably higher than the 2,270 data points left in the thinned dataset. This low point density would also cause a problem during the next step, interpolation.

Once obtaining data point layers of Castle Windy and Ross Hammock that I believe most closely represent these sites, the next step is to interpolate the data points and produce a DEM of the study area. The two interpolation methods that I compare, IDW and Kriging, were discussed in chapter 2. First I tried the IDW method with the Castle Windy data points. The main advantage of this method is that it utilizes less processing power than Kriging, making it ideal for larger datasets. However, since neither Ross Hammock nor Castle Windy are relatively large datasets, processing time was minimal when interpolating for both was minimal regardless of whether IDW or Kriging was used.
Figure 10 A) Unclassified, manually cleaned LiDAR data points thinned by a factor of 8 using the LAStools LASthin function. Total station and RTK GPS points collected by AIST are shown in red. B) LiDAR data points automatically classified as “Ground” through slope-based algorithms.
The DEM resulting from the IDW interpolation method for Castle Windy did not appear to adequately represent the contours of the midden feature, particularly with regard to the steep shoreline present along the northeast corner of the midden. This is to be expected, since the IDW method tends to smooth over sudden drops in elevation due to the interpolation weighting that it assigns based on relative point distance. Considering this, along with the fact that processing time is not a factor in this instance, I decided to proceed to the Kriging method next.

The key factor to consider when using the Kriging method of interpolation is the model that one chooses to form the empirical semivariogram which predicts the unsampled points. There are five of these functions offered within the Kriging tool in ArcMap, they are circular, spherical, exponential, Gaussian, and linear. These functions greatly influence point prediction, and the function which most closely approximates the particular topographic phenomena observed in the given landscape must be selected in order to achieve the best results.

I experimented with all of the Kriging functions when interpolating the data points for both Castle Windy and Ross Hammock. In the case of Castle Windy, I eventually determined that the exponential semivariogram produced the clearest DEMs, since this function handles steep elevations particularly well as it retains the topographic detail of such features. Figure 11 on the next page compares the final DEM produced through the thinning methods outlined here with a DEM interpolated using the NOAA “ground” classified points. A further discussion of these results appears in the next chapter.
Figure 11 A) Digital Elevation Model (DEM) produced with unclassified LiDAR data points thinned by a factor of 8 and manually cleaned of error points. Total station and RTK GPS data points collected by AIST are shown in red, with labels denoting landscape features where points were taken. B) DEM produced with LiDAR data points automatically classified as “Ground” through slope based algorithms.
Ross Hammock, on the other hand, presented a problem when I attempted to interpolate the thinned data points. Neither the IDW nor Kriging interpolation methods worked for the Ross Hammock thinned points, as every attempt to run these methods resulted in an error message stating it’s “not enough points”. This is due to the significant number of points removed through the thinning and manual elimination phases discussed earlier.

However, it was at this stage that I decided to try a new technique for increasing the point density at both sites. Instead of producing DEMs of only the thinned data points, I tried combining the thinned data points with the “ground” classified points in order to raise the point density of both of my study sites and hopefully increase the spatial resolution and topographic detail. The error in the “ground” points that I have cautioned about throughout this thesis does not derive from these points themselves incorrectly representing the ground, but simply due to the fact that often there are not enough of them within a “ground” dataset to produce a DEM that adequately represents the ground surface.

Therefore, one may conceivably increase the ground point density by supplementing the data points classified as ground through algorithms with the manually thinned data points. Through this process, I was able to utilize the thinned data points derived from the Ross Hammock dataset in the creation of a DEM for the site. The data processing chain that I used for this method is revealed in Figure 12 on the next page.

The process outlined in Figure 12 above takes place after the thinned points have been manually cleaned of errors. The first step is to convert the “ground” classified points
from multipart to singlepart features. Next, these points are clipped by the extent shapefile of the mound in order to extract only those points that are relevant. Following this, the “Add Z Information” tool is used to add elevation data as a separate column within the point attributes.

Figure 12 Data processing model for combining thinned data points with “ground” classified points.

The next step is to select only those ground points that do not intersect the thinned data points and export these as a separate shapefile. This is to prevent having duplicate data points within the final dataset, as this could impart an undue degree of weight to these duplicated points during the interpolation process. The last step is to merge the exported ground points with the thinned data points to produce the final, combined dataset which can now be interpolated.
The datasets shown in Figure 12 are from the Castle Windy site. Unfortunately, however, the DEM produced from the combined ground and thinned data points did not feature any appreciable difference when compared to the DEM interpolated from the thinned points alone. Conversely, when the ground and thinned data points were combined for Ross Hammock, the increase in point density meant that the resulting dataset could now be interpolated without error. The differences between these Ross Hammock DEMs also appear minimal, though there are some slight changes discernible at the apexes of the mounds, as shown in Figure 13 on the next page. These and the other results illustrated earlier, however, will be discussed in further detail in the next chapter.
Figure 13 A) DEM of Ross Hammock mound 1 produced with unclassified LiDAR data points thinned by a factor of 8 and manually cleaned of error points. B) DEM of Ross Hammock mound 1 produced with LiDAR data points automatically classified as "ground".
Chapter 6: Discussion and Conclusion

The variability in the results achieved through the application of the thinning methods outlined in the previous chapter to each of my CANA study sites reveals not only the utility of these methods to the archaeologist and/or park manager, but also the need for further LiDAR mapping of both state and federal cultural resources. The complete lack of unclassified points for Seminole Rest, and the mislabeling of LiDAR data that this signifies, is perhaps the most conspicuous of my results. Though this unfortunate occurrence limited the number of archaeological sites to which I could apply the point thinning methods to, it serves to further highlight the advantages that LiDAR data imparts to the study and management of archaeological resources.

These assets are most apparent for the study site where I believe the thinning methods produced the best results, Castle Windy. As Figure 11 in the previous chapter reveals, the DEM interpolated from the thinned data points gives a much more detailed representation of the Castle Windy midden proper than the DEM derived from the NOAA ground classified points. This representation, as the overlaid total station and RTK GPS points collected by AIST illustrate, is more accurate as well. The two RTK GPS points labeled “Mound” in part A of Figure 11 were collected from the apex of the midden, and as the underlying DEM reveals, the thinned data points capture this southern aspect of the midden, whereas the ground classified point completely miss it.
The same is true along the entire western edge of the midden, as the total station and RTK GPS points in Figure 11 indicate where the shoreline and end of the midden is actually located. Again, the thinned point DEM follows these total station and RTK GPS much more closely than the ground classified DEM. Furthermore, the fact that, in this case, the shoreline is what is being misrepresented is very significant from a site management and preservation perspective.

Since Castle Windy is directly adjacent to the shoreline of Mosquito Lagoon, erosion is perhaps the most important factor to consider when developing a management plan for this site. Possessing a map that accurately depicts the site’s topography is crucial in this endeavor, as it can reveal how much of a threat erosion poses and what portions of the site are most susceptible. This information is relevant to Castle Windy since efforts have been and currently are being made to stem the level of erosion occurring at the site as a result of wakes created by recreational boating in the lagoon.

Part of the plan to address this issue is the creation and placement of oyster mats along the tide line bordering Castle Windy in order to help stabilize shore. However, knowing precisely where to position these mats in order to achieve the maximum benefit is paramount to this endeavor. Though the thinned point DEM displayed in Figure 11 can provide assistance in this regard, a slope model, such as the one shown in Figure 14 on the next page, can contribute even more useful information than a DEM. Slope models reveal the degree of elevation change throughout a given area, and are instrumental in determining areas most susceptible to erosion.
Figure 14 A) Slope model produced with unclassified LiDAR data points thinned by a factor of 8 and manually cleaned of error points. B) Slope model produced with LiDAR data points automatically classified as “ground” through slope based algorithms.
Figure 14 compares the slope models derived from both the thinned LiDAR data points (A) and the NOAA ground classified points (B). Once again, the manner in which these two models depict the Castle Windy midden differ greatly, though the thinned point slope model gives a more accurate representation of the midden, to which Figure 11 testifies. In terms of assisting erosion management and site conservation efforts, this thesis has demonstrated that knowing the correct location and degree of slope is critical, and deriving this information from LiDAR data points thinned by the methods outlined here is the best course of action.

Though the application of these methods to Castle Windy was successful, the results achieved from Ross Hammock were disappointing. The differences between the DEM derived from merging the thinned and ground classified points and the DEM interpolated solely from the ground points, as shown in Figure 13 in the previous chapter, are minimal. Though the heights near the apex of the mound near the center of the site appear to be altered slightly, the changes to the topographic expression of the archaeological features are not nearly as dramatic as those witnessed at Castle Windy.

Furthermore, the management imperative for the Ross Hammock site is different than that of Castle Windy, simply due to the respective sites’ locations relative to the shoreline. Since the burial mounds at Ross Hammock lay hundreds of feet away from the shore of Mosquito Lagoon, as opposed to directly adjacent in the case of the Castle Windy midden, erosion is not as much of a factor. However, though still not as dire as at Castle Windy, erosion could eventually threaten the shell midden ridge that runs along the northern and eastern edges of the site boundary. Nevertheless, there are no appreciable differences between the LiDAR signatures of this ridge in the two DEMs in Figure 13.
Though the methods for supplementing the ground classified points with the manually thinned points may be of use in a different application, they did not appear to make much of a difference at either Castle Windy or Ross Hammock. Overall, actually, the DEM derived from just the ground classified points received from NOAA for Ross Hammock appears to represent the site well, as shown in Figure 15 on the next page, and may not even need any sort of manual processing before interpolation.

The raw ground data points for Castle Windy, on the other hand, are clearly in need of some degree of clean up. This is likely due to a relative lack of understory vegetation present at Ross Hammock, which was observed during fieldwork at the site. Less understory vegetation would allow for better returns from the ground surface during LiDAR scans, and thus provide a DEM that more closely approximates the topographic features of a given landscape.

Yet, in both Castle Windy and Ross Hammock, the multipoint layers for these sites can be used as guides for the targeted collection of additional total station RTK data points. The multipoint comparison of Castle Windy in Figure 10 reveals several areas near the apex of the midden that could benefit from such supplemental data points. As previous studies have shown (Pluckhahn and Thompson 2012), additional total station data points are never a bad thing; though the extra time and resources that must be devoted to this endeavor can be restrictive.
Figure 15 DEM of entire Ross Hammock site derived from "ground" classified points.
While the benefits of accurate DEMs for the preservation and analysis of archaeological sites are evident in this study, as well as in the case studies presented in chapter 3, this form of documentation can also greatly aid park managers in complying with the cultural resource management laws mentioned in chapter 4. The Historic American Landscapes Survey (HALS) for instance, specifies current archaeological site features of a landscape that should be documented are spatial organization, visual relationships, natural systems, topography, vegetation, water/hydrology, and buildings/structures (Robinson et al. 2005).

All of these concerns can be addressed through the proper processing and application of LiDAR data. For example, further work at CANA could utilize DEMs derived from the methods specified here to conduct viewshed analyses (Bewley et al. 2005). Such investigations could explore the spatial relationships and sightlines between the archaeological sites present at CANA, thereby studying these features within the context of the surrounding landscape.

National park managers have an obligation to develop management plans for cultural resources within their jurisdiction. The guidelines for HALS stress the utility that GIS and remote sensing methods such as LiDAR can provide in this effort (Robinson et al. 2005). Likewise, the Cultural Resource Geographic Information Systems (CRGIS) initiative, under which HALS is grouped with other cultural resource laws, explicitly encourages the use of GIS technologies by national park managers. Furthermore, the ever widening availability of publicly distributed LiDAR data is overcoming the monetary hurdle associated with this quality of data for many resource-strapped researchers and public land managers.
However, one must also be wary of the pitfalls associated with using publicly
distributed LiDAR data, as has been noted regarding Seminole Rest in this instance. Yet,
as long as one can familiarize themselves with various methods of processing and
analysis of LiDAR data, these difficulties can often be overcome. This thesis suggests
such methods, though adjustments to these processes can and should be made in order
to accommodate the unique topographic traits associated with the particular
archaeological feature one hopes to document. As is particularly evident in the case of
Castle Windy, this thesis reveals just what sort benefits the proper application of these
methods can provide to those responsible for the continued preservation of Florida’s
cultural heritage.
References Cited

American Society for Photogrammetry and Remote Sensing (ASPRS)  

Ashley, Keith H.  

Austin, Robert J., Janice R. Ballo  

Austin, Robert J., Sylvia M. Layman  
1989 *An Archaeological Site Inventory and Management Plan for Volusia County, Florida*. Piper Archaeological Research, INC. Submitted to The Volusia County Planning and Zoning Department, Survey No. 2205.

Axelsson, P.  

Bewley, R.H., S.P. Crutchley, C. Shell  

Brown, A.G.  
2008 Geoarchaeology, the four dimensional (4D) fluvial matrix and climatic causality. *Geomorphology* 101:278-297.

Bullen, Ripley P.  
Bullen, Ripley P., Frederick W. Sleight

Bullen, Ripley P., Adelaide K. Bullen

Chase, Arlen F., Diane Z. Chase


Collins, Lori, Steven Fernandez, and Travis Doering

ESRI

Ehrenhard, John E.
1976  Canaveral National Seashore: Assessment of Archaeological and Historic Resources. Submitted to Florida Division of Archives, History and Records Management, Tallahassee.
Goggin, John M.

Griffin, James B.

Griffin, John W.

Griffin, John W., Dale Benton
1980 Archaeological Survey of South Coastal Volusia County 201 Facilities Addendum. Submitted to Florida Division of Archives, History and Records Management. Tallahassee.

Hardin, Kenneth W., Harry M. Piper

Healy, Paul F., J.D.H. Lambert, J.T. Arnason, R.J. Hebda


The Jaeger Company

Johnson, Robert E., Dana Ste. Claire
Jones, David M. (editor)  

King, Thomas F.  

Lasaponara, R., R. Coluzzi, F.T. Gizzi, N. Masini  
2010 *On the LiDAR contribution for the archaeological and geomorphological study of a deserted medieval village in Southern Italy.* *Journal Geophysics Engineering* 7:155-163.

Lasaponara, R., R. Coluzzi, N. Masini.  

Masini, Nicola, Rosa Coluzzi, Rosa Lasaponara.  

Milanich, Jerald T., Charles H. Fairbanks  

Moore, Clarence B.  


Nelson, Nels C.  

O’Neill, Christopher M.  

Philip, G.M., D.F. Watson  
Piatek, Bruce J.

Pluckhahn, Thomas J., Victor D. Thompson

Robinson, Judith H., Noel D. Vernon, Catherine C. Lavoie

Rolland, Vicki L., Paulette Bond
2000 The Search for Spiculate Clays near Aboriginal Sites in the Lower St. Johns River Region, Florida. Paper presented at the 57th annual meeting of the Southeastern Archaeological Conference, Macon, GA.

Rouse, Irving
1951 *A Survey of Indian River Archaeology, Florida*. In Yale University Publications in Anthropology No. 44. Yale University Press, New Haven.

Russo, Michael, Janice R. Ballo, Robert J. Austin, Lee Newsom, Sylvia Scudder and Vicki Rowland


Russo, Michael

Russo, Michael, Ann S. Cordell, Donna L. Ruhl

Schwadron, Margo
Sears, W.H.  

Sithole, G., G. Vosselman  

Taylor, Bobby J.  

Terrasolid Ltd.  

Young, J.  