Mapping the Landscape for Archaeological Detection, Preservation, and Interpretation: A Case Study in High Resolution Location Modeling from the Blue Mountains of Northeastern Oregon

Trent Skinner
University of Nevada, Las Vegas, trentskinner@gmail.com

Follow this and additional works at: https://digitalscholarship.unlv.edu/thesesdissertations

Part of the Archaeological Anthropology Commons

Repository Citation
https://digitalscholarship.unlv.edu/thesesdissertations/3170

This Thesis is brought to you for free and open access by Digital Scholarship@UNLV. It has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.
MAPPING THE LANDSCAPE FOR ARCHAEOLOGICAL DETECTION, PRESERVATION, AND INTERPRETATION: A CASE STUDY IN HIGH RESOLUTION LOCATION MODELING FROM THE BLUE MOUNTAINS OF NORTHEASTERN OREGON

By

Trent M. Skinner

Bachelor of Arts in Anthropology
University of Nevada, Las Vegas
2014

A thesis submitted in partial fulfillment of the requirements for the

Masters of Arts - Anthropology

Department of Anthropology
College of Liberal Arts
The Graduate College

University of Nevada, Las Vegas
December 2017
This thesis prepared by

Trent M. Skinner

entitled

Mapping the Landscape for Archaeological Detection, Preservation, and Interpretation: A Case Study in High Resolution Location Modeling from the Blue Mountains of Northeastern Oregon

is approved in partial fulfillment of the requirements for the degree of

Masters of Arts - Anthropology
Department of Anthropology

Alan H. Simmons, Ph.D.
Examination Committee Chair

Kathryn Hausbeck Korgan, Ph.D.
Graduate College Interim Dean

Karen Harry, Ph.D.
Examination Committee Member

Liam Frink Ph.D.
Examination Committee Member

Haroon Stephen Ph.D.
Graduate College Faculty Representative
Abstract

Archaeological location modeling (ALM) is an important tool in most survey strategies, and has contributed substantially to economizing efforts to locate and characterize the archaeological record. The increasing availability of high resolution (<3m) airborne light detection and ranging (lidar) data has the potential to refine the application and ultimately the role of ALM. This research tests the precision and accuracy gained by incorporating lidar derived data into an ALM. The site records and other environmental data used in this study were all generated over the last four decades by the resource specialists of the Malheur National Forest. The Weights-of-Evidence (WofE) probability method (Bonham-Carter 1994) was used to produce two ALMs; one based on a 10m digital elevation model (DEM) created from satellite imaging, and the second from a 3m resolution lidar derived DEM. Independent variables (e.g., slope, aspect, distance to water, etc.) commonly used in ALM were largely replaced by index variables (e.g., slope position classification, topographic wetness index, etc.). The final models were classified into areas of high, medium, and low archaeological potential, then cross-validated against a reserved random dataset. Models were then compared using the Kvamme gain statistic and site to area frequency ratio. The 3m model demonstrated a significant improvement over the results obtained from the 10m model and the current probability model used in the study area. A number of factors including model resolution, statistical methodology, and the character of the independent and dependent variables all contributed to the increase in precision and accuracy. The incremental improvement in modeling efficiency demonstrated here will create time and cost saving in the management and preservation of cultural resources, and ultimately contribute to a better understanding of patterns of past human land use.
Acknowledgments

This thesis would not have been possible without the body of work generated by the many archaeological technicians and archaeologists, who spent countless hours over the last four decades surveying and documenting the materials that served as the basis of this work. One in particular, Diane Browning, stands out above the rest.

My thanks to Robert Dickenson, Don Hann, Barbara Millersohr, Katee Withee, and Pete Cadena, the archaeological staff of the Malheur National Forest. Don Hann’s knowledge, ideas, and drive were the impetus for this thesis. Robert Dickenson’s diligent mentoring and instruction in all things GIS provided me with the skills necessary to complete this work. He also kept me on course and motivated to obtain my M.A. degree.

Finally, my sincere thanks to all the members of my committee. I was drawn to UNLV by the work of Alan H. Simmons and Karen Harry, and later had the good fortune for both to serve on my committee. I am deeply indebted to Alan and Karen for campaigning on my behalf as I tried to balance my academic and professional life in order to complete this work. Alan was a constant inspiration, mentor, and supporter throughout my years at UNLV and during summers in Cyprus.
Dedication

I would like to dedicate this work to the women in my life; to my mother, Shirley, who has always served as my anchor; Cristina, my companion, who serves as my rudder, always keeping me on course; and, to my daughter and granddaughters Ashley, Khloe, and Elly, the wind in my sails.
# Table of Contents

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

Acknowledgements ........................................................................................................................ iv

Dedication ........................................................................................................................................v

Table of Contents ........................................................................................................................... vi

List of Tables .................................................................................................................................. vii

List of Figures .................................................................................................................................. viii

1. Introduction ..................................................................................................................................1

   Problem Statement .......................................................................................................................1

   Purpose and Objective ..................................................................................................................2

   Research Question .......................................................................................................................2

   Thesis Structure ............................................................................................................................3

2. Archaeological Location Modeling (ALM) ................................................................................4

   History ..........................................................................................................................................4

   Methods and Theory .....................................................................................................................5

   Issues, concerns, and criticisms of ALM ....................................................................................7

   The Multiple Dimensions of Scale ..............................................................................................8

3. Study Area ..................................................................................................................................10

   Environment ................................................................................................................................12

   Cultural History ............................................................................................................................15

   Archaeology .................................................................................................................................16

4. Materials ....................................................................................................................................20

5. Methods ......................................................................................................................................22

6. Exploratory Analysis ..................................................................................................................32

7. Results ........................................................................................................................................69

   Validation ...................................................................................................................................69

   Model Comparison ......................................................................................................................70

8. Discussion ...................................................................................................................................73

9. Conclusion ...................................................................................................................................79

References ......................................................................................................................................81

Curriculum Vitae ...........................................................................................................................93
**List of Tables**

3.1 Generalized Morphological Typology of Previously Documented Sites .........................18

3.2 Projectile Point Typology ........................................................................................................19

6.1 Group-Wise Summary of Archaeological Sites...........................................................................33

7.1 Results of Current Malheur National Forest Model .................................................................70

7.2 Results of 3m DEM Weights-of-Evidence Model........................................................................70

7.3 Results of 10m DEM Weights-of-Evidence Model.....................................................................70

7.4 Results of 10m DEM Blind Validation Area Weights-of-Evidence Model and ROC Curve
    Validation ......................................................................................................................................71

7.5 Frequency Ratio of Archaeological Sites ................................................................................72
List of Figures

3.1 Regional Context of Study Area ..............................................................................................10
5.1 The Modeling Cycle ................................................................................................................22
5.2 Shannon Diversity Index Example ..........................................................................................28
6.1 Andrefsky’s Generalized Morphological Typology ............................................................32
6.2 Parallel Box Plot of Grouping Analysis...................................................................................34
6.3 Archaeological Location Modeling with a Landscape Ecology Framework ..........................36
6.4 Slope in Degrees and Area Units for 1m, 3m, 10m, and 30m Digital Elevation Models (DEM) ...............................................................................................................................37
6.5 Geology Contrast Values .......................................................................................................38
6.6 Geology Thematic Map .........................................................................................................38
6.7 SAGA Wetness Index Thematic Map, 3m DEM .....................................................................41
6.8 SAGA Wetness Index Contrast Values, 3m DEM .................................................................42
6.9 SAGA Wetness Index Contrast Values, 10m DEM ...............................................................42
6.10 Slope Position Classification, 10m DEM ............................................................................43
6.11 Topographic Position Index (TPI) Values at Three Difference Scales (After Jenness 2007) ........................................................................................................................................44
6.12 Slope Position Classification Contrast Values, 10m DEM ....................................................45
6.13 Slope Position Classification Contrast Values, 3m DEM ......................................................45
6.14 Terrain Ruggedness Index Thematic Map, 3m DEM ................................................................46
6.15 Terrain Ruggedness Index Contrast Values, 3m DEM ........................................................47
6.16 Terrain Ruggedness Index Contrast Values, 10m DEM .......................................................47
6.17 Multiresolution Index of Valley Bottom Flatness Thematic Map, 10m DEM .........................48
6.18 Multiresolution Index of Valley Bottom Flatness Contrast Values, 10m DEM .....................49
6.19 Multiresolution Index of Valley Bottom Flatness Contrast Values, 3m DEM .........................49
6.20 Area Solar Radiation Thematic Map, 3m DEM .....................................................................50
6.21 Area Solar Radiation Contrast Values, 10m DEM ................................................................51
6.22 Area Solar Radiation Contrast Values, 3m DEM ..................................................................51
6.23 Soil Depth, Moisture and Temperature, and Particle Size Thematic Maps ..........................53
6.24 Soil Order Contrast Values ....................................................................................................54
6.25 Soil Order Thematic Map ......................................................................................................54
1. Introduction

Archaeological location modeling (ALM) is an important tool in most survey strategies, and has contributed substantially to economizing efforts to locate and characterize the archaeological record. As survey coverage increases, however, cultural resource management (CRM) efforts begin to shift from an exploratory orientation to one emphasizing management and preservation of identified resources. This preservation focus requires a refinement in the role and application of APM, and the availability of high resolution lidar data has the potential to make this possible.

Problem Statement

The majority of archaeological survey strategies are driven by some form of a probability modeling. The viability of an archaeological predictive model (APM) or ALM is predicated on the resolution and accuracy of the spatial data it derives from. The principal dataset in this equation is the digital elevation model (DEM). Until relatively recently, the majority of DEMs were derived from satellite data, and rendered in a 20~30 meter resolution. Many conventional APM are calculated from a 30m DEM at a 1:25,000 scale; at this resolution acceptable horizontal spatial errors can range from 100~250m (Jaroslaw 2009:2097, 2099).

The resolution of a DEM is critical to APM because the majority of the environmental attributes used in a model are derived from it (e.g., elevation, slope, aspect). Resolution also limits the variety, type, and information potential of environmental attributes. The outcome is often a coarse grained probability map that may aid in reducing the area of physical survey required, but only to a marginal degree. Current mapping standards present a vague and opaque picture of the landscape that stifles anthropological interpretation.
Purpose and Objective

This thesis tests the relative improvement gained in basing an ALM on lidar derived data. Lidar derived data has an accuracy and precision (sub meter) that represents the landscape at the scale of human behavior and decision making. This is paramount, because the task of archaeologists is not simply to identify elements of the modern landscape with the potential to bear cultural material; the ultimate objective of ALM is to understand why people made the decision to habitually occupy certain areas of the landscape (or not), and what factors held weight in the process (Carleton et al 2012:3372).

Research Question

This thesis will answer and discuss the following methodological research question–

1.) Will the use of lidar derived data increase the accuracy and precision of a Weights-of-Evidences predictive model applied to the study area?

*Null Hypothesis (H₀):* 3m lidar DEM does not demonstrate a substantial increase (AUC increase >0.1 / Kvatmme Gain >0.1) over 10m DEM in model efficiency in all categories (Significance of Kvatmme Gain and ROC curve analysis comparison)

*Alternative Hypothesis (H₁):* 3m lidar DEM demonstrates a substantial increase or 10m DEM in model efficiency in all categories (Significance of ROC curve analysis comparison)

Additional underlying queries related to research question one-

a. Which environmental variable(s) of the model present the strongest correlation with location choices? Is there an anthropological grounded causal explanation for the identified correlation?
b. How did variables arrived at solely from lidar data (e.g., landform delineation, flow) compare with variables of the same class calculated using a 10m DEM?

c. Is the resolution of the dependent variable (lithic material) decreasing model precision and accuracy?

**Thesis Structure**

This thesis begins with an overview of ALM including its history, theoretical approaches, and common methodologies. The ALM section ends with a discussion of the influence scale and resolution have on the precision and accuracy of spatial models. This includes a review of other studies that have directly addressed this issue. A description of the study area is presented in the next section. The headwaters of the Middle Fork John Day River, located in the Blue Mountains of northeastern Oregon, serves as the study area for this thesis. The description of the study area includes background on the environment, ethnographically relevant use of the area, and an overview of the region’s archaeology. Details of the 83 previously recorded site and isolates located in the study area are provided. The materials and methods are presented next. This includes a detail description of the Weights-of-Evidence method used to create both the 3m and 10m DEM models. Exploratory analysis details the significance of each variable and the variation seen in the 3m and 10m DEM models. The results section presents the information from the validation and comparison tests of each model. The paper concludes with a discussion of the significance of these results, and their relationship to the methods and theories used in ALM.
2. Archaeological Location Modeling

Archaeological location modeling (ALM—also commonly referred to as archaeological predictive modeling—APM) is defined as a “simplified set of testable hypotheses, based either on behavioral assumptions or on empirical correlations, which at minimum attempts to predict the loci of past human activities resulting in the deposition of artifacts or alteration of the landscape (Kohler 1988:33).”

The principal value of ALM lies is its ability to test hypotheses regarding site location, settlement patterns, and human land use (Neal 2007: 4). Within CRM, APM aids in survey strategies, and allows decision makers to determine whether an area with a high archaeological potential can simply be avoided all together.

History

Early instances of predictive modeling can be found in the habitat and settlement distribution studies of ecology and geography. Gordon R. Willey was the first archaeologist to incorporate the method into his research during the 1950s when he studied early settlement patterns found in the Viru Valley of Peru (Kvamme 2006, Willey 1953). The methodology clearly predates the wide spread use of computers, and early practitioners like Kenneth Kvamme (2006) completed all of the necessary calculation with a map and programmable calculator. A number of major contributions to site location modeling were made in the 1970s by the Southwest Anthropological Research Group, and Ernestene Green, who was the first to apply multivariate statistics to the process in a study of the spatial distribution of Maya sites in the Corozal District of Belize (Verhagen and Whitely 2011, Green 1973).

The use of predictive modeling rose in prominence (in the United States) following passage of the National Historic Preservation Act in 1966, and its subsequent amendments which provided federal agencies with a legislative mandate to identify and inventory cultural resources
located within the lands they manage. With federal land holdings comprising more than 50% of western states, a targeted and systematic approach was necessary to fulfill this mandate. APM was seen as part of the solution, and by the late 1980s, predictive modeling was incorporated into the prefield strategies of most inventory designs. This is what Kvamme describes as the first age of APM (Kvamme 2006). The second age began in the late 1990s with the introduction and widespread use of personal computers, and the introduction of deductive leaning models, largely from Europe and Postprocessual thinking. I would suggest that a third age may be just beginning as part of the geospatial revolution described by Chase et al. (2012).

**Methods and Theory**

"Man may turn which way he please, and undertake anything whatsoever, he will always return to the path nature has once prescribed for him."

-Johann Wolfgang von Goethe 1902

The methods and supporting theory employed in ALM are often described in terms of an objective oriented inductive approach driven by data, or a subjective leaning deductive approach driven by expert knowledge. Every model resides somewhere on a spectrum between the two.

Inductive models were developed out of statistical extrapolation methods, and are often referred to as correlative or data driven in nature (Verhagen and Whitely 2011). They rely heavily on the ecologically based assumption that from a very early period, hominins have been tethered to key resources on the landscape, particularly water, and to a lesser degree tool stone (Foley and Lahr 2015). Other environmental variables have varying degrees of influence on a diverse number of human activities. Inductive ALM models often attempt to describe patterns of persistent and habitual behaviors on the landscape that relate directly or indirectly to those environmental variables. These correlations are then extrapolated into unsurveyed areas by
means of logistic regression. Most models fall within this category, and have been fairly successful over the last four decades. Critics claim this is a form of environmental determinism, is too reductive and dehumanizing, and fails to consider the full range of human behavior. Advocates of the inductive approach like Kvanme (2006) argue that any model is incapable of capturing the full range of human behavior. Humans are bound to the natural environment, but they optimize their navigation and manipulation of it through niche construction. The environment does not determine behavior, but manipulates the patterned occurrence of it across the landscape as prime resources are sought and exploited. This ecological perspective is the most revealing and explanatory approach one can take when discussing human interaction with the environment.

By contrast, deductive approaches to APM are considered more explanatory or theory driven (Verhagen and Whitely 2011). These types of models align more closely with the scientific method, and are usually developed for testing a hypothesis. Simple weighted overlays are often produced in GIS to test hypotheses of settlement location preferences for example (Verhagen and Whitely 2011). This process has shown wide appeal in both new archaeology and Postprocessual camps. However, there are critics of this approach as well. The process is described as being too intuitive in nature and/or overly reliant on expert judgment (Verhagen and Whitely 2011).

Any approach to APM contains a degree of inductive and deductive logic. Verhagen and Whitely (2011) emphasizes that “the “inductive” and “deductive” dichotomy arose in the late 1970s as a historical development and not necessarily as two methodological schools of thought, and as such they should not be thought of as mutually exclusive frameworks.”
Issues, Concerns and Criticisms

During the early years of predictive modeling, methodological, sampling, and statistical issues plagued attempts at spatial analysis, leading to bias, distortion and errors in archaeological research (Schwarz 2006:167). The premise underlying typical aspatial statistical tests (t-test, ANOVA, etc.) is that each variable (observation or sample) is drawn separate and independent of the influences of others of the same type (Schwarz 2006). The autocorrelation found in many forms of spatial data violates this assumption because of Waldo Tobler’s first law of geography (Schwarz 2006):

“Everything is related to everything else, but near things are more related than distant things”

A number of statistical techniques have developed to overcome issues inherent in spatial data. Some form of logistic regression analysis has remained the central pillar of APM since its wide scale adaptation in prefie ld survey strategies during the late 1980s, even though the method has received criticism for violating some of the basic assumptions of spatial data (Carleton 2012, Hatzinikolaou 2006). Weights-of-Evidence and the locally adaptive model of archaeological potential (LAMAP) appear to be two methods with inherent actions for checking undo influences on spatial data calculation and statistical analysis (Carleton 2012).

A final criticism is directed toward certain APMs regarding the nature of what is being modeled. Some misguided APM studies attempt to directly link site locations with past human behavior without first unraveling the tangle of formation processes affecting the archaeological record. Some researchers suggest environmental variables based on the presumed preferences of past societies are anecdotal at best; any association found between sites and present environmental variables is entirely coincidental (Ebert 2006:138). Preservation bias is an important topic, particularly when dealing with surface assemblages.
The Multiple Dimensions of Scale

Three conceptual distinct facets of scale relate to the archaeological record and this research. Each has unique spatial and temporal properties. Phenomenological scale describes the level at which events and processes transpired in the past. Phenomenological scales can vary widely, from the diffusion of ceramic types across a region, to the annual cycle of agricultural practices used in an area, or the chaîne opératoire of a stone tool. Analytical scale describes the scale of investigation, and is at the core of the research question addressed in this thesis. Finally, the level at which meaningful data can be derive from data recorded at the analytical scale is described as the effective scale, or scale of interpretation (Riris 2014).

Airborne lidar provides a dataset capable of informing on each of these facets of scale. Lidar is an indispensable tool for visually locating, identifying, and documenting cultural features in a range of environmental settings. However, its utility as a visualization tool is only the beginning of the technology’s potential contribution to archaeological research. The spatial resolution (~1m) and accuracy (~10cm) of lidar derived DEM make them particularly well suited for ALM.

Lidar derived data has an accuracy and precision (sub meter) that represents the landscape at the scale of human behavior and decision making. This is fundamental to the scale and resolution used to create a model, and its ultimate precision and accuracy. Accuracy describes the ability of a model to reliable predict measurements within a target range. Precision is independent of accuracy, and reflects the proximity of those measurements to each other. In order to ensure all archaeological material is accounted for, accuracy is paramount in CRM applications of ALM. For this reason areas identified as ‘high potential’ in many predictive models encompass major portions of the landscape under study. The object of this thesis is to increase precision without sacrificing accuracy.
Intuition would suggest that the outcome of a predictive model based on a lidar data would provide greater precision and accuracy than one developed using a 10m or 30m DEM. However, a number of recent studies, while presenting substantial improvements in medium and high probability areas, showed mixed results in areas of low archaeological potential; greater precision-less accuracy (Padanyi-Gulyas 2012: 704). The type, resolution and source of independent variables (“evidential themes” in WofE terminology) were cited as the primary cause of the disparity.

Another study with mixed results integrating lidar data into ALM was completed by Verhagen et al (2012) in Netherlands. In this case an object based landform delineation algorithm was used to parse and categorize the landscape. These landform classifications were then incorporated as an independent variable in their ALM.

Another possibility is that inductive modeling methods like logistic regression and Weights-of-evidence (WofE) have a threshold beyond which only diminishing returns are achieved. Incorporating deductive variables, that are developed from firmly grounded theoretically reasoning, is one possible solution to the marginal results of the models discussed so far. A “mobility-shed” (Bruggencate et al. 2015) raster map is one possible deductive variable that may substantially improve a model’s predictive power.

Modeling approaches in ecology discovered that the combination of lidar, aerial photography, and field survey provided a better result. A recent habitat suitability found that each of the data sources complemented the others; each contributing different strengths (Bae 2014:6486).
3. Study Area

Environment

The study area encompasses the headwaters of the upper Middle Fork John Day River, Oregon, which are located in the eastern extent of the John Day Sub-basin, within the Blue Mountain physiographic province of the Mid-Columbia Basin. The province covers much of northeastern Oregon and is bounded in the southwest by the High Lava Plains province, to the northwest by the Deschutes Umatilla Plateau, and in the southeast by the Owyhee Uplands. The northeastern extent of the Blue Mountain province is bordered by the Salmon River Range in Idaho (Orr and Orr 2012: 9, 18-19).

The Blue Mountains are comprised of several relatively small ranges of varying orientation, relief, and geologic origin. The upper elevation of each range is generally composed of older Paleozoic and Mesozoic sedimentary rocks, while the lower extents are draped in
successive lava flows of the middle Cenozoic. The mineral rich Greenhorn Range borders the northern portion of the study area. Time and the elements have sculpted this patchwork of geology into a heterogeneous landscape of narrow drainage systems feeding into confined montane valleys, and slender alluvial plains (Orr and Orr 2012: 20-47, Klinger et al. 2010). The geology and landforms of the study area are discussed in greater detail in Section 6- Exploratory Analysis.

A number of major rivers originate within the Blue Mountains close to the study area and empty into the Columbia River to the north, or its chief tributary, the Snake River to the east. The John Day River is the largest of these draining more than 20,000 square kilometers of northeastern Oregon through its 457 kilometers course from the Blue Mountains to the Columbia River (Klinger et al. 2010, Orr and Orr 2012: 20). The Middle Fork John Day travels from the study area for approximately 120 kilometers before joining the main steam. Vegetation communities within the study area are discussed in detail under Section 6- Exploratory Analysis. In general, the study area is dominated by mixed conifer communities of Ponderosa pine, Douglas fir, and Western Larch. Although still present to a degree, extensive wet meadow complexes once thrived along the headwaters of the Middle Fork John Day River (Klinger et al. 2010: 1). A number of floral species traditionally used by Native Americans are present within the study area, including camas (*Camassia quamash*), chokecherries (*Prunus spp.*), yarrow (*Achillea millefolium*), lupine (*Lupinus spp.*), snowberry (*Symphoricarpos albus*), bitterroot (*Lewisia rediviva*), bisquit root (*Lomatium spp.*), yampah (*Perideridia spp.*), wild strawberry (*Fragaria spp.*), Cascade Oregon Grape (*Berberis nervosa*), currant (*Ribes spp.*), wild rose (*Rosa spp.*), wild onion (*Allium spp.*), big sagebrush (*Artemesia tridentata*), rabbit brush (*Chrysothamnnus spp.*), mountain mahogany (*Cercocarpus spp.*), juniper (*Juniperus spp.*), willow
(Salix spp.), and red-osier dogwood (Comus stolinifera). Ethnographic sources document the use of more than 30 wild plant species within the Harney Basin (Couture et al. 1986: 150).

The study area is home to a variety of mammals, birds, and fish species significant to PreContact populations including elk (Cervus Canadensis), mule deer (Odocoileus hemionus), historically bighorn sheep (Ovis Canadensis), historically bison (Bison bison), coyote (Canis latrans), historically gray wolf (Canis lupus), lynx (Lynx Canadensis) cougar (Puma concolor), beaver (Castor Canadensis), black bear (Ursus americanus), historically grizzly bears (Ursus arctos), several Leporidae and Rodentia species, and a variety of mustelidae such as mink, weasel, badger, skunk, and wolverine. Fish species include Chinnook, steelhead, and other members of the Salmonidae family. Fresh water mussels and eel are also found in the area. A small number of reptile and amphibians are present along with over fifty species of bird.

The climate of the Blue Mountains is generally characterized by cold wet winters and warm dry summers. Elevation in the study area ranges from 1,290m to 1,980m above sea level. The variable relief and diversity of the landscape often leads to localized climate affects (Johnson and Clausnitzer 1992: 3). Significant climatic periods are fairly similar for the Great Basin and Southern Plateau. During the Last Glacial Maximum (24 to 20 KYA) snow fields and glaciers covered upper tributaries of Columbia River and Blue Mountains. Lower elevations largely contained tundra vegetation (Aikens et al. 2011: 152). Between 15,000 to 9,300 cal BP (Terminal Pleistocene Early Holocene (TP/EH)) the Great Basin was covered by roughly 150 pluvial lakes, Harney Basin just to the southwest of the study area contained several lakes supported by the Silveis River which drains the Blue Mountains to the west. Between 8,500 and 7,500 cal BP, pollen evidence suggests conditions on the Plateau were warmer and dryer than those of the present day. Descending timberlines suggest this trend likely reverse in the millennium to follow as moisture increased and temperatures were generally cooler (Aikens et al.
Around 6,000 cal BP, mountain glaciers were reborn in some areas, indicating a further increase in effective moisture. Pollen evidence also suggests that forests continued to move down slope, and replace steppe/grassland environments (Chatters 1998). Between 5,000 to 3,000 cal BP, this trend continued as forest continued to replace grassland steppe at lower elevations. The climate cooled further, while remaining wet, which possibly created the optimal condition for salmon populations of the Plateau. A warming trend occurred between 3,000 to 2,000 cal BP, causing forests to retreat upslope to some extent. This environmental balance has persisted in the region to the present day (Aikens et al. 2011: 152).

**Cultural History**

Climatic variability and cultural history are considered to be closely linked within the Northern Great Basin and Columbia Plateau. However, Plateau environments (and cultures) have at various times resembled and diverged from those of the northern Great Basin over the last 13,000 years (Chatters 2012: 142). Reconstructions of past lifeways in both regions are generally informed by paleoenvironmental data and the broader patterns suggested by it, so the same chronology is employed here (Smith and Morgan 2015: 59). The chronology outlined below corresponds closely with that of the historic environment, and relies heavily on the works of Aikens et al. (2011), Andrefsky (2004), Chatters (2012), and Smith and Morgan (2015). The Great Basin has received the most attention from archaeologist working in the state of Oregon. Because of this specific cultural periods in the area are relatively well defined (Aikens et al. 2011). These cultural traditions are referenced in this chronology whenever possible.
15,000 to 9,300 cal BP: Terminal Pleistocene Early Holocene (TP/EH)

By about 14,500 cal BP, people were present in Oregon. The area’s population grew after 11,000 cal BP, as post glacial warming continued, and summer temperatures increased. Subsistence strategies during this period are characterized by high and wide ranging mobility which likely served as an adaptation to lower precipitation levels and reduced plant diversity (Aikens et al. 2001: 78). The earliest artifacts of a defined age are Clovis fluted points. Surface finds have been recovered in higher elevations of Oregon away from the Columbia and The Dalles (Aikens et al. 2001: 155). A number of other Clovis sites located are located in southern Washington immediately north of the Columbia River. Include among these is a large excavated assemblage at Wenatchee. The most notable Clovis site in Oregon’s Northern Great Basin is the Dietz site located approximately 150km southwest of the study area. Other Clovis material has been recovered from Oregon’s western valleys and coastal region, however these are mostly surface finds, and lack context for precise dating. Clovis examples of this widespread cultural tradition from outside the region have been correlated with c14 dates between 13,200 and 12,800 cal BP. This chronology roughly corresponds with the Paisley Period (>15,700 to 12,900 BP) and Fort Rock Period (12,900 to 9,000 BP) in the Great Basin (Aikens et al. 2001: 155).

9,300 to 5,200 cal BP: The Middle Holocene (MH)

The transition to the Lunette Lake Period (9,000 to 6,000 BP) in the Great Basin is comparable to the shift from the Windust to Cascade phase of the Columbia Plateau (Aikens et al. 2001: 79). Projectile points hold to the leaf shaped form, and assemblages commonly include the remains of small game and waterfowl. Well-formed scrapers are replaced by the infrequent utilized flakes. Hearths are of a shallow and simple form. The record of this period is large
composed of open air sites. The period was relatively dry, and pluvial lakes retreated. Locally produced resources were exploited, and then populations would move on, maintaining a high mobility rate in this period of low moisture and decreased biodiversity. Material recovered at higher elevations in forested landscape suggests deer and elk hunting were an important resource Plateau. New technologies emerged. The Northern Side-notch project point first appears following the eruption of Mt. Mazama in roughly 7,600 BP (Aikens et al. 2001: 79-80)

*Development of Riverine Adaptations during Arid Conditions (9,000-5,000 cal BP)*

The climate cooled further, while remaining wet, which possibly created the optimal condition for salmon populations of the Plateau. The greatest aboriginal fisheries found anywhere in North America developed along the Columbia River during this period (Aikens et al. 2001: 149). Annual subsistence rounds would incorporate both riverine and woodland resources.

*5,200 cal BP to Present: The Late Holocene (LH)*

The Late Holocene period captures the Bergen (6,000 to 3,000 BP) and Boulder Village Periods (3,000 BP to contact) in the Great Basin. Temperatures moderate, and precipitation increases during the Bergen Period then oscillating weather patterns lead to periods of increase precipitation followed by periods of drought during the Boulder Village Period (Aikens et al. 2001: 109). Human populations were relatively high, but varied with the climate. Pithouse village begin to appear during the Bergen period (Aikens et al. 2001: 108). A seasonal pattern of subsistence rounds developed which persists into the Contact period.
Archaeology of the Study Area

Since 1980, 184km² (or 58%) of the study area has been surveyed for cultural resources. These surveys relate to federal compliance requirements for 84 separate undertakings (e.g., timber harvests, range allotments, restoration, etc.) ranging in size from 0.004km² to more than 47km². According to the present Malheur National Forest probability model previous survey covered 63km² of area classified as high probability, 72km² of medium, and 49 km² described as low. These surveys documented all of the sites and isolates used in this study. Three test excavation reports of note were produced in or relatively near the study area. These reports relate to test excavations at 11 different sites, the majority identified early on as the sparse lithic scatters that are the focus of this study.

Forest Highway Road Realignment Project-A road realignment project in the mid-1990s required test excavations be completed at eight previously documented prehistoric sites, roughly eight miles south of the study area. The work was performed between May and August of 1995, by Bill R. Roulette, Julia Wilt, Eric Forgeng, and John Fagan of Archaeological Investigations Northwest. Test excavations were completed at 35GR724, 725, 1680, 1681, 1729, 1730, 1731, and 1732 (Fagan et al. 1996). A variety of lithic and faunal material was documented, along with one stone feature. Two additional sites were tested during this period along highway 26 within the study area and resulted in similar findings.

Established projectile point chronologies for the area suggest the oldest of these sites, 35GR1680, was first occupied ca. 7,000–5,000 BP. The frequency ratio of other temporally diagnostic projectile points coupled with obsidian hydration testing suggests a substantial increase in the use of the area between 5,000–2,000 BP, followed by considerable decline in use from 2,000 BP to Contact (Fagan et al. 1996: iii).
Obsidian sourcing and stylistic similarities suggest the people that occupied these sites had a stronger association with the culture traditions of the Great Basin those of the southern Columbia Plateau. This would suggest that the study area may have served as a primary travel route between the watersheds of the Middle Fork John Day, Malheur, and Burnt Rivers (Fagan et al. 1996: iii).

*The Cannon Site*—is a Late Archaic pithouse village located several kilometers north of the town of Long Creek, near the confluence of Long Creek and one of its tributaries. The site contains thirteen surface depressions ranging from 3 to 7.5m in diameter. In 1985, test excavations were completed in three of the depressions, confirming their characterization as pithouses. Projectile points recovered from the depressions were triangular, with base and corner notching resembling Late Archaic forms of the Columbia Plateau. Underlying the pithouse component in the test units was a lithic assemblage with a projectile point dating to the Middle Archaic Period. These projectile points and other material recovered during testing suggest the site was initially occupied around 4,000 BP (Fagan 1996: 14, Jaehnig 1985).

The study area contains 83 previously documented prehistoric sites and isolates (defined as a concentration of less than ten artifacts) composed primarily of chipped stone. There are also some relatively rare occurrences of ground stone. All of these sites are described as lithic scatters. Lithic scatters are generally characterized as having a low artifact density (defined as approximately one artifact per m²), limited spatial extent (~one half acre), and general lack of residential features. Lithic scatters may contain a subsurface component, but often do not. Artifact diversity is generally low (Rieth 2008: 1).

The most common site type, the lithic scatter, reportedly exhibits the lowest correlative value with specific environmental variables. Data from the Prineville BLM database collected up
to 1990, indicates that lithic scatters increase with relative frequency with an increase in distance from the Columbia River. Lithic scatter density also corresponds with elevation, becoming less frequent at lower elevations. Lithic scatters were also found on the widest range of primary and secondary land forms, and associated with a wide range of water sources (Lebow et al. 1990: 127).

**Generalized Morphological Typology**

Andrefsky’s (2006: 73) generalized morphological typology was used to record the chipped stone material documented at each of the sites and isolates found in the study area. This information is discussed in greater detail in Section 6-Exploratory Analysis. The abbreviations OBS, FGV, and CS found in the table below are the material types obsidian, fine grain volcanics (which includes basalt and ignimbrite), and the crystalline silicate family of macro, micro, and crypto rocks (which includes chert/ flint, jasper, quartz, etc.). The line below the material types (line six) records the total count for each type by material. Line eight describes the distribution (in percentage) of the total type count by material. The bottom line describes the proportion of each type within the total assemblage count for the study area.

<table>
<thead>
<tr>
<th>Chipped Stone Tool</th>
<th>Debitage Flakes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biface Hafted</td>
<td>Nonbiface</td>
</tr>
<tr>
<td>Hafted</td>
<td>Unhafted</td>
</tr>
<tr>
<td>OBS</td>
<td>FGV</td>
</tr>
<tr>
<td>59</td>
<td>1</td>
</tr>
<tr>
<td>Total Hafted</td>
<td>65</td>
</tr>
<tr>
<td>90.8</td>
<td>1.5</td>
</tr>
<tr>
<td>% Total = 65%</td>
<td></td>
</tr>
<tr>
<td>21.4</td>
<td>35.7</td>
</tr>
<tr>
<td>% Total = 245%</td>
<td></td>
</tr>
<tr>
<td>% Total = 245%</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.1
Chronological Typology

The chronology below was synthesized for the Malheur National Forest by Don Hann (2010) through examination of Forest’s substantial lithic record, and references to Justice (2002) and Oetting (1999).

<table>
<thead>
<tr>
<th>Projectile Point Typology</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Side-notched (8,000~4,500 BP)</td>
<td>10</td>
<td>23</td>
</tr>
<tr>
<td>Elko Series (4,000~1,000 BP)</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>Rosegate (1,250~650 BP)</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Gatecliff Split Stem (5,000~2,700 BP)</td>
<td>3</td>
<td>6.8</td>
</tr>
<tr>
<td>Large Stemmed (11,200~8,000 BP)</td>
<td>6</td>
<td>14</td>
</tr>
<tr>
<td>Small Stemmed (8,000~1,000 BP)</td>
<td>1</td>
<td>2.3</td>
</tr>
<tr>
<td>Willow Leaf (10,000~1,000 BP)</td>
<td>1</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>44</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 3.2
4. Materials

Software

All Weights of Evidence calculations were completed in Esri ArcMap 10.3 using the open source Arc Spatial Data Modeler 5 (ArcSDM) toolkit. ArcSDM was initially developed by Don Sawatzky et al. (2010) for the United States and Canadian Geologic Surveys. Terrains attribute and index calculations were completed in Esri ArcMap 10.3, Quantum GIS (QGIS) 2.18, the System for Automated Geoscientific Analysis (SAGA), Geographic Resources Analysis Support System 7.2 (GRASS), and The Relief Visualization Toolkit 1.3 (RVT). Calculations for the Slope Position and Landform Classification indices were completed in ArcMap 10.3 using the Topology Toolbox produced by Tom Tilts at the Great Basin Landscape Ecology Lab. With the exception of ArcMap 10.3, each of these programs and extensions are free and open source and available at the following links:

- Spatial Data Modeler 5: https://github.com/gtkfi/ArcSDM
- The Relief Visualization Toolkit (RVT) is available at: http://iaps.zrc-sazu.si/en/rvt#
- QGIS with plugins for the latest stable releases of SAGA and GRASS is available at: http://www.qgis.org/en/site/.
- Topology Toolbox for ArcMap 10.3: http://www.arcgis.com/home/item.html?id=b13b3b40fa3c43d4a23a1a09c5fe96b9

Digital Elevation Model Data

The 3m DEM was derived from lidar data recorded during a 2009 airborne survey of the Middle Fork John Day River (Aero-Graphics 2010). This data is freely available to the public.
under the State of Oregon Lidar Acquisition Prioritization Plan at:

http://www.oregongeology.org/lidar/.

The 10m DEM was obtained freely from the USGS National Map at: https://nationalmap.gov/.

**Additional Environmental Data**

Soil, vegetation, and geologic properties used in this model were obtained by spatial data developed by respective specialist of the United States Forest Service (USFS). This data is freely available to the public at https://www.fs.fed.us/r6/data-library/gis/umatilla/index.shtml.
5. Methods

This thesis tests the relative contribution of lidar derived data towards archaeological location modeling. This four step process will begin with exploratory analysis of two datasets, the first created from a 3m Lidar derived digital elevation model (DEM), and the second dataset created from a 10m DEM produced from orthographically corrected satellite imaging.

Following exploratory analysis, probability models will be constructed from the suite of explanatory variables that present the strongest correlation to previously documented archaeological sites and isolates. Finally, the models will be validated, and then compared for their precision and accuracy. This section begins with an introduction to the theory underlying Weights of Evidence modeling, followed by a description of the procedures employed in this four step modeling process.
Weights-of-Evidence Modeling

Weights-of-Evidence (WofE) is a data driven method based on Bayesian probability (Lee et al 2012:92). Similar to methods of multiple regression in frequentist statistics, WofE describes the relative predictive power of multiple independent variables in relation to a dependent variable.

The WofE method begins by defining the prior probability. First, a set of similar archaeological phenomenon are identified (e.g., agricultural fields, habitation site, burial mounds), \( A \), whose distribution on the landscape is hypothesized to be influenced to some degree by characteristics of the environment (e.g., slope, distance to water, visibility). Next, a study area, \( S \), containing known occurrences of \( A \) is delineated. Each occurrence of \( A \) is assumed to occupy a small unit area or cell of \( S \). The size of the unit cell is defined by the size of \( A \), and the assumption is that there is not more than one occurrence of the archaeological phenomenon in each unit cell. The total study area is then divided by this unit cell size, which is described as \( N \{S\} \), where \( N \{ \} \) is the notation denoting the count of unit cells. The prior probability, \( P (A) \), can then be calculated by dividing the number of known archaeological phenomenon \( N \{A\} \) in the study area, by the total number of unit cells \( N \{S\} \), or \( N \{A\} / N \{S\} \). In a notional example of a defined study area of 10,000 km\(^2\) (\( N \{S\} =10,000 \)), which is known to contain 200 Bronze Age settlements that average 1 km\(^2\) each (\( N \{A\} =200 \)), the prior probability is defined as \( 200/10,000 = 0.02 \). This is the probability, that lacking any additional information, any one unit cell chosen at random will contain a Bronze Age settlement (Bonham-Carter 1994: 304).

The unconditional prior probability can be updated and a conditional posterior probability produced by introducing new evidence, \( E \), related to the distribution of the archaeological phenomenon within the study area. This can be represented as \( P (A|E) \), were \( P \) is
the probability of the presence of \( A \), the archaeological phenomenon, conditioned on the co-occurrence of \( E \), the new evidential theme. If for example 180 out of the 200 Bronze Age settlements are determined to occur in areas averaging less than 5% slope, which represents only 3,600 km\(^2\) of the total study area, then the potential for locating areas favorable to this type of phenomenon increases substantially. In this example, the conditional probability of locating the phenomenon given less than 5% slope is \( 180/3,600 = 0.05 \), or two and a half time greater than the prior probability (0.02). Conversely, the probability greatly decreases if this new evidence is absent. If only 20 settlements \( (20/200 = 0.1) \) occur in the remaining 6,400 km\(^2\) \( (6,400/10,000 = 0.64) \), then \( 0.1 / 0.64 = 0.15625 \). This means that the probability of locating the phenomenon absent the new evidence is 0.15625 times smaller than the prior probability, or \( 0.02 \times 0.15625 = 0.003125 \) (Bonham-Carter 1994: 305-308).

These conditional probabilities of presence/absences are calculated in WofE modeling using the natural logarithm of odds (logit), and expressed as positive and negative weights \((W^+, W^-)\) using the following equations:

\[
W^+ = \ln \frac{P(E|A)}{P(E|\bar{A})} \quad \text{and} \quad W^- = \ln \frac{P(\bar{E}|A)}{P(\bar{E}|\bar{A})}
\]

In these expressions of Bayes’ rule \( P \) represents the probability, \( \ln \) is the natural log; \( E \) and \( \bar{E} \) represent the presents or absence of an evidential theme, and \( A \) and \( \bar{A} \) represent the presence or absence of the archaeological phenomenon (Agterberg 2014: 142).

Contrast \((C)\) measures the difference between the positive and negative weights, \( C = W^+ - W^- \), and serves as a measure of the overall weight or predictive power of an evidential theme (Sawatzky et al. 2010). This measurement is positive for a positive spatial association, and negative for a negative spatial association (Pradhan 2010). A contrast equal to zero, \( C = 0 \), would
indicate no spatial correlation exists between the archaeological phenomenon and the hypothesized evidential theme. Values near 0.5 are generally considered mildly predictive, between 0.5 and 1 moderately predictive; from 1 to 2 highly predictive, and values over 2 reflect an extremely predictive contrast (Sawatzky et al. 2010).

The studentized contrast \( (\text{Stud C}) \) serves as a measure of \( C \)'s significance and the relative certainty of the spatial association. \( \text{Stud C} \) is calculated by dividing the contrast value by its standard deviation. The standard deviation of contrast \( S(C) \) is calculated as:

\[
S(C) = \sqrt{S^2(W^*) + S^2(W^-)}
\]

Where \( S^2(W^*) \) and \( S^2(W^-) \) are the variances of \( W^* \) and \( W^- \) (Pradhan et al. 2010).

In WofE modeling, multiple binary instances like the example presented here are combined together to generate the final probability, or response map. This conditions the probability of occurrence to multiple values, instead of just one: \( P(A | E_1 \cap E_2 \cap E_i \ldots) \)

Where \( P \) is the probability of and archaeological phenomenon, \( A \), occurring given the intersection of multiple evidential sets \( E_1, E_2, E_i \ldots \) This combination includes both values for different explanatory variables (e.g. distance to water, slope, and ruggedness) and different classes within a single variable (e.g. slope classes of North, South, East, and West).

A final measurement in WofE modeling is that of Conditional Independence (\( CI \)). Conditional independence \( (CI) \) is assumed in WofE, and proposes that evidential themes (e.g. slope, aspect, elevation) are not correlated with each other, conditioned on the presence of the archaeological phenomenon. If this assumption holds true, then the summation of the study area multiplied by the posterior probability value for the model, \( T \), should equal the total number of phenomenon \( N \{A\} \) used in its creation (Bonham-Carter 1994: 315-316). However, when combining multiple explanatory variables, models often violate this conditional to some degree.
(Bonham-Carter 1994: 312, Agterberg and Qiuming 2002). Violations are considered significant if the resulting value is greater than 15% of the expected value (Bonham-Carter 1994: 316). To mitigate the effects of any serious violation of CI, evidential themes are either eliminated or combined, and the results cycled again to generate a new response.

Procedures

*Exploratory analysis-*

Exploratory analysis is a vital first step in archaeological probability modeling. During this stage assumptions about the explanatory power and spatial association of both the dependent and individual independent variables are tested.

*Dependent Variable*

Three steps are necessary to prepare dependent variable data for use in model development. First, sites are grouped into several ‘site types’ according to the diversity, evenness, and size of their assemblages. Second, the spatial information for each dependent variable group is converted from polygons to points. Finally, the entire dataset is divided into two randomly sampled populations: a 70% training set which will be used in model construction, and a 30% reserve set used for cross-validation of the final model.

Surface stone artifact assemblages are the focus of this locational analysis model. The assemblages investigated in this study area are commonly referred to as ‘lithic scatters’ and characterized by their relatively sparse nature, primary composition of chipped stone debitage, tools, occasional ground stone, and general lack of archaeological features. These types of assemblages are normally lumped together as ‘prehistoric sites’ in archaeological location modeling, and tested for a simple binary state of presence or absence. This uncritical lumping of
these modest amorphous sites into one dependent variable can lead to an over generalization of a
model’s information value and predictive potential.

Landscape structure has a differential influence on patterns of human behavior related to
lithic procurement, manufacture, use, maintenance, and discard. Palimpsest factors such as
temporal variation in the type and intensity of activities performed at each location and their
resulting discard patterns may cloud or obliterate any direct inferential links to these behaviors.
However, because of the sparse ephemeral nature of many of these deposits, it is likely that some
portion represent relatively discrete chronologies, and are therefore likely to retain some record
of the behavioral pattern that formed them. Previous studies have demonstrated a clear
association between specific elements of the landscape and variation in the structure of
This association is influenced to varying degrees by both behavioral and natural processes.
Archaeological location models that address the potential relationship between assemblage
composition and landscape elements will have greater precision and accuracy, and ultimately
possess greater utility.

The Shannon Diversity Index ($H$) together with values of evenness and assemblage size
are calculated for each site to characterize the variability of their structure. This will be followed
by grouping analysis of these values in Arcmap 10.3 to determine the optimal number of
dependent variable categories for this study (e.g. binomial vs. multinomial).

The Shannon Index ($H$) is a measure of a population’s diversity and evenness. Diversity
relates to richness, and described the variety of artifact types represented in an assemblage.
Evenness articulates the degree to which counts for those artifacts are evenly distributed
throughout the types found in an assemblage. Variations of this metric have been applied in lithic
research to answer a number of questions related to assemblage variability in both regional and intrasite studies (Odell 2004: 111, Hammond et al. 2015, Bocquet-Appel and Tuffrreau 2009, Kvamme 1988b, Prentiss and Clarke 2008). The Shannon Diversity Index \((H)\) is calculated as:

\[
H = \frac{n \log(n) - \sum_{i=1}^{k} f_i \log(f_i)}{n}
\]

Where \(f_i\) is a vector of artifact counts by material type recorded in Andrefsky’s (2006: 75) generalized morphological typology (see example in table below), \(k\) represents the number of classes, and \(n\) is the total count within the assemblage. If \(f_i = 0\), then the index value is set to 0.

The Shannon Diversity Index and Evenness values are calculated in excel. First, a total artifact count for each site is calculated. In step two, the percentage of each artifact type is calculated by dividing the total count in each artifact class by the total sum for the site. Next, the natural logarithm of each percentage value is calculated, and the result multiplied by the initial percentage value. The final values for each type are then summed up across the site. The absolute value of this sum is the \(H\) value for the site. \(H_{max}\) represents the maximum diversity possible.

This value is held constant for all sites. In the example above, a \(H_{max}\) value of 2.71 may only be achieved by having an equal value (>0) in each of the artifact classes and material types of the typology. Evenness is calculated as \(H/H_{max}\).
Grouping analysis is completed in ArcMap 10.3 to determine the optimal number of dependent variable(s) for this study (binomial vs multinomial). The values of Diversity, Evenness, and Total Count are entered in the attribute field for each site in ArcMap 10.3. In ArcMap 10.3, K means analysis is carried out in data space (with no spatial constraints), and ArcMap’s pseudo F -statistic is run to identified the optimum number of groups. Larger F-statistic values indicate grouping optimization by determining the count with the highest within group similarities, and between group differences (ESRI 2017). The resulting groups are individually tested against the independent variables in the next step to identify any variation in the explanatory power related to each. If these differences are significant, a separate model is created for each group. These models are combined in the final probability map, while retaining the information value for each group.

-Independent Variables

Weights-of-Evidence values for all independent variables are calculated in ArcMap 10.3 using the Spatial Data Modeler (ArcSDM) toolkit, following the methods outlined in the beginning of this section (5.1 Weights-of-Evidence Modeling). Weights are calculated using the 70% training set.

An independent variable is considered significant in WofE modeling if it has a positive or negative contrast greater than two ($-^+C > 2$) and a corresponding Studentized Contrast of greater than forty ($\text{Stud } C > 40$). Pairwise testing of CI is conducted before combining explanatory variables in the final model.

Model Construction-

Independent variables demonstrating a significant contrast ($-^+C > 2$ and $\text{Stud } C > 40$) and conditional independence are combined together to create a final probability model (or response
theme). An omnibus test of conditional independence is run. Models with a $CI$ margin of 15% or less pass, and proceed to the validation step. Any model demonstrating a significant violation of $CI$ will return to exploratory analysis of evidential themes and pairwise testing of their $CI$.

Independent variables will be combined or removed until the final response passes the omnibus test of conditional independence.

**Validation**-

The Receiver Operating Characteristic (ROC) curve (AUC) is a common statistic for assessing the discriminatory capacity and performance of classification models (Mas et al. 2013, Valverde 2012). A ROC captures the sensitivity ($Se$) and specificity ($Sp$) of a model. Sensitivity is instances of presence correctly predicted as present (True Positives). Specificity is instances of absence correctly predicted as absence (False Positives). Performance is measured by the area under the ROC curve (AUC). An AUC of 1.0 represents a perfect model; a value of 0.5 is roughly equivalent to chance and suggests a flawed model.

AUCs are calculated for the 70% training set used to create the models, and the 30% sample retained for cross-validation. AUC validate the power of the model to independently predict potential locations of the dependent variable. AUC values are calculated in the ArcSDM toolkit and compared in R using the pROC package.

**Model Comparison**-

A simple, robust, and informative measure of a model’s accuracy and precision is the Kvamme Gain ($KG$) statistic (Verhagen 2007a: 120, Kvamme 1988a: 329). The Kvamme’s Gain is calculated as:

$$ KG = 1 - \frac{P_a}{P_s} $$
Where \( P_a \) equals the percentage of the study area captured in the zone of interest (e.g. high, medium, or low probability); and \( P_s \) equals the percentage of sites (units) located within that same zone (Verhagen 2007a: 120). A model with a \( KG \) in the 0.6 range is considered reliable enough to guide pedestrian survey.

The KG is used to compare probability models produced at different resolutions (3m and 10m DEM). Comparison is also made between the logistic regression model currently in use on the Malheur National Forest, and the WofE model(s) produced here.
6. Exploratory Analysis

Exploratory analysis is a vital first step in archaeological probability modeling. During this stage assumptions about the explanatory power of individual variables are tested. Undertaken properly, exploratory analysis increases the reliability, precision and accuracy of the final model, while eliminating untested, counterproductive, or superfluous data from the process.

Exploratory Analysis of the Dependent Variable

Grouping analysis divided the 83 previously recorded sites and isolates of the study area into three categories based on values of diversity, evenness, and assemblage size. Diversity relates to richness, and described the variety of artifact types represented in an assemblage. Evenness measures how uniformly artifacts are the distributed across types (Odell 2004: 110, Prentiss and Clarke 2008).

To determine diversity and evenness value for each assemblage, the Shannon Index ($H$) was calculated in Microsoft Excel using the values for artifact count and material type recorded for each site in a modified version of Andrefsky’s Generalized Morphological Typology (Andrefsky 2006: 75-see Figure 6.1). This modification was required to standardize the data recorded at each of these sites over the last four decades (e.g. every record documents flake counts and material types, but not other common lithic attributes like size, stage, etc.) The results were entered in the attribute table for each site in Esri ArcMap 10.3. The variables of $H$, evenness, and total artifact count for each concentration were used to conduct grouping analysis.
In the ArcMap’s Grouping Analysis toolkit, K-means analysis was conducted in data space (with no spatial constraints), and ArcMap’s pseudo F-statistic used to identified the optimum number of groups. Larger F-statistic values indicate the optimal within group similarities, and between group differences (ESRI 2017a). These calculations placed the 83 previously recorded sites and isolates into three groups (see Group-Wise Summaries in Table 6.1).

<table>
<thead>
<tr>
<th>Overall Variable Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count = 83, Std. Distance = 179.35, SSD = 58.17</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Evenness</td>
</tr>
<tr>
<td>Shannon Index</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count = 41, Std. Distance 94.13, SSD = 45.97</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Evenness</td>
</tr>
<tr>
<td>Shannon Index</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count = 40, Std. Distance 60.25, SSD = 12.17</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Evenness</td>
</tr>
<tr>
<td>Shannon Index</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Count = 2, Std. Distance 13.50, SSD = 0.02</td>
</tr>
<tr>
<td>Variable</td>
</tr>
<tr>
<td>Count</td>
</tr>
<tr>
<td>Evenness</td>
</tr>
<tr>
<td>Shannon Index</td>
</tr>
</tbody>
</table>

*Table 6.1 Group-Wise Summary of Archaeological Sites*
Group one (n=41) represents the higher spectrum of diversity and evenness in assemblage type. Assemblage size varies from a low of 2 to a high of 471. Group two (n=40) consists of assemblages characterized by low diversity and evenness values. Assemblage size varies from a low of 1 to a maximum of 315.

A third category identified in grouping analysis contains two relatively large and extremely homogenous assemblages. Both are reported to be primary lithic sources for local tool stone; one jasper and the other basalt. These two sites were removed from locational analysis, because their position appears to be entirely dependent on one variable- the local geology. Including these large sites in the model would likely diminish the information potential of other explanatory variables which possibly played a greater role in the choice of location for the majority of the sites.
After accounting for broader mobility strategies and the influence of natural processes on artifact distribution and visibility, the following general inferences can be reached regarding assemblage types.

High diversity concentrations are characterized by a richness in tool and material type. In general, it can be inferred that high diversity concentrations represent persistent places on the landscape; the higher the diversity, evenness, and accumulation values, the greater the number of activities, and possibly the more routine, and longer the duration of visits to a location across time (Clarkson 2008b, Kandel et al. 2015).

Low diversity concentrations have greater homogeny, and often contain only one tool or material type. The inference is that these assemblages represent singular or relatively short duration activities such as game hunting or processing, or opportunistic exploitation of a secondary lithic sources. These types of activities are presumably less constrained by multiple environmental variables, and are therefore likely to have contrast values in WofE (or other regression) modeling different that high diversity artifact assemblages.

From a resource management perspective, low diversity sites are likely to have limited information potential, and therefore unlikely to be determined eligible for inclusion in the National Register of Historic Places (although there are sure to be exceptions). By parsing them from the higher diversity sites the efficiency of the model increases, possibly saving km² of physical survey requirement.
Exploratory Analysis of the Independent Variables

An independent variable is considered significant in WofE modeling if it has a positive or negative Contrast greater than two ($-^{+}C > 2$) and a corresponding Studentized Contrast of greater than forty ($\text{Stud } C \geq 40$). A summary of significant values for each independent variable are provided, along with a discussion of the qualitative aspects of the results. Exploratory analysis also contributes to the central question of this thesis by providing details of how, and to what degree DEM resolution impacts the calculations and results of individual variables used in probability modeling. Figure 6.3 summarizes important elements of the Landscape Ecology framework discussed earlier, and common proxies for their measurement in GIS. Beginning from the bottom with geology, each of these elements informs on the next, and shapes the
structure of the landscape, and in turn human use of it. This framework provides a productive outline for exploratory analysis.

A primary aim of this thesis is to test to relative contribution of lidar derived data in archaeological location modeling. Attributes calculated from a lidar derived Digital Elevation Model (DEM) will be compared to those generated from geospatial referenced and orthographical corrected satellite imagery.

To determine the optimal resolution to use in modeling for each dataset (lidar vs. satellite) a number of basic terrain attributes (e.g. slope, aspect, ruggedness) were calculated at the 1m and 3m resolution for the lidar data set, and the 10m and 30m for the USGS source satellite dataset. The results demonstrated a redundancy in both datasets. From this point forward exploratory analysis and modeling will compare the 3m lidar derived dataset with the 10m dataset derived from USGS satellite imagery.
Geology

Geology is the initial building block of a landscape's morphology, and a significant variable in characterizing the distribution of prehistoric sites within the study area. The geology is typical of the Blue Mountains, with older Triassic and Permian sedimentary rocks forming the upper elevations and successive lava flows of the middle Tertiary forming the lower. This together with the hydrological cycle of the area worked to form a heterogeneous landscape of narrow drainage systems feeding into confined montane valleys, and narrow alluvial plains (Orr and Orr 2012: 20-47, Klinger et al. 2010).

There is a significant correlation between high diversity assemblages and alluvium (C=2.55 / Stud C= 80.35). Alluvium in the study area is of Pleistocene and Holocene age and composed of poorly sorted fluviatile sediments of silt, sand, and gravels (Ferns et al. 1983). Geomorphological analysis placed the alluvial deposits of the Upper Middle Fork John Day River into four categories based on surface characteristics and their elevation relative to the active channel. The first is the flood plain material that is inundated and regularly reworked. The second alluvial category is terrace surfaces located 3-5 feet above the active channel, which
based on a single radiocarbon date were abandoned by the river 1,000-1,200 years ago. These low lying terraces exhibit narrow shallow channels or scars of abandoned channel meander on their surface. The third alluvium category is older terrace deposits located 6~9 feet above the active channel. These older terraces have a noticeably smoother, planar surface. Category three terraces often support stands of older conifers. Many of the stumps found on this surface type were approaching 400 years in age when cut (1920s~1940s), providing a minimal age range of approximately 500 years. The presence of Mazama ash near the lower extent of these landforms suggests a lower age range of ~7,600 BP. Mazama ash in fluvial deposits may not be a reliable indicator of age by itself, however this date range is also supported by detrital charcoal and pollen testing. The fourth and final alluvium category captures the oldest and highest terraces. These terraces are estimated to range from 8,000 to 10,000 years in age based on qualitative assessment of their pedogenic development (Klinger et al. 2010).

To better understand how cultural material is distributed within these four alluvium categories a Height Above the Nearest Drainage (HAND) (Nobre et al. 2011) raster was generated for the study area. The results indicate that of those cultural units located on alluvial deposits, 63% are in Category two areas (3~5 feet above the active stream channel), 22% are located in the Category Three (6~9 feet) and Four areas, and the remaining 13% are located in Category One areas which are immediately next to or in the active stream channel.

Andesitic basalt underlies the majority of the study area (267 km²), and together with granite (2.1 Km²), displays a strong negative relationship with both assemblage types. The contrast is strongest for the total training sample (C=−2.18 / Stud C= 77). The andesitic basalts are Upper Eocene and Oligocene in age (roughly 28~48 kya). This material originated from lava flows erupting over successive period from surface fissure to the east. Other volcanoclastic
material is captured under this heading and includes mudflow breccia deposits formed from lahars along with tuff deposits composed of poorly sorted rock fragments in a matrix of felsic volcanic ash, sand and silt (Ferns et al. 1983).

Basalt and other fine grained volcanic material of the same age as the andesitic basalt group covers approximately 25 km² of the study area and shows a moderate correlation with both assemblage types (High diversity C=1.46/ Stud C = 43.22; low diversity C=1.63/ Stud C = 21.76). This material served as a local tool stone in the area. Other locally sourced tool stone includes a mix of quartz, jasper, chert and other macro, micro, and crypto crystalline silicates. This mix of material originates primarily from the older variegated rock terrane of Triassic and Permian age argillite, sandstone, and ophiolite material which forms the upper elevations of the Greenhorn range immediately north of the study area (Ferns et al. 1983).

Other results of WofE testing suggest low diversity assemblages are correlated to areas of high clay and mixed geology, however the Studentized contrast does not meet the established threshold for significance (C=2.17 / Stud C= 28.23).

Topographic Wetness

Martijn van Leusen (1993) describes water drainage, soil type, and topology as the three most vital parameters in models of prehistoric land use. Ground offering better access to surface water or soil moisture is of almost universal importance to foraging and agrarian economies found in a variety of biogeographical settings (Bevan et al. 2013, Vogel et al. 2015, Kandel et al. 2015).

The Topographic Wetness Index (TWI) quantifies the influence of terrain on hydrologic processes within a watershed. The TWI provides a relative, steady state measure of the upslope
contribution of water to a particular area. In its most parsimonious form, the index is calculated as:

$$\ln \left( \frac{U}{\tan S} \right)$$

Where $U$ represents the upslope contributing area per unit contour length and $\tan S$ is the local slope (Buchanan et al. 2014). The SAGA Wetness Index is a slight variation on this algorithm, which uses a modified catchment area for the calculation and does not view the flow of water as a very thin film. This results in a more realistic model of water movement, particularly in low lying areas with relatively minor changes in elevation (Boehner et al. 2002).

Calculations of the topographic wetness index at the 3m and 10m resolution both show significant results, but with stark variations between models. In the wettest category of the 3m model, 0.75% of the study area (2.4 km²) contains 8.6% of the high diversity assemblage type (383-10m² cultural units) with a contrast of ($C=2.47 / \text{Stud } C=44.86$). The wettest category in the 10m model, also contains the highest positive contrast ($C=2.14 / \text{Stud } C=64.15$), where less than 4.4% (14.1 km²) of the 317 km² study area contains 28% of the high diversity assemblage type (1,254-10m² cultural units).
Conversely, the driest category in the 10m model covers 20% (63.5 km²) of the study area (63.5 km²) and contains only 2.6% of the high diversity assemblages (114-10m² cultural units), giving a high negative contrast but with a lower Studentized Contrast (C =-2.25 / Stud C=-23.76). In the 3m model, the driest category covers 18 km² more acres than the 10m model, with a lower contrast but a slightly higher Studentized contrast (C = -1.87 / Stud C=-27.29). The most significant variation in the two models is seen in the distribution of the study area over the middle three categories were more than 20 km² of the study area are shifted between categories.

There were no significant contrast results for low diversity assemblages. However, the results clearly demonstrate the averaging effect low diversity assemblages have on the total training sample results for this independent variable. These results also suggest that there is some relationship between the environmental affordance captured in the topographic wetness index and the range of activities that contribute to the formation of low and high diversity assemblages.
Landform Classification

Landform classification segments the terrain of an area into discrete relief units based on terrain geomorphometrics. A number of reliable indices have developed to perform landform classification in specific research fields related to hydrology, ecology, and pedology. Most have the potential to inform on environmental affordance in archaeological location modeling. The results of WofE testing for three of the more reliable landform classification indices are presented here.

Landform-Slope Position Classification (SPC)

The Slope Position Classification is a simple and robust measure of discrete relief units developed by Andrew Weiss (2001). The Topographic Position Index (TPI) serves as the basis of the classification method. The TPI is simply the difference between a cell’s elevation value and the average elevation for a predefined neighborhood of surrounding cells (Jenness 2007). The size and shape of the neighborhood is the most important value to consider when calculating the TPI, and can have profound effects on the results of the SPC as demonstrated in Figure 6.10. In
calculating the TPI for the respective DEMs, a circular neighborhood radius of 33 cells was used for the 3m DEM and a circular neighborhood radius of 10 cells for the 10m DEM. This provided each DEM with an approximately 200m wide TPI classification widow. Slope Position Classification is performed by comparing the results of the TPI against slope and elevation for the same area. Landform units are classified based on the TPI values elevation along the slope. Landform categories include Valleys, Toe Slope, Flat, Mid Slope, Upper Slope, and Ridges (Jenness 2007).

The percentage of study area covered by the six classes of SPC varies significantly between the 10m and 3m DEMs. The 10m DEM appears to outperform the 3m DEM in the SPC for predictive potential. The portion of the study area classified as ‘Valley’ consists of 33.1% (105 km²) in the 3m DEM compared to only 16.4% (52 km²) in the 10m DEM. Area classified as ‘Ridges’ covers 39.1% (124 km²) of the 3m DEM, but only 19.5% (62 km²) of the 10m DEM. The WofE contrast value for the 10m DEM in this class crosses the established significance threshold, however the corresponding Studentized contrast does not (C = -2.46 / Stud C=23.23).

The portion classified as ‘Flat’ is the only area to demonstrate significant results for both the 3m and 10m DEMs. In the 10m DEM, 12.3% (39 km²) or the study area is classified as ‘Flat’ and contains 54.4% of the high diversity assemblage type (2,419-10m² cultural units), resulting
in contrast values of \( (C = 2.15 / \text{Stud } C = 71.54) \). In comparison, the portion classified as ‘Flat’ for the 3m DEM contains only 3.8% (12 km²) of the study area and 24% of the high diversity assemblage type (1,084-10m² cultural units), resulting in contrast values of \( (C = 2.08 / \text{Stud } C = 59.2) \).

A significant difference in the distribution of high and low diversity assemblage types is also seen in the ‘Flat’ class. While the class serves as a positive draw for high diversity assemblage types, it appears to repel those of low diversity. Only 1.9% (16-10m² cultural units) of the low diversity assemblage type is located in the ‘Flat’ class of the 3m DEM, and 4.5% (38-10m² cultural units) in the same class of the 10m DEM.

None of the contrast values for the low diversity assemblage type crosses the established significance threshold \( (-^*C \geq 2 / \text{Stud } C \geq 40) \). However they do diminish some of the contrast values (in the Flat and Ridge classes) for high diversity assemblages below the established significance threshold when the two assemblage types are combined.
Rugged terrain can have a profound effect on a number of social, cultural, and economic factors at multiple scales. Terrain variability has played a significant role in shaping exploitation strategies of nearly every living organism, including humans (Riley et al. 1999, Henry et al. 2015, Kandel et al. 2015). A variety of indices have been developed for quantifying terrain ruggedness. Riley’s TRI is a simple and robust method, and the one applied here (Riley et al. 1999). In this algorithm, ruggedness is quantified as the difference between an individual cell’s value, and the mean of its eight cell neighborhood. The resulting values are classified along a spectrum, from level to extremely rugged. Variations on this algorithm have been used to predict the potential location of archaeological sites and aid in the interpretation of land use strategies in a variety of cultural context (Henry et al. 2015, Warren and Asch 2003, Kandel et al. 2015).

Weights of Evidence values for the study area suggest that ruggedness is a highly predictive of archaeology potential. Extremely rugged terrain covers roughly 8% of the study
area, and is nearly devoid of cultural material. On the other spectrum, level or nearly level
ground covers roughly 37% and contains approximately 88% of the total training data set.

In the WofE results for the Ruggedness index, the contrast values are higher for
calculations performed on the 10m DEM, but the corresponding Studentized contrast values are
lower than those of the 3m DEM. In the most significant class, ‘Nearly level’, the WofE values
for the 10m DEM are (C =3.3 / Stud C=49.47) and for the 3m DEM (C =2.98 / Stud C=53.83).

Although results for both DEMs pass the significance threshold (-C ≥ 2 / Stud C ≥40)
the higher Studentized contrast for the 3m DEM suggests those results area a more reliable

![Figure 6.15](image1)

![Figure 6.16](image2)

predictor in archaeological location modeling. The difference in contrast values is likely a result
of information averaging that occurs within site boundaries when using the 10m raster. Rasters
produced at the 3m resolution capture slight variations in the local terrain found within a site.
This additional information can skew the information results, suggesting greater ruggedness, and
introducing ‘noise’ into the calculations. This same affect was observed to a greater degree in the
WofE values for other indices.
Contrast values for the high and low diversity assemblage types differ significantly for ruggedness. While both datasets follow the same general trend, it’s clear from the results for the ruggedness index that the predictive potential of the high diversity set is diminished when the two sets are combined.

Landform-Multiresolution Index of Valley Bottom Flatness (MRVBF)

The MRVBF index identifies flat areas at lower elevations on the terrain, measuring from just a few meters to several kilometers in area, and characterizes them as ‘valley bottoms’. Valley bottoms serve as buffers to a landscape’s hydrology. Because of this they contribute significantly to depositional and geomorphological processes. Valley bottoms are characterized in the index by two derivatives of slope and elevation-flatness and lowness. Lowness is determined by ranking the elevation of the surrounding area at multiple radii. Flatness is calculated by the inverse of slope. These two measures are combined by multiplication, and scaled to a range from 0 to 1. A complementary Ridge Top Flatness Index is also calculated by
replacing the elevation values for lowness with those ranked in the highest percentile (Gallant and Dowling 2003).

The percentages of study area covered by the five classes of MRVBF vary significantly between the 10m and 3m DEM. The portion of the study area classified as (>3.5) consists of 81.3% (257.8 km²) in the 3m DEM compared to 64.6% (205 km²) in the 10m DEM. The WofE results for this class (>3.5) also show significant variation. In the 3m model, 29.2% of the high diversity assemblage type (1,296-10m² cultural units) is located in this class. By contrast, only 5.7% of the same assemblage type (255-10m² cultural units) is documented in the (>3.5) MRVBF class of the 10m DEM. The resulting contrast values are (C = -2.36 / Stud C = -71.67) for the 3m DEM and (C = -3.4 / Stud C = -52.76) for the 10m DEM. Once again, the 10m DEM presents a stronger contrast result than the 3m DEM, but with a slightly lower Studentized contrast value. The contrast values for both models crosses the significance threshold (C > 2 / Stud C > 40). Although the contrast value is slightly lower for the 3m DEM, the higher Studentized contrast suggests a more reliable product.
The distribution of assemblage types varies significantly in the MRVBF index. Within the 3m DEM, the first three categories (0.0-2.5) contain 40.4% of the high diversity assemblages, but only 5.14% of the low diversity type. Conversely, the (>3.5) classification discussed earlier contains 78.7% of the low diversity assembles, but only 29.2% of type high diversity type. The low diversity assemblage type does not show significance contrast values in any of the five MRVBF classes.

**Solar Radiation**

Solar radiation is fundamental to the energy and water balance of most biophysical processes, and consequently the ecological structure of a landscape (Pierce et al 2005, Butzer 1982: 17-21, Fu and Rich 2002). The amount of solar radiation reaching a given area of the earth’s surface is measured as insolation.

![3m DEM Area Solar Radiation](image)

Until relatively recently aspect served as a common proxy for insolation in archaeological and landscape ecology modeling (Pierce et al. 2005, Deravignone et al. 2015: 339, Langston...
2013: 20). The general hypothesis suggests that in the northern hemisphere, southern oriented exposures receive more solar radiation, and are therefore more appealing for occupation (Langston 2013: 20). However, within the highly heterogeneous landscape of the study area, aspect serves as a poor proxy for solar radiation. To arrive at this important environmental variable the computationally expensive Area Solar Radiation value was calculated for each cell across the project area in ArcMap 10.3.

The Area Solar Radiation algorithm calculates the sum of direct, diffuse, and reflected radiation within a defined area for a set time period, and returns the results as an insolation raster map. Creation of an area insolation map in ArcGIS 10.3 is a four step process: an upward looking hemispherical viewshed is created based on the topography of the study area; next direct radiation for the area is estimated by overlaying the upward looking viewshed onto a direct sunlight map; the process is repeated for diffuse sky map; and finally, these steps are repeated for every location to produce and area insolation or solar radiation map (ESRI 2017b, Fu and Rich 2002).
Area Solar Radiation is one of the only DEM derived environmental attribute showing little or no change when calculated at the 10m or 3m resolution. The largest shift in area between raster classes is 6 km², or only 1.8% of the study area. Processing times however were substantially different. The 10m DEM required roughly 3 hours of processing time, where the 3m DEM required more than 72 hours.

None of the WofE values for Area Solar Radiation crossed the established significance threshold \(- \frac{C}{C} \geq 2 / \text{Stud C} \geq 40\). Results for DEM class and WofE values were only marginally different when calculated at the 3m or 10m resolution. The two classes receiving the least amount of solar radiation contain 16.7% of the study area and only 1% (54-10m² cultural units) of the total training sample. The strongest contrast values for these two classes area found in the 442k-500k category for the 3m DEM \(C = -2.67 / \text{Stud C} = -19.01\), however the Studentized contrast failed to cross the significance threshold. The rasters were still introduced in modelling to counter CI.

**Soil and Sediment**

Soil and sediment are distinct building blocks of an area’s geomorphology. The term sediment refers to deposits formed through the mechanical or chemical weathering of source rock, the transportation and deposition of the resulting particles, and their subsequent postdeposition alteration, which may include pedogenesis, the soil formation process (Rapp and Hill 2006: 25-26). Pedogenesis relies on parent sediment, biota, climate, time, and topological influences such as relief and insolation for soil to form. The term soil designates a stabilized portion of earth-surface material that supports flora and is continuously transformed by biotic and chemical activities and weathering (Rapp and Hill 2006: 38-39). Landforms resulting from these building blocks of geomorphology were touched on in the results for Geology.
Data characterizing the soil and sediment of the study area have been collected by the United States Department of Agriculture and Forest Service for more than four decades (Carlson 1974). This data is currently synthesized in the Blue Mountain Soil Atlas into four categories: soil depth, soil moisture, soil particle size, and soil order. Each category was tested in WofE modeling and found significant, however to prevent CI violations only the most informative layer, soil orders, is included here. The spatial correlation between the three unused categories is shown in Figure 6.23.

**Alifisols**- are alkaline (or sometimes acidic) forest soils which are characterized by a thin A horizon with an underlying clay rich, or argillic B horizon. Alifisols form on relatively young stable land surfaces where no major erosion or pedoturbation events have occurred for thousands of years. As a result of their stability and age, chemical nutrients are commonly retained in Alifisols. Alifisols only form in conditions unfavorable to Mollisol or Spodosol development (Rapp and Hill 2006: 42).

Alifisols cover 23% (km²) of the study area and contain 12.8% (679-10m² cultural units) of the total training sample. The resulting WofE values are only marginally predictive of archaeological potential (C = -0.7 / Stud C = -17.21).
Andisols—commonly form in loose volcanic material such as ash, cinders, or pumice. Andisols often have high organic carbon content which makes them favorable for plant growth, but a low bulk density which results in a low weight-bearing capacity. Andisols are strongly acidic and their physical, chemical, and mineralogical properties often inhibit nutrient exchange. Soil profiles in Andisols are commonly characterized by younger tephra materials being deposited on top of older ones resulting in an “up-building pedogenesis” (McDaniel et al. 2012). Andisols cover the majority of the study area (52.2%-165.6 km²), but contain only 29.8% (43-10m² cultural units) of the total training sample. This results in mildly predictive negative WofE values (C =-0.94 / Stud C=-31.52).

Entisols— are an immature or inhibited soil type that shows little evidence of pedogenesis. Soil development is impeded by erosion or deposition processes outpacing pedogenesis. The surface horizon in Entisols consists of mineral or organic material usually developed over bedrock or slightly altered parent material. The soil profile usually contains few if any diagnostic
horizons. Entisols are commonly located on steep slopes in mountainous regions (Rapp and Hill 2006: 41).

Entisols cover only a minor portion (0.1 km²) of the study area, and lack any association with cultural material. This presents WofE values that fail to pass the significance threshold.

*Inceptisols*- are immature soils similar to Entisols but with slightly more evidence of soil development. Pedogenesis is often inhibited in Inceptisols by a resistant parent material such as volcanic ash. Resistant parent material inhibits development of “normal” soil horizons; an A horizon will form, usually with a weakly developed underlying B horizon. B horizon formation is influenced by the leaching or accumulation of minerals from the parent material. Inceptisols form along the toe slopes and rolling foothills of mountainous areas, usually in sequences of alluvial terraces (Rapp and Hill 2006: 41).

Inceptisols cover 15.7% (50 km²) of the study area, but contain only 0.81% of the total training set (43-10m² cultural units). The resulting WofE values are (C = -3.13 / Stud C = -20.44), suggesting Inceptisols have a strong, but not significant negative correlation with human activities in the study area.

*Mollisols*- contain a deep surface horizon of fertile top soil consisting of dark colored humus rich material. Mollisols commonly form in low lying areas under grasslands or poorly drained hardwood forests. Clay, marl, and basalt are common parent materials. Earthworm activity is often extensive (Rapp and Hill 2006: 41-42).

Mollisols cover only 8.8% (27.9 km²) of the study area, but contain 56.5% of the total training set (2,982-10m² cultural units), resulting in WofE values of (C = 2.6 / Stud C = 93.7). Mollisols are a highly predictive and reliable indicator of archaeological potential within the study area.
Potential Vegetation Group

Potential vegetation (PV) describes the community of plants liable to establish in an area if existing environmental, climatic, and topographic factors are held constant and external interference by humans is absent. Potential vegetation groups may include a number of plant association types. Types are usually labeled according to a dominate overstory and undergrowth species such as “western juniper/big sagebrush”. These communities types are not static; natural disturbance and grown and mortality rates drive plant succession which eventually leads to a climax community. However, abiotic site potential usually results in a high degree of floristic and similarity within seral succession (Powell 2014b: 9, Crowe et al. 2004). The Potential vegetation concept allows for isolation and study of the myriad of small ecology units that form a landscape mosaic (Powell et al. 2007: 1, 7). In a number of cultural contexts, proximity to vegetation zones is the most influential factor in prehistoric site selection (Langston 2013).

<table>
<thead>
<tr>
<th>Physiognomic Class</th>
<th>Upland Forest</th>
<th>Upland Shrub</th>
<th>Upland Herb</th>
<th>Upland woodland</th>
<th>Riparian Forest</th>
<th>Riparian Shrub</th>
<th>Riparian Herb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cold</td>
<td>Cold</td>
<td>Cold</td>
<td>Cold</td>
<td></td>
<td>Cold</td>
<td>Cold</td>
<td>Cold</td>
</tr>
<tr>
<td>Moist</td>
<td>Moist</td>
<td>Moist</td>
<td>Moist</td>
<td>Moist</td>
<td>Warm</td>
<td>Warm</td>
<td>Warm</td>
</tr>
<tr>
<td>Dry</td>
<td>Dry</td>
<td>Dry</td>
<td>Dry</td>
<td>Dry</td>
<td>Low SM</td>
<td>Low SM</td>
<td>Low SM</td>
</tr>
</tbody>
</table>

SM = soil moisture

Figure 6.26 Twenty Potential Vegetation Groups of the Blue Mountains (After Powell et al. 2007: 31)

**Dry Upland Woodland**-Juniper and dry pine woodlands are usually located within foothill zones on gentle to moderate slopes or along ridgetops. Common undergrowth species include bitterbrush, sage brush, bunchgrass, and a variety of forbs including yarrow, biscuit root, and phlox (Powell 2014a: 12, Johnson and Clausnitzer 1992: 126)
Dry Upland Woodland covers only 0.22% (0.7 km²) of the study area, but contains 25.1% (210-10m² cultural units) of the low diversity assemblage type, resulting in significant WofE contrast values of (C =5.04 / Stud C=62.51). Weights of Evidence values are positive for high diversity assemblages as well; however they fail to cross the established significance threshold. The area contains 2.4% (108-10m² cultural units) of the high diversity assemblage type, resulting in contrast values of (C =2.42 / Stud C=24.72).

*Moist Upland Herb*- wet meadows are divided into two types: lotic and lentic wetlands. Lotic areas are the dominate form in the study area, and are commonly associated with rivers, streams, and drainageways. These meadows contain a defined channel and flood plain. Camas and other culturally important plants are commonly found in these meadows (Crowe et al. 2004: 7, 12).

Moist Upland Herb (wet meadows) covers only 1.5% (5 km²) of the study area, but contains 23.8% (1,059-10m² cultural units) of the high diversity assemblage type, resulting in significant WofE values of (C =2.99 / Stud C=84.22). The contrast value is significant for low
diversity assemblages, however the corresponding Studentized contrast is not \( (C = 2.15 / \text{Stud C}=20.41) \).

*Dry Upland Forest*-occur at low to moderate elevation in montane vegetation zones, and support mixed conifer stands dominated by Grand fir, Douglas fir, or Ponderosa pine. Common undergrowth species include elk sedge, pinegrass, birchleaf spiraea, snowberry, ninebark, and bitterbrush. Warm dry forests are the most common type found in the Blue Mountains. Ponderosa pine historically dominated Dry Upland Forests because they are well adapted to a low-severity fire regime which commonly occurs in the area every 5 to 20 years. Dry Upland Forests are usually bounded by moist upland forest along their upper borders, and woodlands and shrublands of the foothills along their lower edge (Powell 2014a: 12). Dry Upland Forest is only mildly predictive of location potential. None of the values for either assemblage type cross the established significance threshold \( (C > 2 / \text{Stud C}>40) \) for this category.

*Dry Upland Shrub/Herb*-Grassland Shrub-are usually dominated by bitterbrush or sagebrush, and may also contain mountain mahogany or the occasional Ponderosa pine or western juniper. Bunchgrasses dominate the herbaceous layer, while prominent forbes include yarrow, rosy pussytoes, fleabanes, and phlox (Johnson and Clausnitzer 1992: 136)

Dry Upland Shrub/Herb results differ considerably for high and low diversity assemblages. The category covers 5.2% (16.5 km²) of the study area, and contains 26.7% (1,190-10m² cultural units) of the high diversity assemblage type, resulting in near significant WofE contrast values of \( (C = 1.90 / \text{Stud C}=55.94) \). By comparison, the same category contains only 2.3% (20-10m² cultural units) of the low diversity assemblage type, resulting in negative contrast values \( (C =-0.81 / \text{Stud C}=-3.57) \).
Moist and Cold Upland Forests – include stands commonly dominated by Douglas fir, Grand fir, pine, or spruce. Undergrowth consists of forbs, along with some mid-height and tall shrubs in warmer environments. Big huckleberry is the dominate undergrowth, however plants such as queencup bead lily, twinflower, false bugbane, swordfern and ginger also commonly occur in this moist zone (Powell 2014b: 10).

Moist and Cold Upland Forest covers 89% (282.3 km²) of the study area, and contains 41.2% (1,830-10m² cultural units) of the high diversity assemblage type, resulting in significant negative WofE contrast values of (C =-2.46 / Stud C=-80.68). The area contains 49% (410-10m² cultural units) of the low diversity assemblage type. The contrast value cross the significance threshold for the low diversity assemblage type, however the corresponding Studentized contrast does not (C =-2.04 / Stud C=-29.54).

Cost Distance to Water

Proximity to water is considered a key factor influencing human locational behavior. Humans need water daily for sustenance and as a key component in many technological
processes; water is awkward and heavy to carry; and, waterways contain and/or draws many desirable, high yield, high return resources (Kelly 2013, Hamilton et al. 2007). The requirement is so fundamental that Whitley (2002) argues an auto-correlative relationship exists between all human activity and distance to water and therefore site selection is not causally conditioned on it. However, the environmental context and procurement strategies adopted by a particular culture commonly result in patterned behavior on the landscape.

The likelihood of movement between two locations is gauged by cost (Llobera 2000). A variety of themes (e.g. social, economic, environmental, etc.) and variables types (e.g. slope, vegetation structure, economic return) can be used to calculate cost. Slope was used as a simple value for cost in this study.

Cost distance to water is classified into low, minor, moderate, and high categories. The percentage of study area covered by the four classes varies significantly between the 10m and 3m DEM. The portion of the study area classified as low consists of 20.9% (66.3 km²) in the 3m DEM compared to 37.2% (118 km²) in the 10m DEM. WofE values were significant for both models in the ‘low’ class, with only marginally differences between them. This is surprising, given the large variance in the land area classified as ‘low’ cost by each model. In the 3m model, 84.2% of the high diversity assemblage type (4,446-10m² cultural units) is located in this class. By contrast, 89.2% of the same assemblage type (4,710-10m² cultural units) is documented in the ‘low’ cost class of the 10m DEM. The resulting contrast values are only marginally different
(C =3.01 / Stud C=79.63) for the 3m DEM and (C =-3.00 / Stud C=56.99) for the 10m DEM. Although, the higher Studentized contrast of the 3m DEM suggests a more reliable predictor.

The minor, moderate, and high classes for both the 3m and 10m DEMs show similar high contrast values, however none of the corresponding Studentized values cross the established significance threshold (−C ≥ 2 / Stud C ≥40).

**Local Dominance**

Local dominance is used in this model to characterize the space syntax of site locations within the landscape. Space syntax is a product of architectural theory and planning, where it is used to quantify the spatial relationships within a structure or settlement (Hiller et al. 1976). Common measures in space syntax include symmetry/asymmetry and distributed /nondistributed. Distributed or nondistributed patterns are used to describe the overall accessibility to a location. A distributed configuration offers multiple routes of access into and between loci. Nondistributed configurations restrict or limit access to a few or only one access route. A location is considered symmetric and integrated when all areas are equally accessible from a given start point. In
contrast, Asymmetrical spaces are only accessible through another space. Asymmetrical spaces are characterized as being isolated and segregated (Hudson 2012, Ferguson 1996, Hiller et al. 1976).

Space syntax has been used to evaluate the spatial arrangement of a number of southwest site including Chaco Canyon, Casa Grande, and several Ancestral Puebloan and Historic Zuni settlements. In each case, space syntax was used to clearly outline the existence of a relationship between social organization and the built environment (Ferguson 1996, Hudson 2012, Hiller et al. 1976). Erin Hudson (2012) carried this concept over into landscape archaeology were least cost path analysis was used to characterize the space syntax of site locations within the landscape.

Spatial syntax serves as a device for characterizing what Margaret Conkey (1984) described as the social geography of hunter gatherer groups. In the context of a foraging society, an isolated and segregated location may be suggestive of a defensive posture or location choices influence by some other negative aspect of social interaction. The location of sites described as distributed and symmetrical may be heavily influenced by other environmental variables,
however their positioning in open settings also suggest a lack of or limited concern for unanticipated or hostile encounters. These location choices may also be influence by other factors such as occupation intensity and duration or the communal nature of the activities. Sites located in open prominent areas (topographic pro may also possibly suggest some form of social status or signaling (Martindale and Supernant 2009, Heyman 2009: 23-24, Llobera 2001, Hudson 2012, Hamilton et al. 2007).

Local dominance developed as a tool to aid in the visualization of lidar data (Hesse 2016, Kokaji and Hesse 2017). Local dominance describes how dominant an observer standing on a particular point (cell)l would of the local surrounding area. The algorithm described the perspective of the observer; the value is greater for higher local elevations and lower for open areas and depressions. Local dominance is calculated for each cell at a specified height (observer height) above the surface, and a specified radius. A height of 1.7m and radius of 15m were used in this model.

Local dominance is a significant predictor of site location potential. The percentage of area place in each class differs significantly between the 3m and 10m DEMs. The area classified as Nondistributed Asymmetrical covers 56.4% (179 km²) of the 3m DEM, and in the 10m DEM 68.1% (216 km²). The resulting WofE values are significant in both DEMs for the total training sample and high diversity assemblage types. The Nondistributed Asymmetrical area in the 3m DEM contains only 5.7% (256-10m² cultural units) of high diversity assemblages, resulting in negative WofE values of (C = -2.54 / Stud C = -53.70). The same class in the 10m DEM contains 12.1% (540-10m² cultural units) of high diversity assemblages, resulting in WofE values of (C = -2.73 / Stud C = -56.63).
The area classified as Distributed Symmetrical produces WofE values that cross the significance threshold (-C ≥ 2 / Stud C ≥ 40) for the 3m DEM, but not the 10m DEM. This class in the 3m DEM covers 15.4% (49 km²) of the study area and contains 66.6% (2,961-10m² cultural units) of the high diversity assemblage type, resulting contrast values of (C = 2.38 / Stud C = 74.85). The same area in the 10m DEM covers only 8.2% (26 km²) of the study area and contains 35.1% (1,559-10m² cultural units) of the high diversity assemblage type, resulting in contrast values of (C = 1.79 / Stud C = 56.84) which do not cross the significance threshold.

Local dominance returned no significant values for the low diversity assemblage type. The relatively even distribution of low diversity results suggests again that the associated behaviors are not strongly influence by environment factors. Information value is diminish when the two assemblage types are combined.
Landsat-7 satellite imaging served as a thematic layer in the probability models developed in this thesis. Satellite imaging can be used to describe the current structure of the landscape, and also detect changes over time. The present land cover may be substantially different than that found in prehistory, and only marginally related to human locational behavior. However, detailing the current land cover and land use of an area is an important element in understanding the visibility and vulnerability of the archaeological record (Giardino 2010).

The Landsat program has operated for the last four decades to document georeferenced images of the earth’s surface. These images generally have a pixel resolution of 15m–60m, and are recorded with a multispectral scanner (MSS) which captures the visible wavebands along with near-Infrared (NIR) and infrared (IR) elements of the electromagnetic spectrum. More than seven satellite missions have successfully contributed to the program since 1972. Landsat-9 is scheduled for launch in 2020, and will extend the program for decades. Each satellite platform used in the program has introduced substantial improvements in technology, while maintaining continuity with the information collected during previous missions. Landsat-7 has an Enhance Thematic Mapper Plus (ETM+) which captures six reflective wavebands of light, plus a seventh in the thermal infrared spectrum. Landsat-7 images have a 30m pixel resolution (Lavender 2015: 12).

The Landsat-7 image used in this model was classified according to the general land use categories provided in Anderson et al. (1976), following a supervised (user driven) process. Classification focused on the density as well as type of ground cover. In ArcGIS 10.3, the composite bands toolbar was used to create a new raster with only those bands showing the greatest contrast in the study area (1, 3, 5, and 7). The Image Classification toolbar was then used
to drawn polygons over known land cover/use types to serve as training areas. Once verified, these training samples served as a class signature file. The Maximum likelihood function was then used to evaluate each pixel in the study area, and assign it to a land cover classification based on the means and variances of each class in the signature file (ESRI 2017c, Scheinsohn and Matteucci 2004, Lavender 2015: ).
The results reveal a structure and distribution similar to the potential vegetation groups described earlier. Meadows, scablands (dry herbaceous ground cover), and exposed surface rock document a significant positive contrast. The greatest contrast was recorded for meadows which cover only 2.4% (7.7 km²) of the study area, but contain 21.8% (970-10m² cultural units) of the high diversity assemblage type, resulting in contrast values of (C = 2.42 / Stud C = 66.53). Forest areas tend to present a negative contrast, with those areas classified as dense mixed conifer stands showing the most significant negative results. Dense mixed conifer stands cover 24.4% of the study area, but contain only 3.4% (155-10m² cultural units) of the high diversity assemblage type, resulting in contrast values of (C = -2.19 / Stud C = -50.85).

The positive and negative contrast values for the Landsat-7 layer were combined together into a binary map. This step was performed to strengthen the Studentized contrast values which is now (C = (-)2.54 / Stud C = (-)83.65) for the respective positive and negative classes.

Landsat Time-series Stacking (LTS) layers multiple classified images from different time periods to identify deviations and trends in land cover (Kennedy et al. 2010). This process is beyond the scope of this thesis, however it an important element of remote sensing in cultural resource management that merits a brief remark. LTS can aid in detailing the processes contributing to the exposure and/or vulnerability of archaeological material. This is particularly important in managed forest lands were a host of human directed and natural events such as logging, fire, or insect infestation are continuously altering the surface conditions of a given area. The LTS process may also contribute to archaeological prospecting by highlighting static patterns on the landscape that possibly indicate previously undocumented cultural features (Giardino 2010). These are important factors to consider when developing an archaeological location model for an area.
Summary of Exploratory Analysis

Exploratory analysis provided a framework to test and refine assumptions regarding the information potential of both dependent and independent variables. Undertaken properly, this important step increases the explanatory power of the final probability model. Exploratory analysis ensures untested, counterproductive, or superfluous data are not incorporated in the final model. The next section will present the validation results for the final model, followed by a comparison of their respective precision and accuracy as calculated through the Kvamme’s gain statistic.
For validation, the AUC for each model was calculated using the R based TOC application (https://amsantac.shinyapps.io/TOCapp/). The cultural units used to describe the archaeological material used as the dependent variable was divided into two randomly selected sets. A 70% training set which was used to build the models, and a 30% validation set. The 30% validation set was not visible to the model until this point. The training and validation sets show identical AUC results for the 3m, 10m, and current Malheur National Forest (MNF) models. AUC values are commonly evaluated on an academic scale. The dotted blue line represents and AUC of 0.5, which suggest a model performing no better than chance.

![Figure 7.1 ROC Curve AUC Results](image-url)
Model Comparison

Current Malheur National Forest Logistic Regression Model

<table>
<thead>
<tr>
<th>Probability</th>
<th>km² of Study Area</th>
<th>% of Study Area</th>
<th>Cultural Units (CU)</th>
<th>% of CU</th>
<th>Kvamme Gain (KG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>76</td>
<td>24%</td>
<td>4,134</td>
<td>78%</td>
<td>0.70</td>
</tr>
<tr>
<td>Medium</td>
<td>112</td>
<td>35%</td>
<td>860</td>
<td>16%</td>
<td>-1.18</td>
</tr>
<tr>
<td>Low</td>
<td>129</td>
<td>41%</td>
<td>283</td>
<td>6%</td>
<td>-5.83</td>
</tr>
</tbody>
</table>

Table 7.1

3m DEM Weights-of-Evidence Model

<table>
<thead>
<tr>
<th>Probability</th>
<th>km² of Study Area</th>
<th>% of Study Area</th>
<th>Cultural Units (CU)</th>
<th>% of CU</th>
<th>Kvamme Gain (KG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>13</td>
<td>4%</td>
<td>2,548</td>
<td>48%</td>
<td>0.92</td>
</tr>
<tr>
<td>Medium</td>
<td>46</td>
<td>15%</td>
<td>2,302</td>
<td>44%</td>
<td>0.66</td>
</tr>
<tr>
<td>Low</td>
<td>258</td>
<td>81%</td>
<td>427</td>
<td>8%</td>
<td>-9.12</td>
</tr>
</tbody>
</table>

Table 7.2

10m DEM Weights-of-Evidence Model

<table>
<thead>
<tr>
<th>Probability</th>
<th>km² of Study Area</th>
<th>% of Study Area</th>
<th>Cultural Units (CU)</th>
<th>% of CU</th>
<th>Kvamme Gain (KG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>22</td>
<td>7%</td>
<td>3,157</td>
<td>60%</td>
<td>0.88</td>
</tr>
<tr>
<td>Medium</td>
<td>36</td>
<td>11%</td>
<td>1,232</td>
<td>23%</td>
<td>0.53</td>
</tr>
<tr>
<td>Low</td>
<td>259</td>
<td>82%</td>
<td>888</td>
<td>17%</td>
<td>-3.82</td>
</tr>
</tbody>
</table>

Table 7.3
A predictive model with a Kvamme gain greater than 0.60 is considered a reliable product. A difference of 0.05 or greater is considered a significant difference. The difference in KG between the 3m and 10m model does not cross this significance difference threshold; however the equally important KG values for medium and low probability do. All of the models presented here pass this accuracy and precision threshold. The KG for the 3m model presents a KG 0.22 greater than the respectable 0.70KG of the current MNF model.

**Blind Validation Area**

<table>
<thead>
<tr>
<th>Probability</th>
<th>km² of Study Area</th>
<th>% of Study Area</th>
<th>Cultural Units (CU)</th>
<th>% of CU</th>
<th>Kvamme Gain (KG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>44</td>
<td>27%</td>
<td>11,642</td>
<td>70%</td>
<td><strong>0.62</strong></td>
</tr>
<tr>
<td>Medium</td>
<td>31</td>
<td>19%</td>
<td>2,593</td>
<td>16%</td>
<td><strong>-0.19</strong></td>
</tr>
<tr>
<td>Low</td>
<td>88</td>
<td>54%</td>
<td>2,311</td>
<td>14%</td>
<td><strong>-3.85</strong></td>
</tr>
<tr>
<td></td>
<td>165</td>
<td>100</td>
<td>16,546</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.4

As a final validation measure the class weights arrived at through calculation of 3m DEM derived variables were applied to a blind validation area. The blind validation area is located at a similar elevation roughly 50km southwest from the study area. Neither this area nor any of the previously documented cultural material found in it were used during model construction. The blind validation area encompasses the Summit Creek watershed, which is a tributary of the Malheur River.
Frequency Ratio of Archaeological Sites

Cultural units form the basis of the dependent variable and are defined in this paper by the density and distribution of artifacts within a site’s boundary. Although artifacts are concentrated within different areas of a site, on average one artifact is found every 10m². Site size and artifact counts vary widely, however the average site contains 68 artifacts. This means the average spatial extent of the sites discussed in this paper is 680m² (or 0.00068km²). This number allows for the determination of the frequency ratio of sites within each class of the probability model. This average equates to 78 sites which is close to the actual number of 83 previously documented sites and isolates.

<table>
<thead>
<tr>
<th>Model</th>
<th>Area (km²)</th>
<th>Sites (68m² Avg)</th>
<th>Frequency Ratio</th>
<th>Area (km²)</th>
<th>Sites (68m² Avg)</th>
<th>Frequency Ratio</th>
<th>Area (km²)</th>
<th>Sites (68m² Avg)</th>
<th>Frequency Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNF</td>
<td>129</td>
<td>4</td>
<td>1:32</td>
<td>112</td>
<td>13</td>
<td>1:8.6</td>
<td>76</td>
<td>61</td>
<td>1:1.3</td>
</tr>
<tr>
<td>3m</td>
<td>258</td>
<td>6</td>
<td>1:43</td>
<td>46</td>
<td>34</td>
<td>1:1.4</td>
<td>13</td>
<td>38</td>
<td>1:0.34</td>
</tr>
<tr>
<td>10m</td>
<td>259</td>
<td>13</td>
<td>1:20</td>
<td>36</td>
<td>18</td>
<td>1:2</td>
<td>22</td>
<td>47</td>
<td>1:0.46</td>
</tr>
</tbody>
</table>

Table 7.5

The next section discusses possible implications for the model results presented here. Model resolution, methodology, and/or the character of the independent and dependent variables are all possible contributing factors. The distribution of assemblage types based on diversity is also addressed further in the next section.
8. Discussion

The 3m model showed a significant improvement over the resulted obtained from the 10m model and the current MNF model (KG 3m = 0.92 / 10m= 0.88 / MNF=0.70). In a final test of performance, each model was compared to the validation set (30%) of archaeological material in the WofE process to determine contrast and Studentized contrast values. The medium classification of both the current NMF model and 10m model failed to reach the previously established significance threshold suggesting lower confidence should be placed in their accuracy. Results in the medium classification for the 10m model were (C= 0.86 /Stud C=26.54), and for the current MNF model (C= -1.03 /Stud C= 27.73).

An increase in survey coverage for medium probability areas is suggested to compensate for model uncertainty seen in the MNF and 10m models. Because of the large body of available data, High probability areas are already modeled with a high degree of accuracy and precision. Medium probability areas are where uncertainty exists in all of these models (although the 3m model crossed the significance threshold for Studentized C but not for contrast C= 1.52 /Stud C= 54.81). Current survey standards for the Forest call for a 100% pedestrian survey of areas modeled as high probability, 40% coverage for medium probability, and 10% for areas classified as low probability. Depending on the extent of a project, and time and resource constraints, medium probability areas should be surveyed at or near the level of high probability areas.

The choice of independent variables likely contributed to the significance of the modeling results. A logistic regression model was produced from both the 3m and 10m datasets. The results for both models were nearly identical to the original WofE model. This suggests the choice of modeling method (WofE or logistic regression) is not a significant factor in a models outcome. However, outliers in logistic regression can affect the outcome significantly (Aidi and Purwaningsih 2013: 2)
Resolution did contribute to a relatively significance difference between the 3m and 10m models, however when either is compared to the current model the difference is stark, particularly in the area classified as low probability. The low probability area is nearly double the size in the current model with only a marginal improvement in the Kvamme gain. The only remaining difference between the models is the underlying variables each set is built on. The current model is built on what could be described as continuous variables like slope and aspect. These continuous variables average out the information potential of the dependent variable. The
3m and 10m models were built on what could be considered discrete variables. Landform and wetness indices take slope and aspect and compartmentalize them into discrete areas on the landscape. This increases the information potential of the model when these evidential layers are compared against the dependent variable. This in turn leads to a greater understand of the distribution of archaeological material within the landscape.

The division of the dependent variable into diversity types likely played an incremental factor in the improvement of model performance. This process demonstrated that even at the relatively low spatial extent of these models (317km²) subtle distinction can be found in the
concentration of archaeological material. In a final display of this subtle distinction, the quantile classification divided the 3m probability model into twenty roughly equal segments of 15km² each (see Figure 8.1). The distribution of assemblage types is presented in Figure 8.3 and the kernel density in Figure 8.2.

The final part of this section departs from the methodological concerns of modeling to discuss possible ecological and anthropological explanations for the distribution of archaeological material. Drawing from a landscape ecology framework, the results of can be summarized in a patch-corridor-matrix model. This model provides a simple spatial language for communicating patterns of human land use in an area with other resource specialists and decision makers. In a patch-corridor-matrix model, a landscape is considered a diverse mosaic that can be delineated into the three categories according to a particular species needs. Patches serve as central places where resources are concentrated, and dwell time and environment affordance are
greatest. Corridors offer connectivity between patches. The background matrix is the dominant element, and supports the patch-corridor arrangement, and promotes the flow of energy through both (Forman 1995, Scheinsohn and Matteucci 2004).

Moist, flat meadows and the ecotone surrounding them functioned as patches for human activity in the study area and likely served as central places with fairly prolonged dwell times. The lack of archaeological features and the relatively sparse nature of most assemblages suggest occupation duration and intensity was limited to a seasonal exploitation strategy. Dry forested areas dominated by western juniper and dry shrub/grasslands are often found immediately adjacent to these meadows, and also performed as patches for human activities. The majority of previously documented archaeological material lies within or in close proximately to these patches. The area within and immediately adjacent to streams and rivers likely functioned as corridors for human movement into the area and between patches. The remainder of the forested landscape served as a background mosaic from which resources were occasionally drawn, but dwell time was limited.

From an anthropological perspective, upland mountain settings like the study area contain abundant resources and were important to PreContact cultures of both the Great Basin and Columbia Plateau. Framed within a resource rank-depression model this area for the past several thousand years has served as a buffer zone between the primary resource patches of several cultures and communities. The resource ranked depression model provides are framework to characterize this period of land use. In its simplest form, the resource rank depression model is analogous to the village life/hunting life contrast described by Hickerson (1965). In this model habitually occupied places are described as core areas or primary patches. The longer a population resides in a place a number of incongruities develop in the distribution of resources.
This is dependent on a number of factors including population size, density, dwell time, and resource variability. Ultimately a foraging induced resource depression develops within the vicinity of habitually occupied places. High ranking prey are presumably pursued opportunistically when encountered. A bioterioration zone develops around the core as highly mobile high ranked resources are driven away from the core and into main resource areas which are described as buffer zones. Presumably under less pressure, prey population in buffer zones steadily increase. Environmental factors like elevation and climatic condition prevent or limit long term habitation of buffer zones, allowing them to persist. This ultimately leads to the development and scheduling of annual subsistence rounds. Buffer zones are exploited by multiple communities. Ecologically, these “resource reservoirs” are similar to the structure described in distribution of wolf populations (Bayham et al. 2012).
9. Conclusion

This research demonstrated the influence a relatively minor change in resolution can have on mapping the structure of the archaeological landscape. A significant gain was achieved by incorporating lidar derived data into ALM. Airborne lidar is increasing in resolution and availability each year. Archaeologists need to be proactive in finding novel ways to exploit this technology.

Future research in the realm of ALM will likely supplement its current role with a focus on preservation. This important element was beyond the scope of this study; however, preservation will increase the scope and utility of ALM when coupled with the other elements of archaeological detection and interpretation presented here. Erosion modeling and change detection are just two of a number of widely supported methods that can direct preservation efforts by identify areas of archaeological vulnerability. The change detection method related to Landsat data briefly discussed either can be applied to lidar datasets as well. A real world example of this is found in the Columbia Gorge Scenic Area of Oregon and Washington, where the 2017 Eagle Creek Fire burned more than 120 km². The burned area likely contains archaeological material significant to the aboriginal populations of the area. Airborne lidar was used to record the Columbia Gorge Scenic Area several years ago. A planned subsequent lidar pass in 2018, will allow the archaeologist for the Scenic Area to apply change detection to the two datasets. This will identify areas vulnerable to loss, and help direct stabilization and/or recovery efforts that can preserve these significant elements of the archaeological record.

An expanded role for ALM can also contribute to better management of our public lands. More than a century of fire suppression on public lands has led to an overabundance of fuels. This has contributed to an increase in the size and intensity of fires, and sparked a paradigm shift in management direction. Land managers are striving for landscape resiliency, and a return of the
biome to its historic range of variability. Archaeological data can inform on this historic range of variability and become a valuable tool to other resource specialists when presented in the Landscape Ecology framework outlined earlier. Some outstanding examples of this and related concepts can be found in the works of Sam Turner (2007), Mike Smith et al. (2011), and Don Hann (2006).

The use of lidar is leading to a geospatial revolution in archaeology, similar in scope and impact to that introduced by radiocarbon dating (Chase et al. 2012). This research is part of the multitude of small efforts contributing to that revolution. More importantly, it contributes to a better understanding of how humans relate to and interact with their environment. And, on a practical level, even an incremental improvement in ALM economizes the time and cost requirements for continuing survey, monitoring, and avoidance measures.
References

Aero-Graphics, Inc.

Agterberg, Frederik P.
2014 Geomathematics: Theoretical Foundations, Applications and Future Developments

Agterberg, Frederik P. and Graeme F. Bonham-Carter

Agterberg, Frederik P. and Qiuming Cheng

Aidi, Muhammad Nur and Tuti Purwaningsih

Aikens, C. Melvin, Thomas J. Connolly, and Dennis L. Jenkins
2011 Oregon Archaeology. Oregon State University, Corvallis, Oregon


Andrefsky, William Jr.

2006 Lithics: Macroscopic approaches to analysis. Cambridge Manuals in Archaeology. Cambridge University Press
Bae, Soyeon, Bjoern Reineking, Michael Ewald, and Joerg Mueller

Bevan, Andrew, Enrico Crema, Xiuzhen Li, and Alessio Palmisano
2013 “Intensities, interactions and uncertainties: some new approaches to archaeological distributions,” in Computational Approaches to Archaeological Spaces, edited by Bevan A. and Lake, M.

Bocquet-Appel, Jean-Pierre and Alain Tufreau
2009 “Technological Responses of Neanderthals to Macroclimatic Variations (240,000-40,000 BP),” Human Biology, Special Issue on Demography and Cultural Macroevolution, Issue 2, Article 10, Vol. 81, pp. 287-307

2002 Soil regionalization by means of terrain analysis and process parameterization, Research Report No. 7, European Soil Bureau

Bonham-Carter, Graeme F.

Burtchard, Greg C. and Stephen C. Hamilton

Burnett, Paul

Butzer, Karl W.
1982 Archaeology as human ecology. Cambridge University Press

Carlson, Garwin T.
1974 Malheur National Forest Soil Resource Inventory. Region 6, United States Forest Service, Department of Agriculture
Carleton, W. Chris, James Conolly, and Gyles Ianonne

Chase, Arlen F., Diane Z. Chase, Christopher T. Fisher, Stephen J. Leisz, and Wieis, Hampel

Chatters, James C.


Chatters, James C., Steven Hackenberger, Anna M. Prentiss, and Jayne-Leigh Thomas

Clarkson, Chris


Conkey, Margaret W.

Crowe, Elizabeth A., Bernard L. Kovalchik, and Marie Jeannette James
2004 *Riparian and wetland vegetation of central and eastern Oregon*. Oregon Nature Heritage Information Center, Institute for Natural Resources

Couture, Marilyn D., Mary F. Ricks, and Lucile Housley

83
Deravignone, Luca, Hans Peter Blanholm, and Giovanna Pizziolo

Environmental Systems Research Institute (ESRI)


Fagan, John L., Bill Roulette, and Terry Ozbun

Ferguson, T. J.

Ferns, M.L. et al.

Foley, Robert A., and Marta Mirazón Lahr

Ford, Arianne and Craig J.R. Hart

Gallant, John C. and Trevor I. Dowling
Giardino, Marco J.

Green, L. Ernestene

Hamilton, Marcus J., Bruce T. Milne, Robert S. Walker, and James H. Brown

Hammond, Heidi, Leandro Zilo, and Alicia S. Castro
2015 “Stratigraphic lithic assemblages from shell middens on the northern coast of Santa Cruz (Patagonia, Argentina),” *Quaternary International*, 2015, pp. 1-10

Hann, Don
2006 “Devil’s in the Details: Using Archaeological and Historical Data to Refine Ecosystem Models at the Local Level,” *USDA Forest Service Proceedings RMRS-P-42CD*, pp. 505-513


Hatzinikolaou, Eugenia G.

Hesse, Ralf

Heyman, Marjorie A. W.

Hickerson, Harold
1965 “The Virginia Deer and Intertribal Buffer Zones in the Upper Mississippi Valley,” In *Man, Culture and Animals: The Role of Animals in Human Ecological Adjustments*, edited by Anthony Leeds and Andrew P. Vayda, Publication 78, Americaon Association for the Advancement of Science, pp. 43-66
Hiller, B., A. Leaman, P. Stansall, M. Bedford

Hudson, Erin

Jaehnig, Menferd

Jaroslaw, Jasiewicz and Iwona Hildebrant-Radke

Jenkins, Dennis L. and Thomas J. Connolly
1990 *Archaeology of Indian Grade Spring: A Special Function Site on Stinkingwater Mountain, Harney County, Oregon*. University of Oregon Anthropological Papers No. 42, University of Oregon, Corvallis, Oregon

Jenness, Jeff
2007 *Some Thoughts on Analyzing Topographic Habitat Characteristics*. Unpublished information paper, Jenness Enterprises, Flagstaff Arizona

Johnson, Charles Grier Jr. and Rodrick R. Clausnitzer
1992 *Plant Associations of the Blue and Ochoco Mountains*. R6-ERW-TP-036-92, United States Department of Agriculture, Forest Service, Pacific Northwest Region

Justice, Noel D.

Kandel, Andrew W., Michael Bolus, Knut Bretzke, Angela A. Bruch, Miriam N. Haidle, Christine Hertler, and Michael Marker

Kelly, Robert L.
Kennedy, Robert E., Zhiqiang Yang, and Warren B. Cohen

Klinger, Ralph E., Elaina H. Gordon, and Toni E. Turner

Kokalj, Žiga and Ralf Hesse

Kohler, T. A.

Kvamme, Kenneth L.


Langston, Lucinda M.

Llobera, Marcos


Martindale, Andrew and Kisha Supernant
2009  Quantifying the defensiveness of defended sites on the Northwest Coast of North America, Journal of Anthropological Archaeology, Vol. 28, pp. 191-204

Mas, Jean-Francois, Britaldo Soares Filho, Robert Gilmore Pontius Jr., Michelle Farfan Gutierrez, and Hermann Rodrigues

McDaniel, P.A., Lowe, D.J.; Arnalds, O., Ping, C.-L.

McDonald, Stan A.

Neal, Mark Lynton
2007  A Predictive Probabilistic Model of Village Site Location within the Santa Ynez Valley, California. Unpublished MA Thesis, University of California, Santa Barbara

Nelson, Greg C.
2011 “Height Above the Nearest Drainage – a hydrologically relevant new terrain model,” 

Odell, George H.

Oetting, Albert C.

Orr, Elizabeth L. and William N. Orr

Padanyi-Gulyas, Gergely, Mate Stibranyi, Gabor Mesterhazy, and Marton Deak
2012 “Familiar Road, Unfamiliar Ground. Archaeological Predictive Modelling in Hungary,” 
CAA2012 Proceedings of the 40th Conference in Computer Applications and Quantitative Methods in Archaeology, Southampton, United Kingdom

Pierce, Kenneth B. Jr, Todd Lookingbill, and Dean Urban

Powell, David C.
2000 Potential Vegetation, Disturbance, Plant Succession, and Other Aspects of Forest Ecology. Technical Publication F14-SO-TP-09-00, United States Department of Agriculture, Forest Service, Pacific Northwest Region


Powell, David C.; Johnson, Charles G. Jr.; Crowe, Elizabeth A.; Wells, Aaron; Swanson, David

Pradhan, Biswajeet, Hyun-Joo Oh, and Manfred Buchroithner

Prentiss, Anna Marie and David S. Clarke

Rapp, George Jr. and Christopher L. Hill

Rieth, Christina B.
2008  Current Approaches to the Analysis and Interpretation of Small Lithic Sites in the Northeast. New York State Museum.

Riris, Philip George Constantine
2014  Exploring the spatial structure of pre-Columbian cultural landscapes in the Alto Parana (Misiones province, Argentina). Unpublished PhD Thesis of the University of Southampton

Swartz, Kevin R. and Jerry Mount

Sawatzky, Don, Gary Raines, and Graeme Bonham-Carter

Scheinsohn, V, Matteucci, SD
2004  “Spaces and species: archaeology, landscape ecology and spatial models in northern Patagonia,” Before Farming, Article 2, pp 1-11
Scudder, S.J., J. E. Foss, and M. E. Collins  

Schmitt, Erich  
2010 Weights of Evidence Mineral Prospectivity Modelling with ArcGIS. EOSC 448 Directed Studies. Unpublished University of British Columbia directed studies material

Smith, Geoffrey M. and Christopher Morgan  

Smith, Mike J., Paolo Paron, and James S. Griffiths  
2011 “Mapping the Late Holocene Landscape Evolution and Human Impact – A Case Study from Lower Khuzestan (SW Iran),” *Geomorphological Mapping Methods and Applications: Developments in Earth Surface Processes*, Vol. 15, pp. 551-575

Solimano, Paul S. and Daniel M. Gilmour  

Turner, Sam  

Valverde, Alberto Jimenez  

Van Leusen, Martijn  

2002 Pattern to process: methodological investigations into formation and interpretation of spatial patterns in archaeological landscapes.

Verhagen, Philip  

Verhagen, Philip and Thomas G. Whitely

Vogel, Sebastian, Michael Marker, Domenico Esposito, and Florian Seiler

Warren, Robert E. and David L. Asch

Weiss, Andrew

Wescott, Konnie L., and R. Joe Brandon
1999 *Practical Applications of GIS for Archaeologists: A Predictive Modelling Toolkit*.

Whitley, Thomas G.

Willey, Gordon R.
Curriculum Vitae
Trent M. Skinner
(October, 2015)

Department of Anthropology
University of Nevada, Las Vegas
4505 S. Maryland Pkwy
Las Vegas, NV 89154-5003
Email: skinnt1@unlv.nevada.edu

Education

University of Nevada, Las Vegas, Las Vegas, NV
Anthropology, M.A. -Archaeology Concentration In progress (Fall 2017)
Anthropology, B.A. May 2014

Archaeological Field Experience

2014 Ais Giorkis, Cyprus (excavation). Principal Investigator: Dr. Alan H. Simmons
University of Nevada, Las Vegas. Crew supervisor. Summer (May –July)

2013 Ais Giorkis, Cyprus (excavation). Principal Investigator: Dr. Alan H. Simmons
University of Nevada, Las Vegas. Crew supervisor. Summer (Jun –July)

2013 Ais Giorkis, Cyprus (excavation). Principal Investigator: Dr. Alan H. Simmons
Lemba Archaeological Field School, University of Edinburgh. Instructor: Dr. Paul Croft.
Student crew member. Summer (May–Jun)

Associations

Society for American Archaeology
Oregon Archaeological Society
Association of Oregon Archaeologists
Computer Applications & Quantitative Methods in Archaeology