ABSTRACT

Title of Thesis:	MACHINE LEARNING APPROACHES TO ARCHAEOLOGICAL PREDICTIVE MODELING IN THE AGE OF WILDFIRE, LAKE TAHOE BASIN MANAGEMENT UNIT, CALIFORNIA AND NEVADA
	Maximilian Anthony van Rensselaer, Master of Professional Studies in Cultural and Heritage Resource Management, 2023
Thesis Directed By:	Dr. Matthew Palus, Cultural and Heritage Resource Management

Machine learning is a powerful tool for archaeological prediction mapping. This thesis compares machine learning approaches to Middle and Late Archaic archaeological prediction in the Lake Tahoe Basin Management Unit, California and Nevada. Specifically, the analysis seeks to answer whether logistic regression, Random Forest, or Maximum Entropy models perform better at archaeological prediction. The explanatory variables used to predict site presence include elevation, slope, aspect, distance to streams, land cover, soil, and geology. Of all three models, Maximum Entropy produced the most accurate predictive models based on combined diagnostic metrics. Predictive modeling is a valuable tool in preventative archaeology, where identifying and mitigating adverse effects to archaeological sites in a time-efficient manner is critical. Environmental challenges such as uncontrolled wildfires provide an impetus for indigenous communities, management agencies, and researchers to employ predictive modeling approaches in preventative cultural and heritage resource management applications.

MACHINE LEARNING APPROACHES TO ARCHAEOLOGICAL PREDICTIVE MODELING IN THE AGE OF WILDFIRE, LAKE TAHOE BASIN MANAGEMENT UNIT, CALIFORNIA AND NEVADA

by

Maximilian Anthony van Rensselaer

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Professional Studies
in Cultural and Heritage
Resource Management
2023

Advisory Committee:
Professor Matthew Palus, Chair
Professor Jonathan Resop
Professor Kathryn Lafrenz Samuels
Professor George Hambrecht

Foreword

This research is the product of my experience as a Fire Archaeologist and Resource Advisor with the Bureau of Land Management in northern Nevada and as a Wildland Firefighter with the State of Nevada. As a Fire Archaeologist, I surveyed large swaths of land for cultural resources ahead of planned fuel treatments such as native reseeding projects and fuel break construction. Often, I would work with the Fuels/Fire Division to identify an Area of Potential Effects for the project, allowing me the opportunity to select and survey an area that would benefit from fire resilience while also providing the best chance to identify and protect archaeological sites.

I planned and executed cultural resource surveys in post-burn areas such as the Martin Fire, which burned 439,230 acres in Nevada in 2018. That was the largest fire in Nevada's history and the largest in the country that year. Planning a survey strategy across an area that big, the need for a better way to devise a sampling strategy became apparent to me. The techniques explained in this thesis could have made my life easier. To paint the picture of life as a Fire Archaeologist, you are hiking through ashen landscapes without a tree in sight and the sun bearing down on you. Suddenly you find yourself at the edge of an unrecorded archaeological site with 10,000 artifacts distributed across multiple saddles and draws. The once-hidden archaeological site reveals itself as the sage and cheatgrass have burned away, leaving obsidian flakes gleaming and shimmering in the sunlight.

As a Wildland Firefighter (shout out to Battleborn Helitack), we landed our helicopter in the middle of a lithic scatter, set up helicopter supply drop sites within archaeological sites, and I came across numerous flakes digging fireline in remote areas that probably hadn't been surveyed for cultural resources yet. On the job, the priority is protecting life and property, and there isn't time to stop to take a GPS point, then call up the local land manager and suggest they get someone out there to record a site. That's the job of a Resource Advisor on an incident. However, those experiences opened my mind to the need for archaeological prediction using GIS before and after wildfires. Before the fire, we can identify those areas most likely to contain unrecorded sites and prioritize surveys there. After the flaming front has passed, new sites rise from the ash, and known sites can be updated and monitored for damage. When the fires burn as big as they tend to these days, the cultural resources work after a fire can be daunting. A predictive model such as the one outlined in the following pages provides a way to more efficiently plan for and manage effects to cultural resources in the context of fire management.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AICc – Akaike Information Criterion

DEM – Digital Elevation Model

GLR – Generalized Linear Regression

GIS – Geographic Information Systems

LTBMU – Lake Tahoe Basin Management Unit

MaxEnt – Maximum Entropy

RF – Random Forest

Chapter 1: A View of Things to Come

Purpose of the Study

The United States Forest Service's Lake Tahoe Basin Management Unit (LTBMU), located in portions of California and Nevada, is home to rich cultural history both in the names and places that continue to be important to the Washoe people and others, as well as in the cultural remnants of Washoe ancestors that occupied the region. However, the regional threat of increasing wildfires and the environmental degradation caused by landslides, flooding, and destructive processes associated with high-intensity burning present problems for the integrity of cultural resources in the LTBMU.

As a Fire Archaeologist with the Bureau of Land Management, Winnemucca District, I witnessed first-hand the difficulty associated with monitoring and updating previously recorded archaeological sites affected by wildfires that burned scores of acres, as well as the challenge of identifying new sites exposed by the burned landscape. What is the best way to prioritize archaeological mitigation and data recording across hundreds of thousands of acres with limited time and resources? Predictive modeling offers one avenue towards directing those efforts in a streamlined, efficient way.

This thesis will explore Geographic Information Systems (GIS) approaches to creating predictive models for Middle Archaic and Late Archaic archaeological sites in the LTBMU. GIS-based spatial analysis is essential to reduce archaeological risk and support decision-making in cultural resource management (Danese et al. 2013). Predictive models can aid indigenous groups, researchers, and government agencies in responding to or mitigating natural disasters that have the potential to affect cultural resources, allowing these groups to prioritize salvage archaeology and cultural resource protection most efficiently when faced with limited time and resources.

Research Questions

In addition to providing a valuable tool for cultural resource preservation, this thesis uses graphical representations to identify differences in land use strategies between the Middle and Late Archaic periods. How do Washoe ethnographic observations of land use correlate to the results of Middle and Late Archaic GIS investigations? How do land use strategies differ between the Middle and Late Archaic, and what environmental variables contributed most to differential site selection? What type of model is most effective at archaeological prediction in the study area? The explanatory environmental variables involved in the predictive analysis of archaeological sites include slope, aspect, elevation, distance to streams, geologic unit, soil type, and land cover. Other predictive models have been developed using similar environmental variable constraints, including a statewide predictive model for Washington that included all seven explanatory variables used in this study and one additional variable, landform (WISAARD 2009).

Previous archaeological studies in the LTBMU have investigated relationships between Middle and Late Archaic site selection patterns (Heizer and Elasser 1953; Elston 1971; Davis, Elston, and Townsend 1973; Elston et al. 1977). Davis, Elston, and Townsend (1973) developed a model to predict the number and location of archaeological sites for the Fallen Leaf Lake area of the Eldorado National Forest within the LTBMU. However, due to the computational power limits, that investigation could not produce a predictive model for the entire LTBMU. This thesis will create multiple high-resolution predictive models for the entirety of the LTBMU.

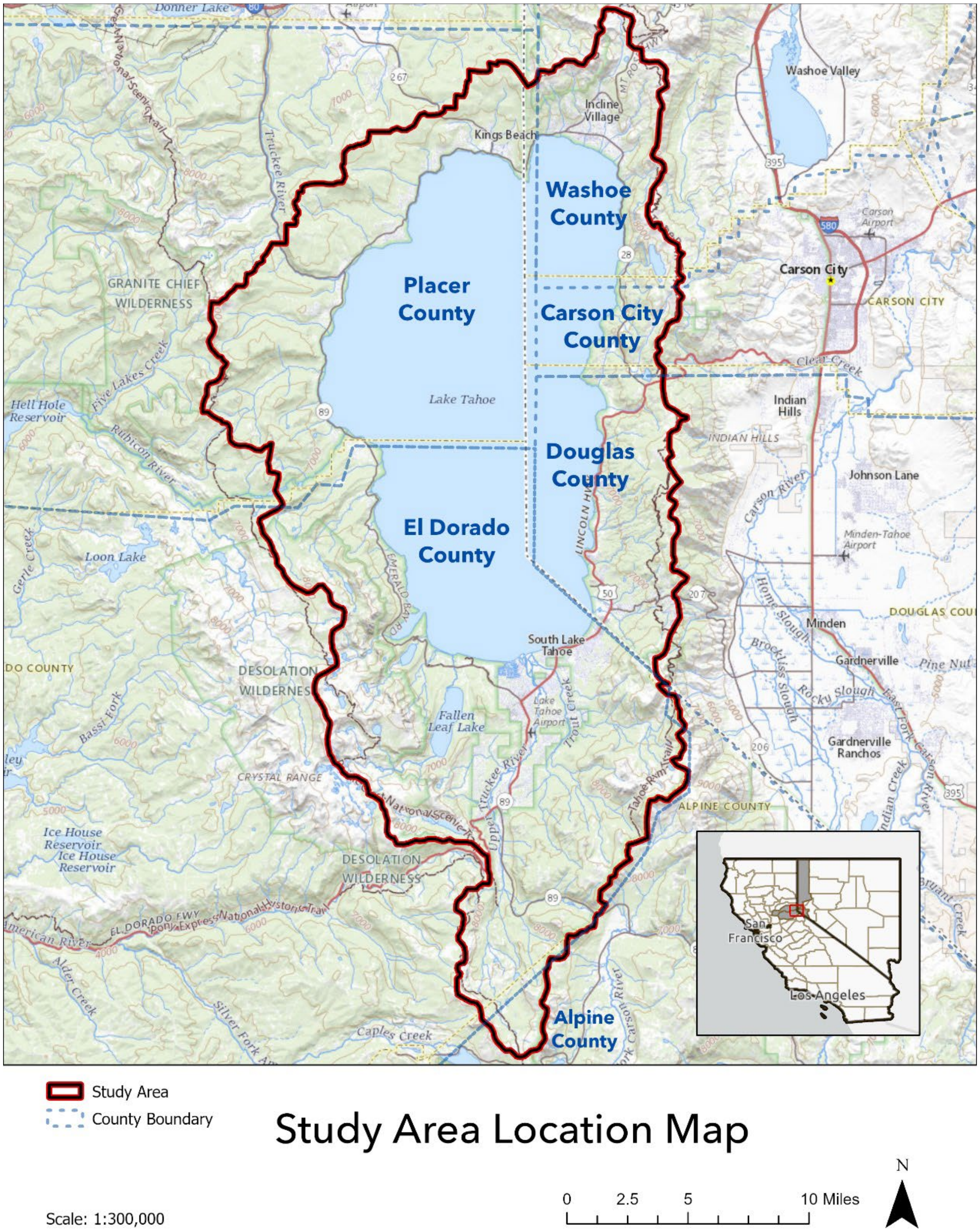

Figure 1: Study area location map (map by athor)

Study Area Overview

The Lake Tahoe Basin Management Unit (LTBMU) is a National Forest designated by the United States Forest Service and is located within parts of

California and Nevada. It includes portions of Placer, Alpine, and El Dorado Counties in California, and Washoe, Carson City, and Douglas Counties in Nevada. The LTBMU encompasses over 154,000 acres of land and was created in 1973 from sections of the Toiyabe, Eldorado, and Tahoe National Forests to create a unified approach to land management as development in the region increased (USDA 2022). The LTBMU represents the western extent of the Great Basin region, including valleys, basins, lakes, and steep mountain ranges of the Basin and Range Province (USGS 2004). The LTBMU is also part of the Sierra Nevada geomorphic province, consisting of a nearly 400-mile-long tilted fault block with high, rugged peaks (California Geological Survey 2002). Elevation within the LTBMU ranges between 10,883 feet and 6,170 feet.

The Washoe Tribe has historically managed the land now contained within the LTBMU. Although Washoe ethnography will be discussed later in this thesis, it is important to mention colonization forced the Washoe people out of their lifeways and homeland around Lake Tahoe. Rhiana Jones, Environmental Program Director for the Washoe Environmental Protection Department, said that Washoe people "weren't allowed to continue their seasonal migration to and from Lake Tahoe, for their summer camps. So, they stopped, essentially, managing the land" (Nevada Today 2021). However, efforts such as the MÁYALA WÁTA Restoration Project at Meeks Meadow, located within the LTBMU, are examples of traditional indigenous ecological knowledge being used to influence and guide modern land management practices within the LTBMU (Nevada Today 2021). The basis of the models described in this thesis, including the variables used to predict archaeological site location, was developed in consultation with the former Tribal Historic Preservation Officer of the Washoe Tribe of Nevada and California, Darrel Cruz, now retired.

GIS and Statistical Analysis

This analysis uses data from previously recorded prehistoric archaeological sites in the LTBMU that were assigned to either Middle or Late Archaic temporal categories based on diagnostic lithic artifacts obtained from site records. Portions of the LTBMU remain unsurveyed, covered by dense vegetation, or may erode to expose previously unseen cultural remains. To fully understand site location trends and the potential presence of previously unrecorded prehistoric archaeological sites, this study employs computer algorithms and maps to visualize spatial relationships. Gleaning trends in land use and settlement strategies via predictive modeling can direct future identification efforts.

In answering questions related to settlement and land use patterns, this thesis will also use analytical metrics to evaluate relative model fit and performance between three model types, including Generalized Linear Regression (GLR), Random Forest (RF) and Maximum Entropy (MaxEnt). What are the strengths and weaknesses of each model type, and which method is most useful for GIS-based archaeological investigations? ArcGIS Pro provides built-in tools for these three model types that allow for predicting a dependent variable, the presence of archaeological sites, with explanatory variables.

Machine Learning

This thesis uses three forms of machine learning: logistic regression, RF, and MaxEnt. Regression models are a standard statistical analysis method, and the "predominant statistical technique in constructing archaeological predictive models" (Wachtel et al. 2018:29). RF and MaxEnt are relatively new GIS tools derived from algorithms that leverage twenty-first-century computational power. Machine learning-based modeling approaches use input data to train a model to evaluate relationships between explanatory and dependent variables, allowing for the analysis and visualization of spatial data (Castiello and Tonini 2021:110). Machine learning models can use built-in complex algorithms to make accurate predictions, even with smaller input datasets (Castiello and Tonini 2021; Wachtel et al. 2018). In similar archaeological predictive modeling analyses, MaxEnt and RF approaches have produced efficient, reliable, and valuable predictive models compared to regression (Castiello and Tonini 2021; Wachtel et al. 2018).

Gaps in the Archaeological Record

While theoretical approaches create avenues for evaluating subsistence and settlement patterns, gaps in the archaeological record require further testing to establish more evidence and understanding of cultural change. These research objectives will require the continued search for evidence of paleoenvironments, buried archaeological deposits, and "extraction of subsistence and chronological data from surface assemblages" (Elston and Zeanah 2002: 121). However, crucial paleoenvironmental evidence can be lost in the western United States, particularly in the Eastern Sierra Front, where devastating wildfires are becoming more common. Paleoenvironmental evidence, the integrity of surface assemblages, and buried sites exposed by erosion are all impacted by the growing threat of large wildfires. Based on principles of human behavioral ecology, this thesis will apply archaeological predictive modeling to a study area threatened by wildfire, the Lake Tahoe Basin Management Unit. The analysis will include standard regression techniques and machine learning approaches, comparing the results of predictive models that use logistic models of generalized linear regression (GLR), Random Forests (RF), and Maximum Entropy (MaxEnt).

Threats to the Archaeological Record

The anthropogenic use of fire is a critical cross-cultural phenomenon used for thousands of years to restore and maintain ecosystems (Mason et al. 2012: 187). However, in North America, the prevailing doctrine of immediate fire suppression in most cases of wildfire events has "disrupted ecological processes, elevating the risk of wildfire, insects, and disease, affecting the health and availability of resources on which the Tribes depend" (Mason et al. 2012: 187). Wildfire, and its inevitable fire suppression methods, can physically destroy archaeological sites. Severe fires often impact the ability of researchers to recover critical data that helps to inform gaps in the archaeological record.

Figure 2: Prehistoric archaeological site in a post-wildfire environment (photo by author)

Additionally, hot and destructive wildfires can permanently alter valued traditional indigenous cultural spaces and the modern communities in which the current indigenous communities live. In 2021, the Tamarack Fire burned 68,637 acres of land in Alpine County, California and Douglas, Lyon, and Nevada County Nevada. The fire forced evacuations in Woodfords, home of the *Hung-A-Lel-Ti*, the southern community of Washoe Tribe members. Firefighters' efforts to save Woodfords were successful, but the incident highlighted the threat wildfires pose to Washoe communities in the region.

Summary by Chapter

Chapter 2 provides a theoretical overview of the work carried out in this thesis. Gaps in the archaeological record require further investigation to link cultures in transition affected by changing environments, cultural shifts, and adaptive strategies. Environmental threats also affect the ability to identify, record, and preserve finite archaeological resources that have the potential to inform understanding of past human behavior. Behavioral ecology is a theoretical framework that can model optimal choices by human agents in achieving fitness-related goals and the costs, benefits, and constraints associated with those goals (Broughton and O'Connell 1999).

The introduction of the bow and arrow, a key feature separating the Middle and Late Archaic periods investigated in this thesis, is one example of behavioral optimization associated with improving the efficiency and handling of resources. Technology such as Geographic Information Systems (GIS) that integrates statistical analysis can aid behavioral ecological research, such as adaptations in settlement patterns over time, providing a lens through which to view changing adaptive strategies over the long term. Recent machine learning-based GIS approaches offer an efficient and accurate way to implement modeling solutions to archaeological problems.

Chapter 3 provides a more detailed introduction to the study area and a prehistoric overview of the region, as well as a discussion of Washoe ethnography. The chapter begins by noting the utility of culture phases and culture history if only to distinguish cultural differences in the context of behavioral ecology while discarding notions of cultural evolution. The study area is described in terms of its location within geographic and modern political boundaries and within one of the three major zones for plant and animal resources in the Sierra Nevada region, the Boreal zone.

Cultural chronology is separated into four significant periods of the Great Basin and Tahoe Basin: the Prearchaic, Early Archaic, Middle Archaic, and Late Archaic. These changing periods correspond with a general trend from hunting-based societies to more diverse subsistence strategies that may result from changing environments, subsistence opportunities, migration patterns, and population pressure.

The primary focus of this thesis is the Middle Archaic and Late Archaic. Cultural phases corresponding to the Middle Archaic in the Tahoe Basin include the Martis Complex, comprised of the Early Martis and Late Martis phases. The Late Archaic period corresponds to the Kings Beach complex, comprised of the Early Kings Beach and Late Kings Beach phases. Technological and subsistence differences distinguish these phases from one another in the archaeological record.

Washoe lifeways continue even as colonization forced significant changes. The end of Chapter 3 discusses ethnographic observations of the Washoe from the early to mid-twentieth century, including the cultural, spiritual, and mythological importance of Lake Tahoe reflected in Washoe subsistence and settlement.

Chapter 4 outlines the methods used in the GIS analysis, including ArcGIS Pro spatial modeling tools: Generalized Linear Regression (Logistic Regression), Forest-based Classification and Regression (Random Forest), and Presence-only Prediction (Maximum Entropy). The chapter describes the acquisition and classification of prehistoric sites in the LTBMU into Middle Archaic and Late Archaic datasets used to create the GIS models. Additionally, information regarding sources of the explanatory variables such as elevation, slope, aspect, distance to streams, land cover, geology, and soils is discussed.

Chapter 5 presents the graphical and statistical results of Middle and Late Archaic GIS analyses separated by model type. The chapter includes metrics that will be used to evaluate model fit and performance, as well as numbers and percentages of relative archaeological presence and absence points between the models. Regression models are presented in binary outputs for Middle and Late Archaic datasets. Diagnostic metrics for GLR results include validation accuracy, Deviance Explained, a percentage of archaeological presence explained by the explanatory variables, and Akaike's Information Criterion (AICc), a relative measure of model fit and performance.

The predictive power of the models is compared using validation accuracy for Middle and Late Archaic datasets. Calibration data is used to train the model, while validation data is used to test the model. Therefore, calibration accuracy is expected to be higher than validation accuracy, and high validation accuracy indicates a good model fit. The RF and MaxEnt results also provide measures of predicted archaeological presence and absence in the study area and relative explanatory variable importance rates for the Middle and Late Archaic.

Diagnostic metrics specific to the MaxEnt models include Omission Rate and AUC. Omission Rate measures how frequently the model incorrectly predicted an archaeological presence or absence point. The AUC value measures the model's accuracy in predicting known presence points as presence locations. Additionally, MaxEnt results include rates of archaeological presence percentage throughout the study area and variable coefficients that measure the strength of each variable in relation to the dependent variable, archaeological presence.

Chapter 6 is the analysis chapter that presents interpretations of the results chapter, evaluating each of the models for overall performance and connecting the results of the models to archaeological literature. The chapter is broken into two halves: quantitative analysis of model metrics and qualitative analysis of the maps produced by each model. This quantitative analysis includes a discussion of each model's prediction power and the strengths and weaknesses of different modeling approaches. This section also analyzes explanatory variables, showing the possible benefits of including variables that aren't typically regarded as the strongest predictors of archaeological site location. The qualitative analysis details the connection to Washoe ethnography and answers twentieth-century debates, providing supporting evidence to archaeological observations of Middle and Late Archaic settlement patterns in the Tahoe Basin by leveraging twenty-first-century computational power.

Chapter 7 provides a summary of the overall conclusions of the thesis, as well as acknowledging some of the limitations of the analysis, and suggests paths for future research. This thesis will not be research for the sake of research, but can be a helpful tool that might aid indigenous communities, management agencies, and archaeologists in preserving and protecting cultural resources in the Lake Tahoe Basin.

Chapter 2: Theory, Practice, and Computational Power

Chapter Outlook

This chapter discusses gaps in archaeological knowledge and contends that further testing can help close these gaps. Because archaeological resources are nonrenewable, environmental challenges such as wildfire, erosion, and flooding threaten their preservation. Theoretical approaches such as behavioral ecology help interpret archaeological evidence once recovered. The theoretical framework of behavioral ecology seeks to map the networks of behavior, fitness, and environmental change to model cultural change through time and can be correlated to ethnographic evidence. Spatial modeling is ideal for behavioral ecological modeling of subsistence, settlement, and technological change. Additionally, spatial modeling is helpful for preventative archaeology because predictive analysis can prioritize resources towards protecting and preserving finite cultural and heritage resources.

Theoretical Approaches

Theoretical approaches are critical for directing researchers toward essential contributions to understanding the human past and developing material culture. In combination with methodological tools to extract information from data, theoretical guidelines such as behavioral ecology can help bridge the gap between archaeological and ethnographic data and unify a general theory of human behavior with the forces of natural selection (Codding and Bird 2010). In this thesis, I will apply behavioral ecology, preventative archaeology, and statistical extrapolation methods to my research.

Behavioral Ecology

Behavioral ecology is a comprehensive theoretical framework that seeks to model optimal approaches to possible fitness-related goals and benefits, costs, and constraints to achieving those goals (Broughton and O'Connell 1999). The perspective aims to model the relationship of humans to the natural environment, understanding them as part of a cohesive ecosystem. This relationship can be examined clearly in the context of resource exploitation, settlement patterns, and demographic arrangements (Jochim 1976: 9)

In the context of behavioral ecology, a *goal* is an optimal rate of resource acquisition; costs and benefits to attaining that goal are measured as *currency* (e.g., caloric expenditure); and *constraints* include characteristics of the social and environmental context such as available technology or the distribution and density of resources (Winterhalder and Smith 2000: 54). Models developed from behavioral ecology concepts can "link behavior to expected material outcomes" (Codding and Bird 2015: 9).

Behavioral ecology applies evolutionary ecological models to the study of the diversity of human behaviors (Winterhalder and Smith 2000: 51). Additionally, optimization of behaviors among a set of alternative behaviors is a product of natural selection that works to improve the fitness of the organism (Broughton and O'Connell 1999: 154). Behavioral optimization can take the form of altering settlement patterns or expanding diet breadth to reduce caloric costs associated with foraging, and technology changes related to improving the efficiency of handling resources, such as the proliferation of basketry and the introduction of the bow and arrow in the late Holocene in California (Broughton and O'Connell 1999: 155).

Behavioral ecology attempts to answer how socioenvironmental factors influence the choice of behaviors among a set of alternatives, and how natural or sexual selection will affect the cost and benefits of those behavioral choices (Winterhalder and Smith 2000: 52-54). Practical strategies for survival, typically observed as a measure of optimization such as efficient resource acquisition, should be more likely to be passed to the next generation as "evidence of how natural selection has shaped decision making" (Codding and Bird 2015: 10).

Behavioral ecological approaches rely on empirical archaeological data to test theoretical hypotheses that are typically based on mathematical or graphical models (Winterhalder and Smith 2000: 52). Archaeological material can be damaged, and the ability to recover material culture can be affected by natural processes such as wildfire, erosion, and flooding.

Behavior ecology theoretical approaches allow researchers to model subsistence and settlement strategies by simulating and modeling diet breadth, patch choice, and risk (Elston and Zeanah 2002: 112). This theory attempts to model human behavior by evaluating caloric cost and return in the context of paleoenvironments, the relative ranking of resources that existed based on paleobotanical evidence, and modeling risks associated with resource selection and procurement strategies (Elston and Zeanah 2002: 112). These predictions can be tested empirically and correlated with ethnographic observations to infer the plausibility of decision-making regarding subsistence. Despite the focus on quantitative analysis associated with behavioral ecology, ethnographic observations remain essential for validating models and bridging the gap between theory and archaeological data (Winterhalder and Smith 2000: 52).

Preventative Archaeology

Preventative archaeology is a research method designed to reduce archaeological risk, or the loss of finite, non-renewable resources by directing research efforts before development or environmental degradation occurs.

Archaeological risk occurs as a product of hazard, vulnerability, and exposure (Danese et al. 2013: 42). An example of a hazard is a natural risk factor such as a wildfire event. Vulnerability is the value at risk, such as an archaeological site, artifact, or ecofacts that can be exposed to the hazard. GIS-based predictive models are a valuable tool to provide decision support for "defining survey priorities and facilitating discoveries; saving time and money; especially over large areas" (Danese et al. 2013: 42).

<u>Models of Prehistoric Occupation in the Tahoe Basin</u>

Robert F. Heizer and Albert Elasser (1953) first detailed two distinct archaeological complexes in the region, the Martis complex and the Kings Beach complex. Based on observations made during excavations and surface collections, Heizer and Elasser characterized Martis sites as basalt-oriented, with rare or absent obsidian, and typified by large, heavy, and crude atlatl projectile points. Heizer and Elasser posited that most Martis sites fell within the Transition Zone between 2,500-6,000 feet in elevation and were located along game trails and productive seed gathering areas (Heizer and Elasser 1953: 19).

Meanwhile, Heizer and Elasser defined Kings Beach sites as dominated by obsidian and silicate rock types, with rare occurrences of basalt, and the presence of small, lightweight projectile points used for bows (Heizer and Elasser 1953: 20). To Heizer and Elasser, Kings Beach Complex sites were representative of a fishing and seed-gathering economy that began in AD 1,000 and continued until historic Washoe encounters (Heizer and Elasser 1953: 21). Importantly, Heizer and Elasser believed Kings Beach and Martis sites were mutually exclusive in terms of site location (Elston 1971: 10).

Robert Elston (1971) revised the cultural sequence of the Tahoe Basin and made different conclusions compared to Heizer and Elasser (1953). Elston, too, identified a change in intensity of occupation of sites between the Martis and Kings Beach complexes. However, while some sites were occupied with less intensity, other sites in the Tahoe Basin were continually occupied and not abandoned between the two complexes (Elston 1971: 138). Additionally, Elston did not observe the scarcity of basalt artifacts during the Kings Beach that was proposed by Heizer and Elasser. With a renewed cultural sequence based on artifact typology, Elston's research supported a link between Martis and Kings Beach sites, identifying overlap in site locations and continual occupation between the two complexes (Elston 1971: 140).

GIS and Statistical Extrapolation Methods

Geographic Information Systems (GIS) based approaches using predictive modeling can provide utility in maximizing the potential to prioritize research in strategic areas before the loss of a finite and non-renewable resource due to environmental degradation and change. Predictive models aim to study known environmental or social factors that influence an archaeological site's position in space to generate a set of rules that can be used to predict the location of undiscovered sites (Danese et al. 2013: 43). Common factors include slope, elevation, aspect, distance to water, viewshed, and landform. GIS models use spatial analysis to allow "investigation into the spatial distribution of phenomena" and the relationships between phenomenon and the relevant factors (Danese et al. 2013: 43). The results of the model attempt to express "a probabilistic relationship between human behavior and prior existing spatial conditions" (Verhagen and Whitley 2011: 71).

Predictive modeling applies to behavioral ecology theoretical approaches because human behavior should follow a broad pattern relating to optimization of fitness in an ecological context. Effective predictive models should be able to identify such patterns (Broughton and O'Connell 1999: 158).

Spatial Models

Statistical extrapolation methods, also referred to as "inductive" or "correlative" models rely on known site data to make comparisons to environmental factors, and then extrapolate those predictions to areas where no site data exists, typically using logistic regression or machine-learning algorithms (Verhagen and Whitley 2011: 51). This method of correlation can be useful for exploratory data analysis and understanding parameters for defining cause-and-effect relationships for a particular behavior, that can contribute to the development of theory building (Verhagen and Whitley 2011: 90). However, the research design of a spatial analysis model should be grounded in archaeological theory to produce meaningful results. Often, GIS models attempting to predict archaeological site locations incorporate all known sites without respect to the differing population densities and subsistence practices that would affect land use decisions through time (Yaworsky et al. 2020: 2). Therefore, predictive models incorporating statistical extrapolation methods should "limit archaeological response data to sites from the same time period" to capture the behavioral adaptations that would lead to a pattern of land use and site selection (Yaworsky et al. 2020: 2).

Reductionist models that appeal to the "lowest common denominator of suitability for all behaviors or events" over many different time periods fail to provide a meaningful explanatory framework (Verhagen and Whitley 2011: 73). Additionally, environmental predictor data should conform to theoretically grounded expectations of what factors influenced settlement decisions (Yaworsky et al. 2020: 2). Drawing from behavioral ecology, explanations of human behavior and predicted outcomes are linked when environmental factors are correlated to land use strategies (Yaworsky et al. 2020: 2).

Logistic regression is a local regression model useful for investigating spatially varying relationships between dependent and independent variables resulting in a spatial distribution of model coefficients. This regression technique is well-established in statistical theory and has been applied to infer spatial trends in archaeological data (Yaworsky et al. 2020).

Machine Learning

Machine learning approaches such as RF and MaxEnt are flexible processes of predictive power that arise from an iterative process that uses data to calibrate an algorithm and results in a classification model. RF is a form of classification and regression tree analysis in which a decision support system based on an algorithm is used to predict an unknown outcome, for example, the location of an archaeological site. Predictor variables partition the dataset into smaller groups to homogenize the value of a response variable (Yaworsky et al. 2020: 9). These subsets can be further divided based on additional predictor variables resulting in a decision tree. The benefit of RF is that many trees are created, reducing the risk of overfitting the model as a function of built-in randomness, which decreases variance in the model. Decreasing variance increases the accuracy of the model's predictive capability to identify unknown archaeological sites based on input variables.

MaxEnt models work by estimating probabilities of occurrences over a geographic area based on "presence-only observations and environmental variables" (Galletti et al. 2013: 46). This form of model can create complex predictive spatial distributions based on a low sample size of known archaeological sites, serving as the presence-only observations. The concept of Maximum Entropy begins with an assumption of equal distributions of presence points across the study area, constituting a distribution of maximum entropy (Yaworsky et al. 2020: 8). Prediction points are constrained by environmental variables that increase density at the variable value based on presence points (Yaworsky et al. 2020: 8). Uniformity of the mathematical distribution implied by maximum entropy is leveraged against the constraints of the non-uniform environmental variables, resulting in a relative rate of occurrence of a feature's presence across the landscape (Yaworsky et al. 2020: 9).

MaxEnt models have resulted in compelling predictions and understanding of relationships between environmental and archaeological data and provide a valuable supplement to archaeological field data (Galletti et al 2013, Healy et al 2016, Yaworsky et al 2020).

Machine learning models are practical tools for archaeological predictive analysis. While they cannot be 100% accurate, these tools can be leveraged to identify general spatial trends in archaeological data. Accurate predictions of archaeological presence and absence are effective strategies to prioritize archaeological identification and mitigation efforts, especially in preventative archaeology where funds and efficiency should be directed toward the most value.

Chapter 3: Past Understanding and Cultural Continuity

Chapter Outlook

This chapter frames the GIS analysis presented in this thesis by summarizing the culture history of the Lake Tahoe Basin, focusing primarily on the Middle and Late Archaic periods, for which a wealth of archaeological evidence exists compared to the Prearchaic and Early Archaic periods. An outline of human occupation in the Lake Tahoe Basin follows a brief examination of the utility of the culture-historical model. The discussion traces subsistence, settlement, and technology developments that coincide with the environmental shifts that forced adaptive cultural change. Last, this chapter addresses Washoe ethnography, helping to elucidate the behaviors reflected in the archaeological record. An effective predictive model will reflect not only observations from material culture but also the behaviors and worldview of the Washoe people.

Culture Phases and Culture-History

Archaeologists use cultural phases to compare and distinguish periods of prehistory from one another (Ringhoff and Stoner 2011: 5). Gordon Willey and Philip Phillips used the term "phase" to mean "an archaeological unit possessing traits sufficiently characteristic to distinguish it from all other units similarly conceived, whether of the same or other cultures or civilizations, spatially limited to the order of magnitude of a locality or region and chronologically limited to a brief interval of time" (Willey and Phillips 1958: 22). This idea of a culture phase is related to the culture-historical theoretical paradigm that advocated for an evolutionary model of cultural development. These theories of cultural evolution implicitly justified colonialism in the United States using the rationale that one group of humans had advanced more than another, with domination and conquest the ultimate result of that advancement (Mergen 2014: 57).

While evolutionary theories of culture have rightfully fallen out of favor in American archaeology, the term "phase" is still valid when referring to unified characteristics that define an episode of cultural occupation. In the 1960s, spatial relationships between distinctive styles of prehistoric projectile points resulted in named projectile point types with known temporal ranges that informed the understanding of distinct cultural phases within the western Great Basin (Elston 2013: 128, d'Azevedo 1986: 466). In addition to assessing cultural phases based on artifact morphology, archaeologists divide prehistory by gleaning differences in adaptive strategies reflecting changing environmental conditions and changes in material culture (Ringhoff and Stoner 2011: 6, Thomas 2020: 18).

The Washoe tribe occupies the Sierra Nevada region surrounding Lake Tahoe, however modern day Washoe communities exist outside of the LTBMU, primarily near Reno and Carson City. The Sierra Nevada region has three major zones for plant and animal resources (d'Azevedo 1986: 466-467). These three zones include the Boreal zone, the Transition zone, and the Upper Sonoran zone. Only the Boreal Zone is located within the LTBMU. The Boreal Zone ranges from 6,000 to 10,000 feet and contains forests, lakes, streams, and meadows where large and small game and fish could be exploited (d'Azevedo 1986: 467). Vegetation within the Boreal zone includes "pinon-juniper, yellow pine, fir and hemlock cover, and microthermal and mesothermal climates" (Elston 1971: 4). The study area is located entirely within the Boreal zone.

Prehistoric Overview

Relative and absolute dating techniques, including the Law of Superposition, projectile point chronologies, radiocarbon dating, and paleoclimatic data provide evidence for diverse cultural occupations in the Sierra Nevada (Ringhoff and Stoner 2011: 28; Lindström et al. 2021). Archaeological data, landscape and ecology, and ethnography inform about the changing lifeways and adaptations made by the Washoe people. Washoe territory was bounded in the east by Reno and Carson City in Nevada, continued north to Honey Lake in California, and extended west to Mt. Whitney, the highest point in the Sierra Nevada, and south to Antelope Valley (Elston 1971). Elston et al. (1994) identified four commonly used prehistoric cultural periods for the northern Sierra and western Great Basin: Prearchaic (12,000-8,000 BP), Early Archaic (8,000-4,000), Middle Archaic (4,000-1,300 BP), and Late Archaic (1,300-150 BP). According to Ringhoff and Stoner (2011), these cultural periods correspond to environmental periods and climatic events including the Early Holocene (10,500-7,500 BP), Middle Holocene (7,500-3,500 BP), and the Late Holocene (3,500-Present). Additionally, within each of the prehistoric cultural periods, archaeologists have identified phases within each period. These include the Tahoe Reach (>10,000-8,000 BP), Spooner (8,000-4,000), Early Martis (4,000-3,000 BP), Late Martis (3,000-1,300 BP), Early Kings Beach (1,300-700 BP), and Late Kings Beach (700-150 BP) (Duke et al. 2010).

Archaeological reconstructions of prehistoric lifeways for the Prearchaic, Early Archaic, Middle Archaic, and Late Archaic in the western Great Basin and Tahoe Basin have been summarized by various academic and cultural resource management sources (Heizer and Elasser 1953; Elston 1971; Davis, Elston, and Townsend 1973; Elston et al. 1977; d'Azevedo 1986; Elston et al. 1994; Duke et al. 2010; Ringhoff and Stoner 2011; Ascent Environmental et al. 2014; Buitenhuys 2020; Lindström et al. 2021; Lindström and Blom 2021). Generally, archaeological evidence suggests inhabitants shifted from hunting-based societies to more diverse subsistence strategies, possibly because of changing environments and subsistence opportunities, migration patterns, and population pressure.

Prearchaic Period (Tahoe Reach Phase)

The Prearchaic period is characterized by short occupations at sites that did not result in substantial artifacts or ecofacts. People hunted big and small game and gathered quickly accessible plants that did not involve processing or grinding, two features of later periods. Towards the end of the Prearchaic period and the end of the last ice age, the climate was warmer and more arid as glaciers retreated and pluvial lakes dried up. Artifact assemblages associated with the Prearchaic period Tahoe Reach cultural phase typically consisted of large, stemmed projectile points, and bifacial and unifacial tools often made from basalt or metamorphic rocks (Duke et al. 2010, Ascent Environmental et al. 2014). Land use was focused near lakes and drainages, with occupations represented in the areas of Squaw Valley, Martis Valley, Truckee, Long Valley, and Truckee Meadows (Buitenhuys 2020). Overall, archaeological evidence for the Prearchaic period in the LTBMU is sparse (Elston 1971:135).

Early Archaic Period (Spooner Phase)

The Early Archaic retains similarities to the Prearchaic, such as focusing on large game for subsistence. Diagnostic artifacts marking the Spooner Phase of the Early Archaic period include Humboldt Concave-based and Pinto projectile point types (Elston 1971; Buitenhuys 2020:10). Overall, there exists a dearth of archaeological evidence for the Spooner Phase, especially compared to later cultural phases. Populations were low, and sites were primarily on valley floors near permanent streams. Occasionally, archaeological sites from the Early Archaic have been found in upland areas. The climate continued its warming phase between 9,000 to 5,000 BP during the altithermal period, and “droughts of varying length and intensity” occurred (Ringhoff and Stoner 2011:7). With the return of a cooler and wetter period and subsequent rising water levels, sites from the Spooner Phase may have been inundated resulting in a sparse archaeological record (Lindström and Blom 2021:13). Spooner Phase has been described as sporadic and inferred by the use of the mano and metate (Elston 1971:135-136).

Middle Archaic Period

The Middle Archaic period featured a cooler and moister climate than during the Early Archaic period as the altithermal ended. This led to the expansion of forests and woodlands, as well as rising water levels in lakes and rivers (Ascent Environmental et al. 2014: 9, Lindström et al. 2021:18). Rising water levels and the likely development of more marshes and lakes would have resulted in increased waterfowl and fish populations, attracting hunters looking to exploit subsistence opportunities (Duke et al. 2010: 6). Archaeological evidence supports the assertion that hunting became even more critical for subsistence; the hunters utilized atlatls and dart points to exploit the valley margins for game.

Base camps from which people operated were typically located at the confluence of multiple resource locales, including quarries and hot springs, while simultaneously utilizing "winter pit houses with hearths and cache pits" (Ringhoff and Stoner 2011: 8). In addition to more diverse house structures, increased cultural complexity during the Middle Archaic is supported by evidence of craft specialization, production of perishable goods such as textiles, a variety of projectile points morphologies, and trans-Sierran trade of toolstones, especially basalt (Ascent Environmental et al. 2014: 10). Because Washoe territory encompassed all routes between the Great Basin and California through the Sierra, the Washoe were "centrally located for trade" (Kleam 2019: 26).

Populations increased during the Middle Archaic period more than any other time. The Martis Complex represents "the most intensive period of prehistoric occupation in the region" (Lindström et al. 2021: 18). Settlements would have corresponded with seasonal fluctuations, with occupations in the uplands during the spring and summer, and movements to lower elevations during the fall and winter (Ascent Environmental et al. 2014: 11). This type of settlement pattern would have afforded the greatest flexibility between hunting and seed collecting. Midden deposits near hot springs and buried single houses dating to the Early Martis also represent the earliest known occupations of the high Sierra, unlike the more transient deposits associated with the Tahoe Reach (Buitenhuys 2020: 11). Most Martis Complex sites have been discovered in the Transition zone, between 6,000 and 2,500 feet in elevation (Ascent Environmental et al. 2014: 11). However, the high Sierra within the Boreal zone of the LTBMU remains an opportunity to identify previously undiscovered evidence of Middle Archaic upland occupation.

Early Martis Phase (4,000-3,000 BP) and Late Martis Phase (3,000-1,300 BP)

The Early Martis is demonstrated by the intensive use of permanent camp sites and villages and dense deposits of artifacts, associated with higher population densities and sedentism (Davis, Elston, and Townsend 1974: 18). The Early and Late Martis are marked by the use of basalt artifacts and corner-notched points, including diagnostic projectile points such as Martis Contracting Stem, Martis Split Stem, and Elko series (Elson 1971; Wachter and Bloomer 2009: 7, Buitenhuys 2020: 11). Subsistence evidence indicates a variety of dietary sources including large and small game, as well as seed processing. Archaeological observations have revealed boatstones (atlatl weights), scrapers, drills, punches, and millingstone equipment such as handstones, slabs, mortars, and pestles (Duke et al. 2010: 7; Ascent Environmental et al. 2014: 10, Buitenhuys 2020: 11). The Late Martis differs little from the Early Martis, except that new forms of projectile points are seen in artifact assemblages, including the Martis Corner-Notched, Elko Corner-Notched, and Elko Eared styles (Elston 1971:137; Buitenhuys 2020:11).

Late Archaic Period

The Late Archaic period, comprised of the Early Kings Beach and Late Kings Beach phases, makes up Kings Beach Complex in the the northern Sierra. The Late Archaic corresponds to a warmer, drier climate beginning around 2,000 years ago in which lake levels in the western Great Basin and California subsided (Ascent Environmental et al. 2014:11). The Kings Beach Complex is distinguished from earlier phases as a result of technological innovation, including the introduction of the bow and arrow associated with the Rose Spring and Eastgate projectile point types, changing subsistence strategies related to the introduction of the bedrock mortar, and slightly altered settlement patterns (Elston 1971:138; Duke et al. 2010:7).

Early Kings Beach Phase (1,300-700 BP) and Late Kings Beach Phase (700-150 BP)

Rather than exploiting large game utilizing the atlatl spear thrower and complex stone tool technology, Kings Beach Complex archaeological sites indicate people focused on smaller game by using new bow and arrow technology and simpler stone tools (d'Azevedo 1986:466). With the emergence of the bow and arrow, projectile point morphology changed, including the appearance of the Rosegate and Gunther series in the Early Kings Beach, while the Late Kings Beach saw the arrival of Desert series projectile points (Elston 1971:139; Buitenhuys 2020:11). Other types of tools, such as scrapers, drills, and battered cobbles became less common during these phases (Buitenhuys 2020:11). Additionally, material types changed as chert and obsidian became relatively more common in artifact assemblages, indicating a decrease in reliance upon basalt as a toolstone compared to the Middle Archaic (Duke et al. 2010:7, Buitenhuys 2020:11).

Fishing and seed processing became significant sources of subsistence as more diverse resources were exploited in a greater number of ecological zones and with greater intensity (d'Azevedo 1986:466, Ringhoff and Stoner 2011:43). Sites dated to the Early Kings Beach show greater numbers of pestles and bedrock mortar features, reflecting the intensification of seed and nut processing. This trait persisted into the Late Kings Beach. However, archaeological evidence suggests permanent base camps and winter villages were not used as regularly or intensively during the Late Archaic (Elston et al. 1977:17). This change towards a more restricted settlement pattern could have occurred as a result of an extended dry period that reduced the availability of resources, and resulted in lower overall regional population carrying capacities (Elston et al. 1977:18). The Late Archaic period lasted until shortly after European-American contact. Many Late Archaic archaeological inferences correspond with Washoe ethnography (Elston 1971:138).

Washoe Ethnography

Detailed accounts of Washoe living conditions, preferences, and oral tradition are available from tribal sources, as well the ethnographic information that anthropologists recorded, primarily in the early part of the twentieth century (Washoe Tribe 2023; d'Azevedo 1986; Downs 1966). This information establishes a baseline that archaeologists use as a frame of reference to contextualize information gathered from archaeological investigations (Ringhoff and Stoner 2011:31). However, ethnographic information cannot alone establish facts about people that lived thousands of years ago. Culture is not static, and the knowledge of today cannot project indefinitely into the past. Ethnographic analogy alone is not suitable to provide a behavioral framework because paleoenvironments, landscapes, and resources differed so dramatically from the ethnographic present (Elston and Zeanah 2002:105). Despite the limitations of ethnographic information, it can help illuminate understanding of the past and informs about the plausibility of trends in subsistence and settlement patterns.

Lake Tahoe is the heart of the Washoe people, geographically, spiritually, and culturally. The word Tahoe is a mispronunciation of the Washoe word for lake, *Da ow* (Washoe Tribe 2023:5). The lake contains the remains of the nest of Ong, a gigantic monster bird who once terrorized the Washoe (Downs 1966:1-2; Washoe Tribe 2023:10-11). Each lake, river, creek, stand of trees, and rock outcrop is not only named, but also a location associated with a creation story, a shaman's deeds, or the mischievous activities of the long and the short tailed weasels (Down 1966:2). Despite a series of adaptations necessitated by white intrusion, the Washoe have retained their identity.

Families are the core of Washoe, usually numbering between five to 12 people who worked together and relied on each other (Washoe Tribe 2023:6). Winter camps were generally composed of four to 10 family groups. These groups moved together in a seasonal migration pattern, adapting their lives to the changing environment. Washoe societal groups were differentiated into three broader settlement groups, located north, east, and south of Lake Tahoe. In the spring, these groups would migrate from the lowlands to the shores of Lake Tahoe (Washoe Tribe 2023:6).

Fishing was a primary subsistence activity in the spring, taking place along creeks and streams during spawning runs to Lake Tahoe (Downs 1966:16). Additionally, plant and vegetable resources were available and widely disbursed, but often only in small amounts, especially in spring. Various resources were available from spring through summer, including strawberries and wild onions at higher elevations. Many people traveled to the high country to gather as much as possible for the winter (Downs 1966:18). Fishing remained critical to subsistence until the end of summer when people would begin to travel to the valleys east of the Sierra (Down 1966:19).

Based on Washoe ethnographic observations, archaeological evidence of prehistoric occupations could be expected around the shores of Lake Tahoe, where larger groups of Washoe would meet annually. Additionally, the streams and creeks in the Lake Tahoe Basin would have a higher probability for archaeological remains based on use associated with fishing, which constituted a critical subsistence activity from spring until summer. Subsistence activities such as gathering high-elevation strawberries and wild onions, as well as additional roots, vegetables, and plants comprising the Washoe diet, would also be represented in the archaeological evidence. Effective GIS predictive models should reflect these use cases during the Middle and Late Archaic periods, detailing the cultural continuity of the Washoe people in the region since the beginning of time.

Chapter 4: Methods to the Madness

Research Design

This thesis research uses spatial modeling tools to identify explanatory variables associated with a dependent variable, the location of Middle Archaic and Late Archaic archaeological sites in the Lake Tahoe Basin Management Unit. Physical characteristics of the land should influence decision-making related to prehistoric settlement, occupation, and resource exploitation locales, reflected by the deposition of archaeological evidence. The physical features influencing site patterning include elevation, slope, aspect, proximity to streams, soil and geology type, and land cover.

ArcGIS Pro software includes tools to assess explanatory variables relative to a predicted dependent variable. The tools explored by this research include logistic expressions of Generalized Linear Regression (GLR), Random Forests (RF), and Maximum Entropy (MaxEnt).

Generalized Linear Regression

Regression analysis is a standard statistical approach involving evaluating relationships between two or more features to predict an occurrence (Esri 2023a). The Generalized Linear Regression tool in ArcGIS Pro allows for modeling relationships and graphical representation of the prediction output. The tool involves the selection of Input Features, consisting of site and non-site polygons. Non-site polygons were randomly generated within the study area and used for all models. Binary classification models predicting archaeological presence and absence can be created by selecting Binary (Logistic) as the Model Type parameter of the Generalized Linear Regression tool in ArcGIS.

Regression models require presence and absence values for the dependent variable, unlike a tool such as MaxEnt , which only requires presence features to model prediction effectively. Explanatory variable values such as elevation, slope, aspect, distance to stream, soils, geology, and land cover were assigned to each Input Feature. Non-sites, the absence features representing part of the dependent variable, were randomly generated. The size of non-site features was based on the average size of all archaeological sites within the Middle Archaic and Late Archaic datasets.

The values for each input feature were calculated as an average using the Zonal Statistics tool. The Create Fishnet tool was used to divide the study area into a 100-meter by 100-meter grid, and dependent variable values were assigned to each feature in the grid. Categorical or binary predictions based on the Input Features were then generated as output predictions using the fishnet grid across the entire study area.

Random Forests

The Forest-based Classification and Regression tool is an adaptation of the random forest algorithm developed by Leo Breiman and Adele Cutler that is used in ArcGIS Pro (Esri 2023b). The parameters include Input Training Features, from which a Variable to Predict is chosen. Input Training Features consist of the site and non-site polygons used for all models, including the known archaeological sites and randomly generated non-sites used to train the model. The Variable to Predict is categorical, consisting of a binary value representing the presence or absence of an archaeological site. Predictions are based on Explanatory Training Rasters, which are the explanatory variables described later in this chapter, including elevation, slope, aspect, distance to streams, soils, geology, and land cover. Predictions were assigned to the same fishnet grid used to generate GLR predictions. To test the RF models, 30% of calibration data was withheld as validation data, resulting in a validation accuracy score for each Middle and Late Archaic RF model.

Maximum Entropy

The Presence-only Prediction (MaxEnt) tool in ArcGIS consists of parameters for Input Point Features and Explanatory Training Rasters. Input Point Features are derived from the dependent variable, the centroid of archaeological site boundaries. Explanatory Training Rasters include the explanatory variables used for this study. Given a set of known observations, such as the location of previously recorded archaeological sites, the tool allows for predictions about the location of archaeological sites that have not been previously identified, without the need to create randomly generated absence points or polygons representing non-sites. The tool generated prediction results in the study area in raster format without the use of the fishnet grid. The MaxEnt tool performed cross-validation by setting the Resampling Scheme parameter to Random, resulting in k-folds cross-validation and a subsequent validation accuracy score for each model run. Validation accuracy for Middle and Late Archaic MaxEnt models was assessed by averaging the percentage of correctly classified presence locations for each validation model run.

Archaeological Site Data

The California Historical Resources Information System (CHRIS) and Nevada Cultural Resource Information System (NVCRIS) are state repositories for archaeological data. Records searches were requested from the North Central Information Center (NCIC), a division of CHRIS that holds cultural resource investigation reports and resource records for El Dorado and Placer Counties. NVCRIS is an online system that allows the selection and download of cultural resource investigation reports and resource records for the entire state. A digitized boundary of the LTBMU was used to select reports and resources within the study area.

After receiving resource records within the study area, only those sites assigned a temporal period based on projectile point typology were selected for this study. According to David Hurst Thomas (2013: 133), typological analysis of projectile points "remains critical to our understanding of the archaeological record, particularly the interrelationship of paleoclimatic and human behavioral evidence." Based on a literature review (Elston 1971; Leventhal 1977; Waechter and Bloomer 2009), artifact typologies were ascribed to either Middle Archaic or Late Archaic periods. Site records not exhibiting sufficient resource documentation to designate a particular projectile point type were excluded. In particularly difficult cases, William Bloomer was consulted to provide expertise regarding Middle Archaic and Late Archaic projectile point typologies.

Archaeological sites that contained both a Middle Archaic and a Late Archaic component were used to train both the Middle Archaic prediction models and the Late Archaic models. Isolated occurrences of projectile points were excluded from the study because isolates do not represent evidence of settlement or occupation strategies in a discrete area, but only represent transient human behavior or transposed secondary contexts due to erosion.

Table 1: Diagnostic projectile points of the Middle and Late Archaic

Temporal Period	**Projectile Point Types**
Middle Archaic *Early and Late Martis Phases* 4,000 – 1,300 BP	Martis series, Elko series
Late Archaic *Early and Late Kings Beach Phases* 1,300 – 150 BP	Desert Side Notch series, Cottonwood, Rose Spring

Explanatory Variables

Explanatory variables consist of observations that influence the prediction of a dependent or response variable, such as the existence of an archaeological site in a defined area. Explanatory variables included in this study are elevation, slope and aspect, distance to streams, soils, geology, and land cover (**Table 2**).

Table 2: Explanatory variable data sources

Variable	**Data Source**	**Attribute**
Elevation	Tahoe Regional Planning Agency DEM	Raster value in meters
Slope	DEM	Derived using Slope tool in ArcGIS Pro
Aspect	DEM	Derived using Aspect tool in ArcGIS Pro
Distance to Streams	USGS NHD Streams Dataset	100-foot stream vector buffer converted to a raster.
Soils	USGS NRCS SSURGO	Map Unit Key ("MUKEY") attribute
Geology	USDA NRCS Geospatial Date Gateway	Geologic Unit ("ORIG_LABEL") attribute
Land Cover	USGS EROS NLCD	Land Cover Class ("NLCD_Land_Cover_Class") attribute

Elevation

The digital elevation model (DEM) was obtained from Mason Bindl, GIS Analyst with the Tahoe Regional Planning Agency. This dataset was acquired via lidar remote sensing and has a spatial resolution of 0.5 meters. The DEM represents the topographic surface of the earth free of vegetation, buildings, and other surface objects (USGS n.d.). Relative elevation may be important to settlement and occupation site selection due to opportunities for viewshed and site defensibility or may be related to specific resource locales (GeoEngineers 2009: 10)

Slope

The study area's slope was calculated from the lidar DEM using ArcGIS Pro tools. Slope is a measure of steepness, or rise over run, relative from one cell to a neighboring cell. Generally, archaeological preservation is better on shallower slopes due to increased potential for deposition due to downslope weathering on steep slopes (GeoEngineers 2009: 12). Additionally, more desirable habitation locations are located on flatter slopes, although steeper slopes may be utilized for specific resource collection opportunities, such as lithic quarries or plant and herb gathering (GeoEngineers 2009: 12). Slope was calculated in ArcGIS Pro as a percentage rather than in degrees.

Aspect

Aspect is a measure of the steepest downslope direction at a discrete surface location and is derived from each cell of the surface of a raster, such as a DEM. Aspect in the study area was calculated using the DEM in ArcGIS Pro. Aspect represents solar illumination, as south-facing slopes receive more sun exposure than north-facing slopes. Sun exposure and shade may affect vegetation communities, influencing the type of plants or herbs that grow in a particular area and that could affect animal or human behavior. Additionally, people may select habitation sites preferentially by the relative amount of sun exposure. Aspect is measured in degrees between 0-360, with 0 and 360 degrees representing true north, while flat slopes have no direction and are given a value of -1.

Distance to Streams

The USGS maintains the National Hydrography Dataset (NHD), representing a water drainage network including rivers, streams, lakes, and other water features. This comprehensive dataset is maintained regularly and is mapped at a scale of 1:24,000. In consultation with Darrel Cruz, former Tribal Historic Preservation Officer for the Washoe Tribe of Nevada and California, all areas within 100 feet of streams throughout the study area were identified as preferential for settlement, occupation, and resource locales. From the NHD dataset clipped to the study area, all stream networks were selected, and a 100-foot buffer was applied to them, resulting in a dataset that can be utilized to identify the relative importance of proximity to streams during spatial modeling processes relative to the location of known and predicted archaeological sites. The explanatory variable is the raster created representing all cells within 100 feet of streams, and all cells outside of the 100 foot buffer.

Soils

The United States Department of Agriculture (USDA) Natural Resource Conservation Service (NRCS) maintains a Soil Survey Geographic (SSURGO) Database product used for "national, regional, and statewide resource planning and analysis of soils data" (USDA n.d.). This data is the most comprehensive and detailed soil data developed by the National Cooperative Soils Survey (NCSS). Soil data is available for the entirety of the study area. The development of soils depends upon conditions such as climate, the effects of humans and other organisms, time, and the origins of the parent material (GeoEngineers 2009: 29). Soil distributions may influence models of archaeological potential (GeoEngineers 2009: 29).

Geology

The USDA NRCS also maintains a Geospatial Data Gateway that contains a library of high-resolution vector and raster datasets for environmental and natural resources data, including geology (USDA 2022). Geologic datasets were downloaded for Nevada and California, then merged and clipped to the study area in the ArcGIS Pro environment. Geologic data can provide archaeological insight regarding tool stone sources, with implications for preferred lithic procurement areas that may have influenced lithic procurement activities and settlement strategies, with implications for potential local and regional trade networks (GeoEngineers 2009: 16). Additionally, geologic types may affect rates of weathering and erosion, potentially impacting archaeological preservation (GeoEngineers 2009: 16).

Land Cover

The USGS Earth Resources and Observation Science (EROS) Center is responsible for collecting and maintaining the National Land Cover Database (NLCD), which is “the definitive land cover database for the United States” (USGS 2018). This product provides coverage for the entire study area. Land classes within the study area include Evergreen Forest, Shrub/Scrub, Woody Wetlands, Emergent Herbaceous Wetlands, Herbaceous, Barren Land, Hay/Pasture, Open Water, and Developed. While some land characteristics may have changed since the Middle Archaic and Late Archaic periods, and classes that exist because of modern activity such as Developed and Hay/Pasture did not influence where prehistoric people made settlement decisions, land cover may be an important variable in understanding general patterns of settlement, occupation, and activity locales. Additionally, land cover may contribute to the effectiveness of archaeological pedestrian surveys in identifying prehistoric sites.

Chapter 5: Dense and Dry

Chapter Outlook

This results chapter will present diagnostic metrics returned from GLR, RF, and MaxEnt Middle Archaic and Late Archaic models. The metrics include rates of predicted archaeological presence and absence, values for model fit and performance, and measures of variable explanatory power. While some diagnostic metrics differ in name between each model, direct comparisons can be made between the models by assessing validation accuracy. This chapter details results from calibration datasets that are used to train models and validation datasets that are used to test the models. This chapter consists of dense and dry statistics, and the subsequent analysis chapter will evaluate the diagnostic metrics introduced in this results chapter.

Generalized Linear Regression

The Generalized Linear Regression tool calculated potential presence and absence prediction locations based on calibration groups for the Middle Archaic and Late Archaic periods. The Middle Archaic dataset, consisting of 40 archaeological site polygons and 40 non-site polygons, and the Late Archaic dataset, consisting of 25 archaeological site polygons and 25 non-site polygons, were each split into calibration and validation data groups. Seventy percent (%) of each dataset was randomly selected for a calibration group, and the remaining 30% was selected for a validation group. Categorical GLR models output results in a binary classification. A binary output indicates only a yes or no predictive value for each location.

Middle Archaic – Binary GLR Model

Out of 86,291 100-square-meter cells making up the study area, the Middle Archaic GLR model (**Figure 2**) predicted 51,811 absence locations, or 60% of the study area, while 34,480 cells (40%) were identified as potential Middle Archaic presence locations (**Table 3**). Of the 13 archaeological sites that were part of the validation data, 11 sites were located within cells that were identified as potential Middle Archaic presence locations, while two sites were within absence locations. Of the 11 randomly determined non-site polygons, five were located within or mostly within potential presence locations, while six were located entirely within absence locations (**Table 4**). The Deviance Explained metric, a percentage of the dependent model variance explained by the explanatory variables, was 27.1%, and the Akaike's Information Criterion (AICc) metric, a relative measure of model fit and performance, was 75.5 (**Table 5**). Lower AICc values indicate better model performance.

Table 3: GLR predicted archaeological presence and absence

Period	Presence Cells	Presence Percent of Study Area	Absence Cells	Absence Percent of Study Area
Middle Archaic	34,480	40%	51,811	60%
Late Archaic	33,661	39%	52,630	61%

Table 4: GLR validation data results for the Middle Archaic

Category	Predicted Presence Locations	Predicted Absence Locations	Accuracy
Sites (n=13)	11	2	70.8%
Non-Sites (n=11)	5	6	

Table 5: GLR performance metrics for the Middle Archaic

Metric	Measure
Deviance Explained	27.1%
Akaike's Information Criterion (AICc)	75.5

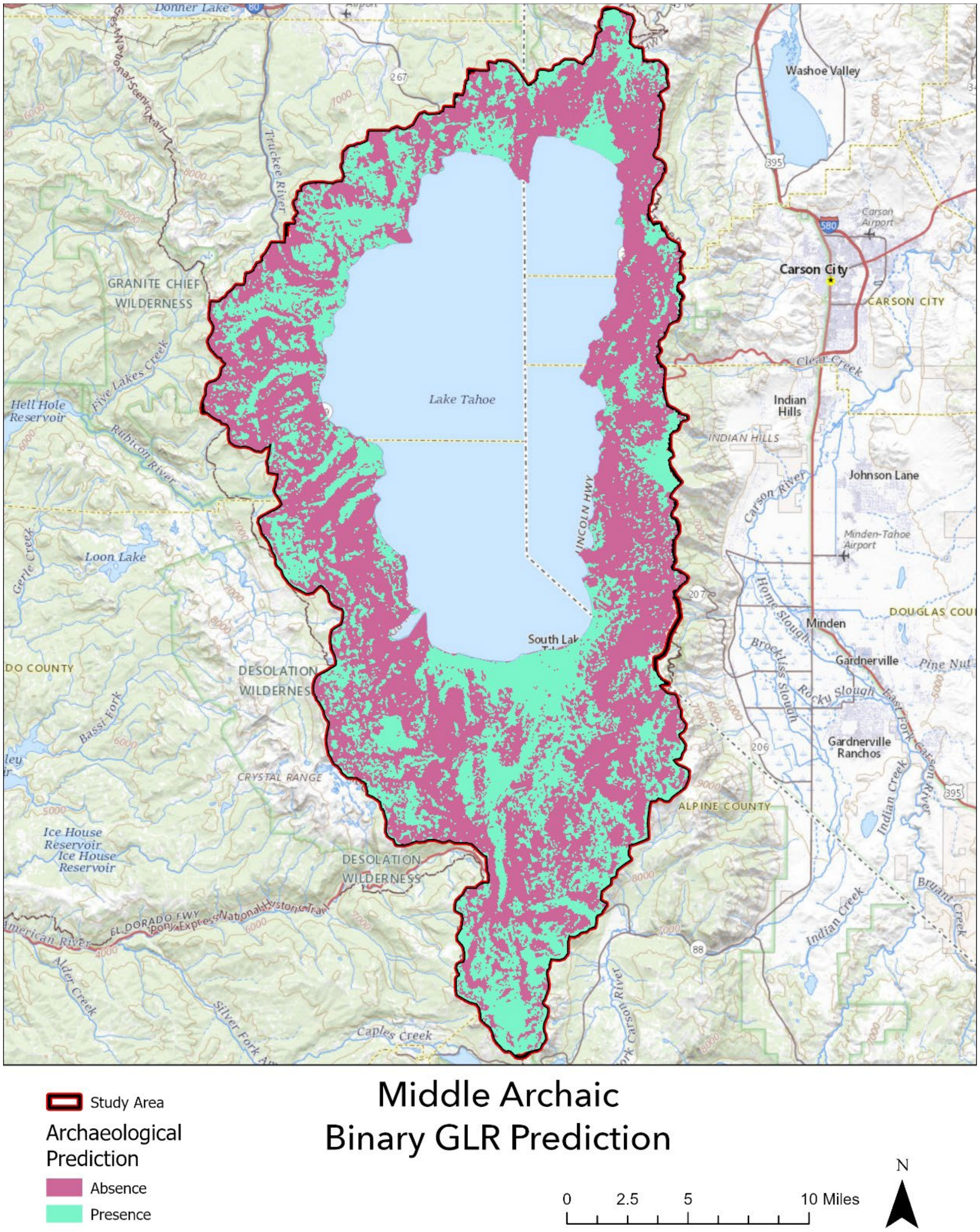

Figure 3: Middle Archaic categorical GLR prediction (map by athor)

Late Archaic – Binary GLR Model

Out of 86,291 100-square-meter cells making up the study area, the Late Archaic GLR model (**Figure 3**) predicted 52,630 absence locations, or 61% of the study area, while 33,661 cells (39%) were identified as potential Middle Archaic presence locations (**Table 3**). Of the eight archaeological sites in the validation data, all eight were located entirely or mostly within cells identified as potential Middle Archaic presence locations. Of the seven randomly determined non-site polygons, five were located within or mostly within potential presence locations, while two were located entirely within absence locations (**Table 6**). The Deviance Explained metric, a measure of the proportion of dependent model variance explained by the explanatory variables, was 58.13% and the Akaike's Information Criterion (AICc) metric, a relative measure of model fit and performance, was 41.63 (**Table 7**).

Table 6: GLR validation data results for the Late Archaic

	Predicted Presence Locations	**Predicted Absence Locations**	**Accuracy**
Sites (n=8)	8	0	
Non-Sites (n=7)	5	2	66.6%

Table 7: GLR performance metrics for the Late Archaic

Metric	**Measure**
Deviance Explained	58.13%
Akaike's Information Criterion (AICc)	41.63

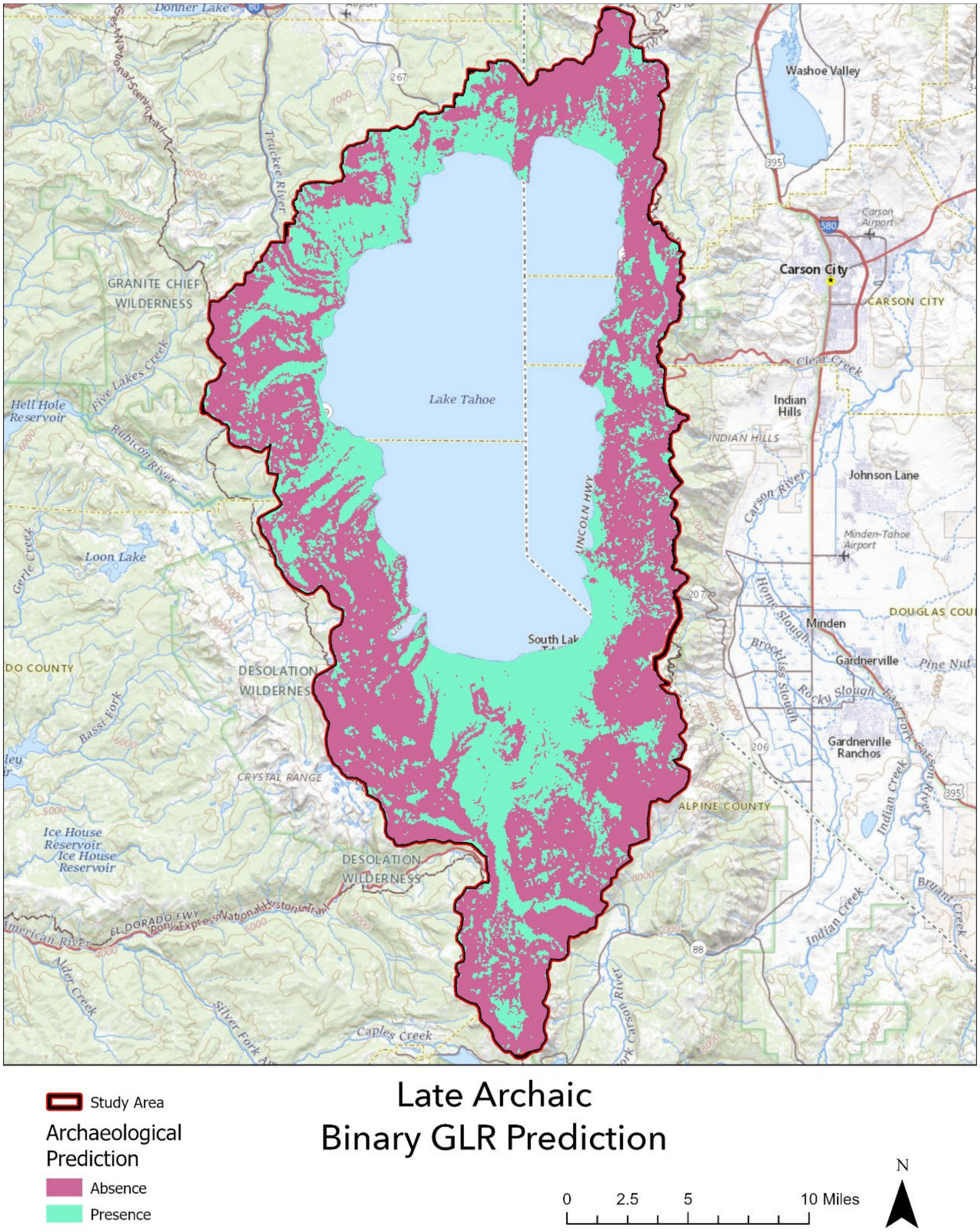

Figure 4: Late Archaic categorical GLR prediction (map by athor)

Random Forest

Model Calibration and Validation

The Forest-based Classification and Regression tool used to calibrate the prediction output was run using the same parameters for both Middle Archaic and Late Archaic datasets. The Middle Archaic dataset consisted of 40 Middle Archaic site boundary polygons, and 40 non-site polygons. The Late Archaic dataset comprised 25 Late Archaic site boundary polygons, and 25 non-site polygons. Of the polygons in either dataset, 70% of the data was used for model calibration, and 30% was used for model validation. The tool used 50,000 decision trees to generate predictions and included 100% of the available calibration data per tree. For each dataset, 10 cross-validation model runs occurred, and the calibration results were based on the model run with the highest accuracy value, indicating the best model fit determined by the tool.

Once the model was trained using the calibration data, the tool determined an accuracy score of the calibration data. Upon training the model with calibration data, predictions were made for the validation data and compared to archaeological sites' observed presence or absence, resulting in a validation accuracy score. Calibration accuracy for the Middle Archaic dataset was 95%, and validation accuracy was 62% (**Table 8**). Calibration accuracy for the Late Archaic dataset was 97%, and validation accuracy was 80% (**Table 8**).

Table 8: Random Forest calibration and validation accuracy scores

Period	Calibration Accuracy	Validation Accuracy
Middle Archaic	95%	62%
Late Archaic	97%	80%

For both Middle Archaic and Late Archaic model runs, the resulting feature class derived from the trained model was used to predict archaeological presence and absence across the entire gridded study area.

Middle Archaic Model

Out of 86,291 100-square-meter cells making up the study area, the Middle Archaic RF model (**Figure 4**) predicted 56,958 absence locations, or 66% of the study area, while 29,333 cells (34%) were identified as potential Middle Archaic presence locations (**Table 9**). The most important variables in order of importance in determining site presence were soils, elevation, slope, land cover, distance to streams, aspect, and geology (**Table 10**).

Table 9: Random Forest prediction archaeological presence and absence

Period	**Presence Cells**	**Presence Percent of Study Area**	**Absence Cells**	**Absence Percent of Study Area**
Middle Archaic	29,333	32%	56,958	68%
Late Archaic	27,603	34%	58,688	66%

Table 10: Random Forest variable importance for the Middle Archaic

Variable	**Rate of Importance**	**Percent of Importance**
Soils	0.71	15%
Elevation	0.71	15%
Slope	0.71	15%
Land Cover	0.67	14%
Distance to Streams	0.65	14%
Aspect	0.65	14%
Geology	0.58	12%

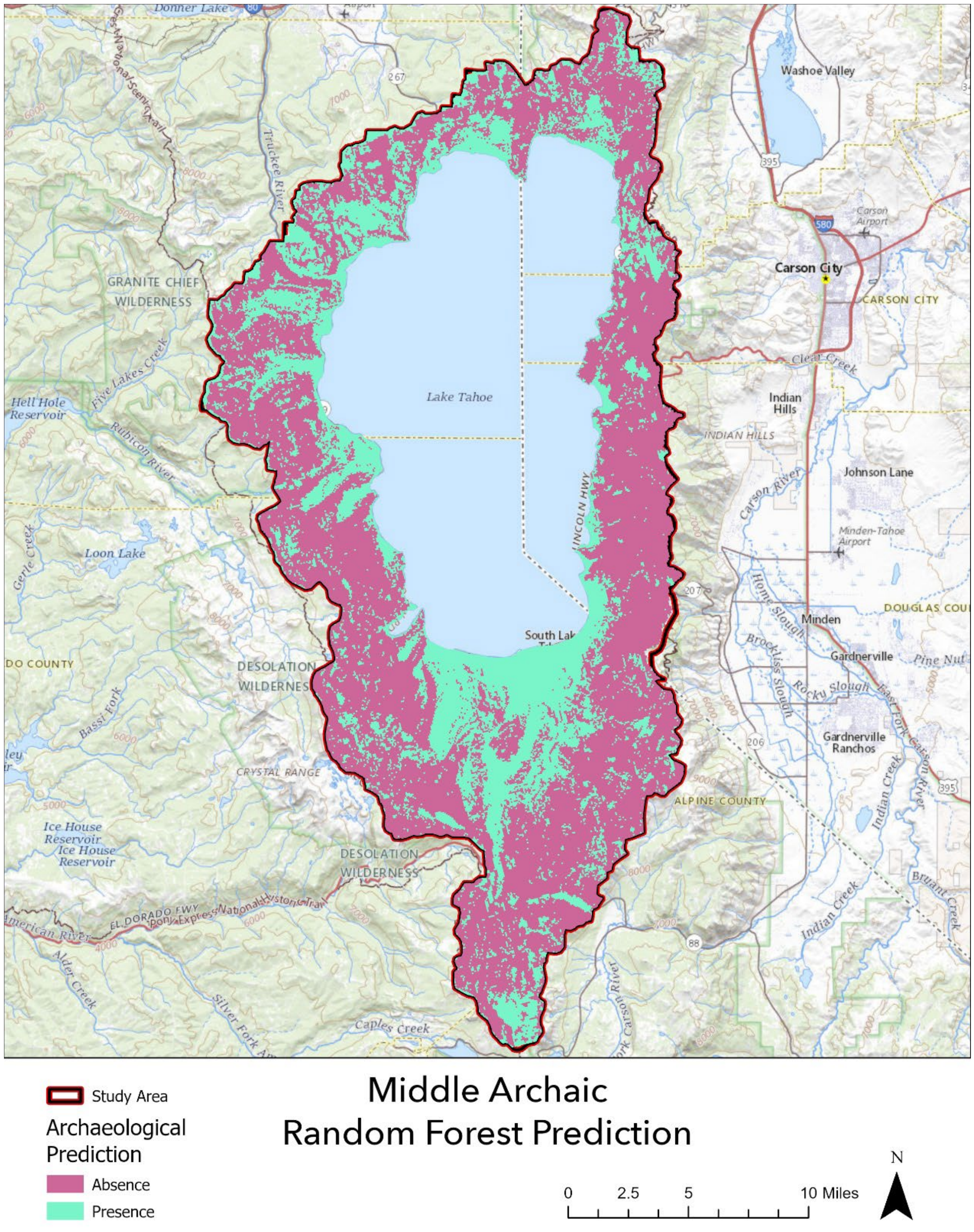

Figure 5: Middle Archaic Random Forest prediction (map by athor)

Late Archaic Model

For the Late Archaic, the RF model (**Figure 5**) predicted 58,688 absence locations, or 68% of the study area, while 27,603 cells (32%) were identified as potential Middle Archaic presence locations (**Table 9**). The most important variables in determining site presence were soils, geology, aspect, elevation, distance to streams, land cover, and slope (**Table 11**).

Table 11: Random Forest variable importance for the Late Archaic

Variable	Rate of Importance	Percent of Importance
Soils	0.39	15%
Geology	0.37	15%
Aspect	0.36	14%
Elevation	0.35	14%
Distance to Streams	0.35	14%
Land Cover	0.35	14%
Slope	0.32	13%

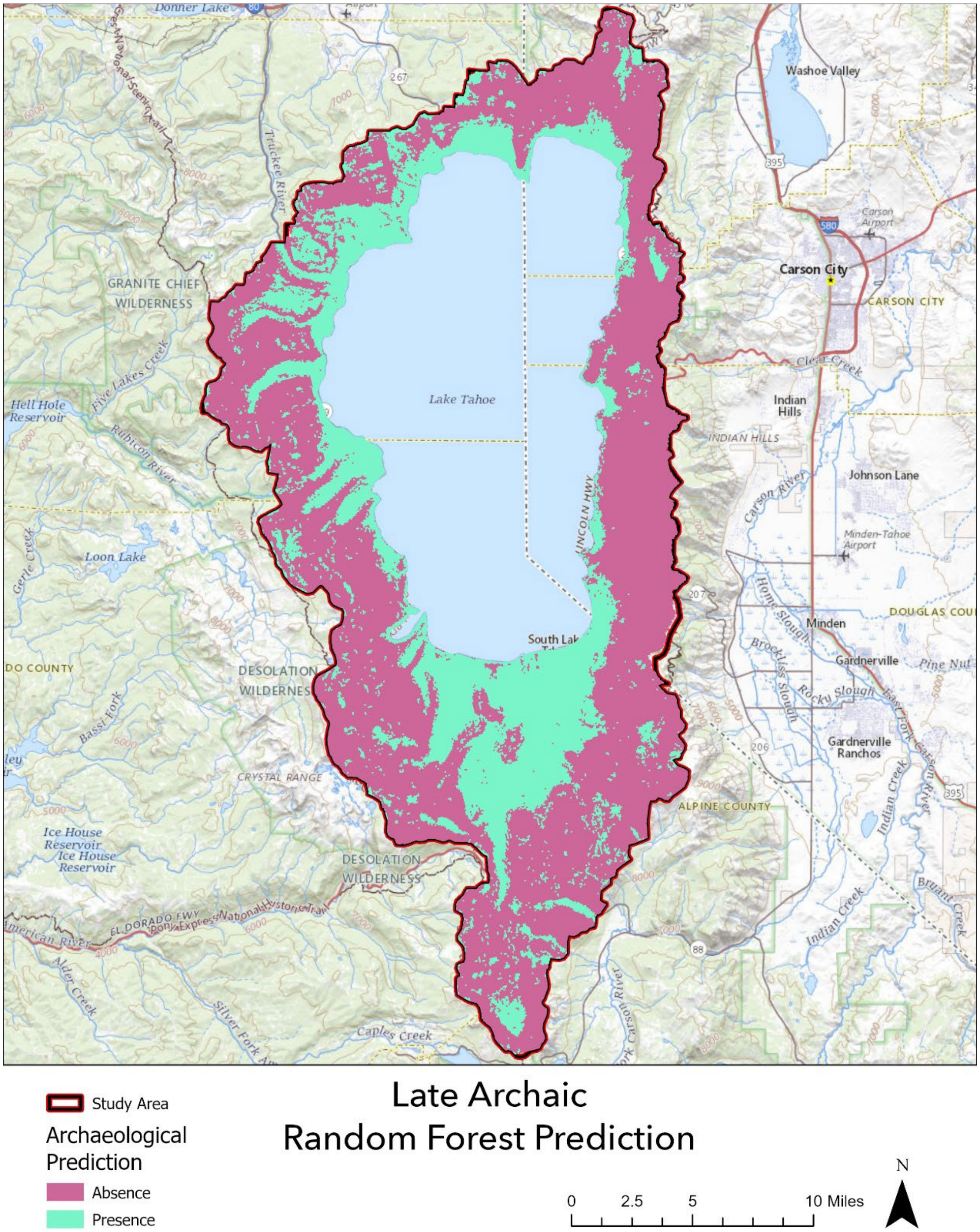

Figure 6: Late Archaic Random Forest prediction (map by athor)

Maximum Entropy

Model Calibration and Validation

The output of the Presence-only Prediction (MaxEnt) tool is a continuous value raster over the entirety of the study area. The basis functions parameter used to transform the explanatory variables for use in the Middle Archaic and Late Archaic models were Original (Linear), Squared (Quadratic), and Pairwise interaction (Product).

The Original basis function employed a linear relationship between dependent and explanatory variables and is the default option to investigate simple model relationships. The Squared basis function transforms each variable by squaring the value to represent non-linear, quadratic relationships between environmental variables and archaeological presence, serving to investigate more complex relationships.

The Pairwise interaction basis formula multiplies pairs of explanatory variables to identify interactive relationships, where archaeological presence results from the interplay of two explanatory variables. The three basis functions were employed to create a more robust model output resulting in better fit and prediction power. While the Pairwise interaction basis function is useful for modeling multiple interacting variables, the function can obscure the interpretability of explanatory variables (Esri 2023c). To interpret explanatory variable importance, the Middle Archaic and Late Archaic models were run again using only the Original basis functions, resulting in an interpretable coefficient table (**Tables 14 and 17**).

Cross-validation built into the MaxEnt tool in ArcGIS Pro includes a random resampling parameter, and when selected, results in ten resampling groups used to validate every model. Averaging the percent of known presence points correctly classified for all cross-validation runs results in a single accuracy score for each

Middle and Late Archaic model. A standard cutoff value for presence probability of 0.5 was used for both Middle Archaic and Late Archaic models. This point is a number between 0.01 and 0.99 that signifies the value at which a presence or absence point is determined and helps to evaluate the model based on a comparison of training data and known presence points. The prediction values generated by the MaxEnt models were grouped into two categories representing the likelihood of a potential presence location: 0-0.5 and 0.5-1.0.

Middle Archaic Model

The Middle Archaic MaxEnt model (**Figure 6**) resulted in an Omission Rate of 0.23, indicating that 23% of presence points were misclassified as absence points, and the AUC (area under the ROC curve), a value indicating overall model fit and performance, was 0.89 (**Table 12**). AUC values closer to 1.0 represent better prediction power, while scores closer to 0.5 represent poorer model performance. 48,524 background points were used to train the model, and 8,753 (18.0%) of those points were identified as potential presence locations, and the accuracy of the model in predicting presence at known archaeological sites using validation data was 65.8% (**Table 13**). The model employing only the Original basis function resulted in a table that produced coefficients for each dependent variable (**Table 14**). Because none of the coefficients were zero, each variable influenced the model output.

Table 12: Maximum Entropy performance metrics for the Middle Archaic

Category	Value
Omission Rate	0.23
AUC	0.89

Table 13: Maximum Entropy rate of presence for the Middle Archaic

Background Points	Presence Points	Presence Percentage	Accuracy
48,524	8,753	18.0%	65.8%

Table 14: Maximum Entropy coefficient values for the Middle Archaic

Dependent Variable	**Coefficient**
Distance from Streams	-0.2850
Geology	-0.0217
Slope	-0.0647
Land Cover	0.0362
Soils	0.0087
Aspect	0.0008
Elevation	0.0021

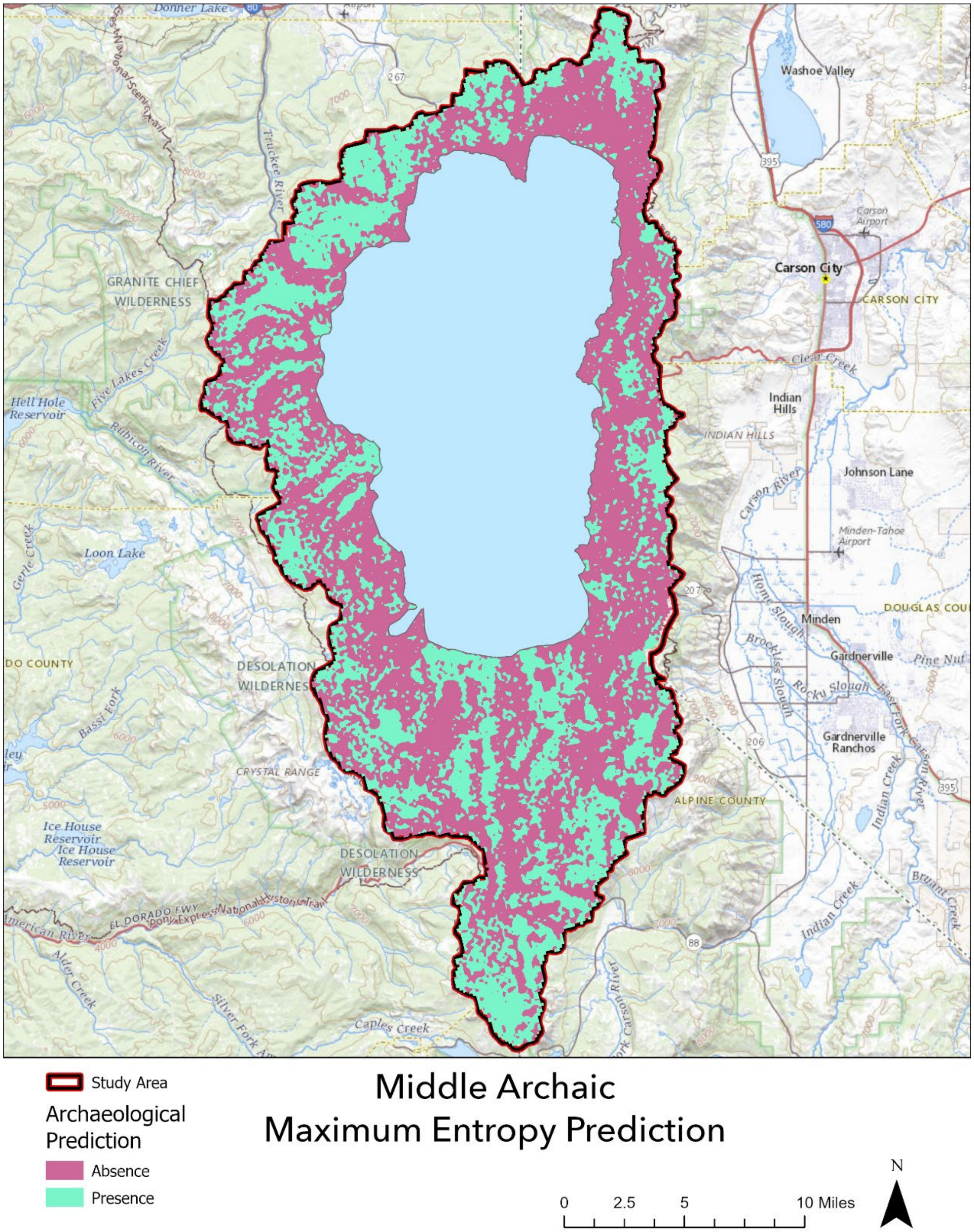

Figure 7: Middle Archaic Maximum Entropy prediction (map by athor)

Late Archaic Model

The Late Archaic MaxEnt model (**Figure 7**) resulted in an Omission Rate of 0.13, indicating that 13% of presence points were misclassified as absence points, and the AUC value was 0.90 (**Table 15**). 47,049 background points were used to train the model, and 9,389 (19.9%) of those points were identified as potential presence locations, and a 79.9% accuracy rate in predicting presence at known archaeological sites using validation data (**Table 16**). The model employing Original and Squared basis functions resulted in a table that produced coefficients for each dependent variable except Distance from Streams and Elevation (**Table 17**).

Table 15: Maximum Entropy performance metrics for the Late Archaic

Category	**Value**
Omission Rate	0.13
AUC	0.90

Table 16: Maximum Entropy rate of presence for the Late Archaic

Background Points	**Presence Points**	**Presence Percentage**	**Accuracy**
47,049	9,389	19.9%	79.9%

Table 17: Maximum Entropy coefficient values for the Late Archaic

Dependent Variable	**Coefficient**
Distance from Streams	N/A
Geology	-0.0097
Slope	-0.0202
Land Cover	0.0407
Soils	0.0142
Aspect	0.0011
Elevation	N/A

Chapter Review

This chapter presented the raw results concerning Middle and Late Archaic archaeological prediction based on GLR, RF, and MaxEnt models. These results included values for archaeological presence and absence rates, model fit and performance, and variable importance. The meaning of these diagnostic metrics will be further explained and compared between models in the next chapter. Additionally, the results of each model were presented as graphical outputs in this chapter. The maps produced by each model were presented as values of binary presence and absence. In Chapter 6, I compare the map results for different models and observe general trends in archaeological presence or absence and differences between GLR, RF, and MaxEnt map results.

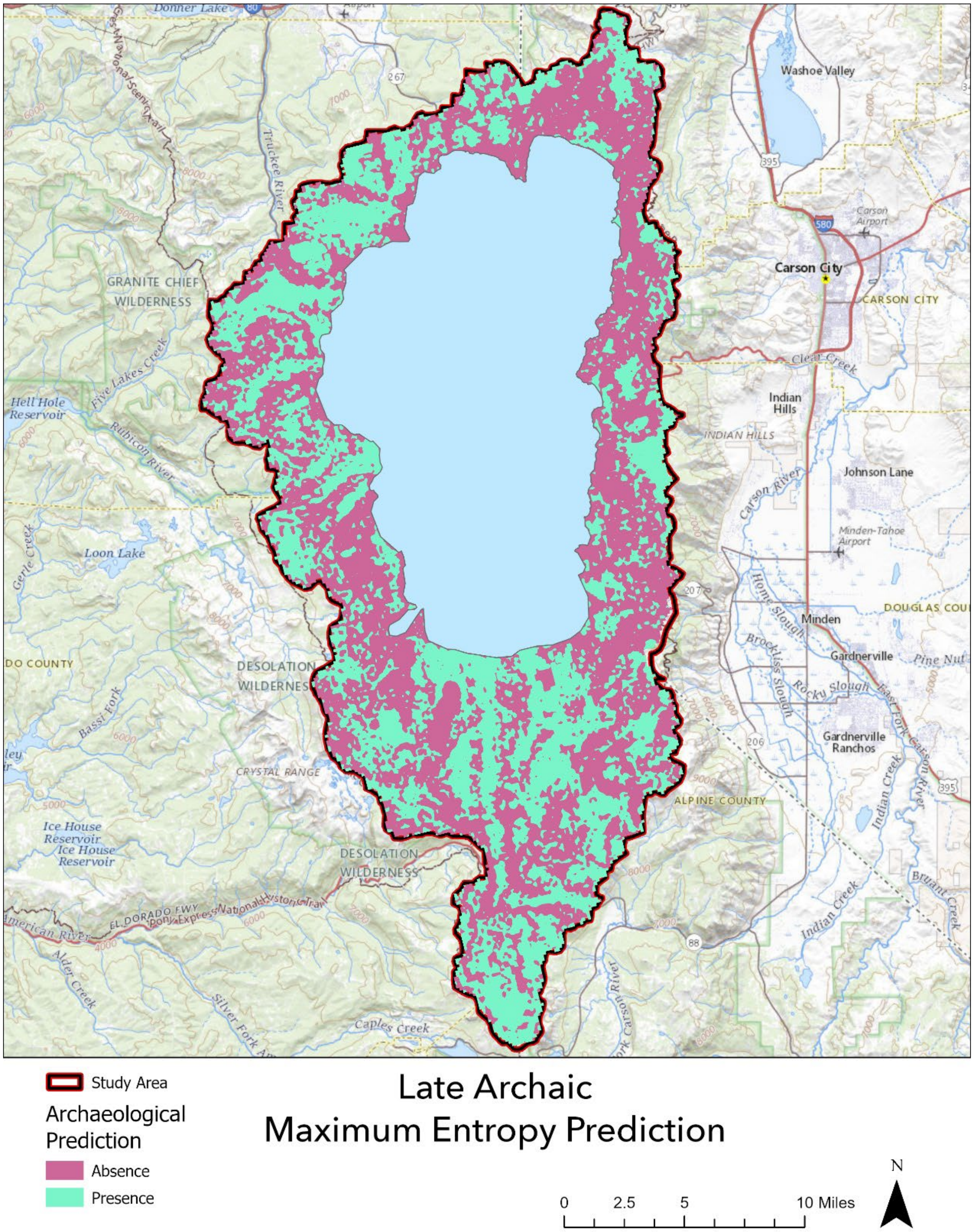

Figure 8: Late Archaic Maximum Entropy prediction (map by athor)

Chapter 6: Putting it All Together

Chapter Outlook

The previous chapter provided metrics for each model that will used in this chapter to assess the best fit for the prediction of Middle and Late Archaic sites in the Lake Tahoe Basin. Each model type results in unique evaluation metrics, but relative comparisons can be made between them. Additionally, this chapter will connect model predictions with archaeological evidence from published literature to assess the utility and significance of the model predictions.

Quantitative Analysis of Model Metrics

This section will compare prediction values, variable importance, performance, and fit metrics for the GLR, RF, and MaxEnt Middle and Late Archaic models. While the same performance metrics, except for accuracy, are not shared between GLR, RF, and MaxEnt, relative comparisons can be made in terms of goodness of fit and predictive value between the three model types, as well as within each model type by comparing results for the Middle and Late Archaic datasets. The best model will return good fit, accuracy, and predictive value for both the Middle Archaic and the Late Archaic. MaxEnt was the most robust model predicting archaeological presence based on the high accuracy of both Middle Archaic and Late Archaic model runs (**Table 18**). Random Forest was the next most-effective model based on combined validation accuracy, but had a larger discrepancy between the Middle and Late Archaic validation accuracy compared to MaxEnt models. GLR results between the Middle and Late Archaic performed similarly, but with the lowest combined validation accuracy.

Table 18: Direct comparison of model accuracy

Model Type	Period	Validation Accuracy	Combined Accuracy
GLR	Middle Archaic	70.8%	68.7%
	Late Archaic	66.6%	
RF	Middle Archaic	62.0%	71%
	Late Archaic	80.0%	
MaxEnt	Middle Archaic	65.8%	72.8%
	Late Archaic	79.9%	

GLR Performance

Between GLR, RF, and MaxEnt models, GLR was the weakest model according to output metrics. The AICc value, a measure of model fit and performance, indicated better performance of the Late Archaic model compared to the Middle Archaic model. However, in assessing validation data importance, the Middle Archaic dataset had a higher accuracy in predicting presence or absence of archaeological sites compared to the Late Archaic. While the GLR Late Archaic model AICc value was better than the Middle Archaic, the validation accuracy did not reflect better model performance of the Late Archaic model.

RF Performance

The RF models resulted in relatively high accuracy prediction power for the training and validation data. Validation data accuracy is a better measure of model prediction power because it tests the prediction accuracy of features not used to train the model and is therefore less biased. Validation accuracy for the Middle Archaic was 0.62, or 62%. This value indicates that the model correctly predicted the presence and absence of 62% of the site and non-site features not involved in training the model.

However, the Late Archaic RF model outperformed the Middle Archaic model. Validation accuracy for the Late Archaic was 0.80, or 80%. Unlike the GLR models with a relatively narrow prediction power gap and model fit between the Middle and

Late Archaic, the RF models had a wider range of prediction power according to validation accuracy scores. Based on diagnostic metrics, RF is a better tool for Late Archaic archaeological prediction than GLR in the LTBMU, but GLR produced better results using the Middle Archaic data.

Top variable importance differed between the Middle and Late Archaic within the RF model comparison. In order of importance, the variables of the greatest importance for the Middle Archaic model were soils, elevation, slope, land cover, distance to streams, aspect, and geology. All Middle Archaic RF variables had an importance percentage between 12-15%. These percentage scores indicate that all variables had a significant contribution to the output result of the model.

Variable scores for the Late Archaic RF model had a different order of importance: soils, geology, aspect, elevation, distance to streams, land cover, and slope. Once again, all importance percentages were similar, varying between 13-15%. Distance to water, slope, and soils are generally three of the most important elements of a predictive model in archaeology (Verhagen 2018; Yan et al. 2021). However, this RF analysis shows that under particular geographical settings and as culture change occurs, variable importance may change over time, and variable importance may not be restricted to the most commonly held notions of predictive power.

MaxEnt Performance

MaxEnt models for the Middle and Late Archaic performed similarly well. The metric used for evaluating MaxEnt model fit and performance is the AUC value. This value measures how good the model is at predicting known presence points as presence locations (Esri 2023d). The AUC for the Middle Archaic was 0.89, and the AUC for the Late Archaic was 0.90. Values closer to 1.0 indicate an excellent fit, while values around 0.5 means the model does not perform accurate prediction (Esri

2023d). Additionally, the combined validation accuracy of both MaxEnt models slightly outperformed the GLR and RF models. Out of all three model types tested, MaxEnt showed a good fit between both models with the least discrepancy between the Middle and Late Archaic in terms of performance metrics, indicating MaxEnt was the most effective predictor for both Middle and Late Archaic archaeological presence among GLR, RF, and MaxEnt models.

Additionally, by integrating different basis functions in the model calculation, MaxEnt can produce complex, comprehensive results for a large area with multiple explanatory variables using only presence points. One disadvantage to integrating multiple basis functions to the MaxEnt predictive models for the Middle and Late Archaic is that the explanatory variable diagnostics could have been more easily interpretable. As a solution, the MaxEnt models were run again using only the Original (Linear) basis function to provide an example of dependent variable coefficients, or the relative strength of variable relationships in determining a prediction result.

The Middle Archaic run showed that the order of most important variables included: distance to streams, slope, land cover, geology, elevation, aspect, and soils. A different order of variable importance emerged from the Late Archaic run: slope, land cover, soils, geology, and aspect. The Late Archaic model output diagnostics did not include distance to streams or elevation coefficient values. The absence of these coefficient values is likely due to a processing error, as distance to streams is an important variable in all other model runs. However, all coefficient values generated were greater than zero, indicating variables that influenced model output.

Presence-only modeling is particularly useful for archaeological predictive modeling. While GLR and RF required the use of absence locations, which involves generating a random set of points for which the actuality of archaeological site presence is not known, MaxEnt requires only the use of presence locations. This feature may help to avoid bias in model results by incorrectly training a model to predict an absence location using a randomly generated point where an archaeological site may actually exist.

Qualitative Analysis of Model Graphics (See Figures 2-7 in Chapter 5)

Generally, there are similarities between GLR, RF, and MaxEnt models in terms of Middle Archaic and Late Archaic presence locations. The presence locations are clustered around meadows, lakes, creeks and streams, and the shoreline of Lake Tahoe. These results agree with Washoe ethnographic observations: sites would be located near water used for daily maintenance and fishing such as springs, streams, or lakes, or near productive resource exploitations areas such as meadows and open forest (Davis, Elston, and Townsend 1973:5-9).

GLR, RF, and MaxEnt models showing concentrated presence locations at the shoreline of Lake Tahoe correlate with Washoe ethnography, mythology, and folklore. Because Lake Tahoe comprised the geographic and social center of the Washoe world, "almost every bay, inlet, and stream mouth has a legendary or mythological association" (Downs 1966:16-17). Every year in the spring, groups of young men and women would travel to Lake Tahoe from winter villages, living in caves and rockshelters and harvesting whitefish and plants until the rest of the tribe would join them in larger base camps later in summer (Davis, Elston, and Townsend 1973:3). Higher probabilities of presence locations around Lake Tahoe reflect the more intensive use that would have occurred based on the importance of Lake Tahoe to Washoe culture.

Late Archaic presence locations are more clustered than Middle Archaic presence locations. While they generally share similar prediction trends, the Middle

Archaic GLR, RF, and MaxEnt models show more overall presence locations scattered over a more dispersed area compared to the Late Archaic models, which show fewer and more discrete presence locations. Late Archaic occupation in the Lake Tahoe Basin is less intense and less regular compared to the Middle Archaic, a phenomenon that is probably the result of a drier climate and a decrease in regional population carrying capacity (Elston et al. 1977:2). The model results showed a more constricted pattern of occupation conform with interpretations suggesting a reduced Late Archaic population density in the Tahoe Basin.

However, the Middle and Late Archaic prediction locations are not mutually exclusive, evidence that goes against the existence of two chronologically separate patterns of land use, as was theorized by Heizer and Elasser (Heizer and Elasser 1953:19-20). The results of these models indicate some similarity and overlap between Middle and Late Archaic predicted site locations, a result supported by Robert Elston's conclusion that cultural continuity existed between the Middle Archaic Martis complex and the Late Archaic Kings Beach complex (Elston 1971:140). Additionally, some of the sites used to train the models were multicomponent sites, indicating that base camps were repeatedly occupied during the Middle Archaic and again in the Late Archaic.

Summary

This chapter has compared metrics of fit and performance for GLR, RF, and MaxEnt models for the Middle and Late Archaic. MaxEnt was found to deliver the most consistent results between the Middle and Late Archaic. Presence-only prediction models such as MaxEnt are particularly adapted to archaeological modeling because no randomly generated and potentially false absence locations are needed to deliver predictions. However, certain advantages belong to RF modeling scenarios, such as the ease of interpreting dependent variables. Overall, machine learning algorithms such as RF and MaxEnt provide utility to archaeological modeling due to their ability to produce predictions that account for complex systems over large areas with limited datasets.

This chapter has also connected output model results to Washoe ethnography and archaeological observations from the Lake Tahoe Basin. The Middle Archaic represents a period of population growth and expansion, while the Late Archaic archaeological evidence suggests lower population densities and more restricted settlement patterns. The models produced by this study correlate to those observations, reflecting wider distributions of Middle Archaic presence locations compared to Late Archaic prediction locations, which are fewer and more restricted. Therefore, temporal trends in the distribution and density of archaeological sites are supported by this analysis. Additionally, the constraining parameters dictating site choice according to ethnographic observations, such as proximity to certain types of natural resources, are supported by model results.

Chapter 7: Looking Back, and a Way Forward

The Study Area

The LTBMU is a federally designated area comprising approximately 154,000 acres of land, distributed between two states and six counties. The Lake Tahoe Basin is home to over 10,000 years of human history and remains a place of cultural and spiritual importance to the Washoe descendants that continue to inhabit the region. Archaeology, paleoecology, and ethnography help to reconstruct the lifeways of Washoe ancestors in the region, as environmental changes reflect the cultural and technological changes made by past peoples. Previous archaeological research in the region has defined periods, including the Prearchaic, Early Archaic, Middle Archaic, and Late Archaic, and cultural phases such as the Spooner, Martis, and Kings Beach complexes, that coincide with larger environmental trends during the Holocene. While much archaeological data has been recorded, the protection and preservation of cultural resources remain a priority.

Challenges and a Solution

The proliferation of large wildfires in the west presents a challenge to cultural and heritage resource management. Fire is a critical aspect of human society, and indigenous anthropogenic burning has a long history in California as a tool for landscape management (Cuthrell et al. 2012). A century of fire suppression in the United States, combined with land clearing, logging, invasive species, and human-based ignitions invites destructive fires (Pyne 2021:106). Destructive fires that burn with great intensity, as well as fire suppression activities, have the potential to alter or destroy prehistoric cultural resources, including lithics, ceramics, rock art, and ground stone (Ryan et al. 2012). While prescribed burning and principles of traditional ecological knowledge related to burning serve numerous benefits, and fire is a necessary component of the natural environment, management tools such as

GIS provide a solution to help protect cultural resources from the unwanted effects of wildfires, as well as landslides, erosion, and flooding.

Practical Use Cases

This thesis presents a targeted spatial modeling approach that utilizes statistical analysis and machine learning to identify patterns of prehistoric land use during the Middle Archaic and Late Archaic in the LTBMU. The results of this study compare different model types that are evaluated based on statistical metrics and qualitative analysis tied to the archaeological and ethnographic literature of the area. In constructing models that show areas of the archaeological site location probability with the highest accuracy, this thesis provides an effective management tool for indigenous groups, local, state, or federal management agencies, and archaeologists that wish to preempt or mitigate damage to Middle and Late Archaic prehistoric cultural resources that occur as a result of natural disaster.

Management decisions informed by spatial analysis offer an avenue for time-efficient and cost-effective strategies to assess and mitigate damage to cultural resources through survey and monitoring by prioritizing the most likely affected areas. Machine learning-based spatial modeling approaches, especially MaxEnt and RF, provide advantages in producing more accurate models compared to traditional regression. With accurate models that are effective at producing archaeological predictions across large areas, archaeological site monitoring and preservation efforts can be boosted especially in fire management scenarios.

Efficient and effective management decisions are critical when responding to protect known archaeological resources from fire damage, and during post-wildfire rehabilitation work. The ability to identify areas of higher archaeological probability when designing an Area of Potential Effects for a native reseeding project in a burned area, for example, can result in a project that maximizes benefit of not only the ecosystem, but also provides the greatest benefit to the protection and preservation of archaeological sites. Because many archaeological sites are exposed when wildfires occur, resulting in a landscape of shimmering obsidian flakes, the potential for looting increases as well. Carrying out post-wildfire vegetation rehabilitation projects in higher probability locations of archaeological site presence can help to reduce instances of looting as vegetation will grow to obscure exposed archaeological sites. Federal land management agencies such as the Bureau of Land Management, Forest Service, and National Park Service, who are responsible more managing biological and cultural resources on public lands, would benefit greatly from predictive modeling to better manage archaeological sites in post-wildfire recovery efforts.

Contribution to Theory

GIS modeling is an optimal lens to view cultural change as it relates to the principles of behavioral ecology. This analysis shows that cultural change coincided with environmental shifts as climates changed during the Early, Middle, and Late Holocene. Cultural change includes technological innovation and settlement patterns, two factors involved in this analysis. Technological innovation informs the Middle and Late Archaic datasets used to construct the models, while settlement patterns are the output of each model. As behaviors associated with land use became more effective regarding resource selection and procurement strategies, and environmental pressures such as a drier climate forced changes in behavior, restricting settlement patterns during the Late Archaic compared to the Middle Archaic. These decision-making processes are an example of patch-modeling theory in which behavior is optimized based on the seasonal and spatial distribution of resources. The cost-benefit analysis of effective survival strategies leads to changes in the sociocultural realm, resulting in material change reflected in the archaeological record. Machine-learning GIS modeling approaches are effective agents to visualize the change in choice, behavior, and material patterns.

The results of this thesis support Robert Elston's conclusion that Middle and Late Archaic sites were not mutually exclusive, as was proposed by Heizer and Elasser (Elston 1971; Heizer and Elasser 1955). Despite a more restricted settlement pattern during the Late Archaic that was likely influenced by environmental factors, significant overlap in land use occurred during the Middle and Late Archaic, suggesting cultural continuity between the two periods. The existence of sites with both a Middle Archaic and Late Archaic component provides further evidence of the continual occupation of sites. Ethnographic evidence indicated that productive locales such as those near meadows, lakes, creeks, streams, and the shoreline of Lake Tahoe were strong factors influencing land use. Based on a qualitative analysis, RF and MaxEnt models from the Middle and Late Archaic mirror the importance of those locales, predicting stronger rates of archaeological site occurrence near them. This thesis has sought to build bridges between archaeological theory, archaeological data, and ethnographic observations.

Limitations and Future Research

A limitation of this thesis is the sites that are excluded from the analysis. Only prehistoric sites with a diagnostic stone tool component that could be ascribed to the Middle or Late Archaic were selected for the GIS analyses. This exclusion was performed to create the strongest association of sites to a cultural period, limiting obfuscation of land use patterns between different periods. In this way, by not including all known sites, including those associated with different population densities and subsistence practices, the model results capture specific behavioral adaptations regarding land use and site selection. However, including only sites with a stone tool component may not effectively predict prehistoric sites and land use patterning related to fishing, plant gathering and processing, and other sites with spiritual or ritual association.

Future GIS research in the region might focus on prehistoric sites that contain ground stone, or those with rock art components for results more specifically suited to those land use questions. Alternatively, radiocarbon dating might be implemented systematically to produce more generalized outputs representing multiple types of activities and land use patterns. The methods presented in this thesis can be tested in other areas and regions to assess the utility and effectiveness of machine learning-based modeling approaches compared to linear regression, once the standard of archaeological modeling. Ideally, the map outputs produced in this thesis will be used in the field to help strategize and prioritize efforts to prevent or mitigate adverse effects to cultural and heritage resources.

Interdisciplinary predictive modeling approaches may allow for the most robust prevention and rehabilitation efforts in the context of wildfire management. For example, archaeological predictive models could be combined with predictive models that simulate the greatest burning intensity based on vegetation type, historic fire regimes, annual rates of precipitation, wind trends, slope, and aspect. Wildfire predictive models and archaeological predictive models could be overlaid to identify the combination of greatest archaeological sensitivity with highest probability of intense burning. Then, prior to the instance of an uncontrolled wildfire or a prescribed burn that may affect cultural resources, these areas could be surveyed to record site data, providing a baseline to monitor changes to cultural resources in the event of a wildfire. GIS modeling approaches leverage comprehensive data and computational power provide strategies to more effectively manage cultural resources in the context of fire management.

Bibliography

Ascent Environmental, Power Engineers, Inc., Parus Consulting, Inc., and Natural Investigations Company, LLC.

2014 *Cultural Resources Evaluation Report for CalPeco 625 and 650 Electrical Line Upgrade Project, Nevada and Placer Counties, California.* Submitted to the Lake Tahoe Basin Management Unit, South Lake Tahoe, California, Tahoe National Forest, Nevada City, California, and United States Army Corps of Engineers, Sacramento District. On file at North Central Information Center, Sacramento, California.

Broughton, Jack M., and James F. O'Connell

1999 On Evolutionary Ecology, Selectionist Archaeology, and Behavioral Archaeology. *American Antiquity* 64(1): 153-165.

Buitenhuys, Connor

2020 *Historic Property Survey Report for the West Shore Lake Tahoe Project at Meeks Creek Bridge Along State Route 89 in El Dorado County, California.* California Department of Transportation, Cultural Resources – South, Marysville, CA. On file at North Central Information Center, Sacramento, California.

California Geological Survey

2002 California Geomorphic Provinces Note 36. Electronic document, accessed December 22, 2022, https://www.conservation.ca.gov/cgs/Documents /Publications/CGS-Notes/CGS-Note-36.pdf.

Castiello, Maria Elena and Marj Tonini

2021 An Explorative Application of Random Forest Algorithm for Archaeological Predictive Modeling. A Swiss Case Study. In *Journal of Computer Applications in Archaeology* 4(1):110-125.

Codding, Brian F., and Douglas W. Bird

2015 Behavioral ecology and the future of archaeological science. *Journal of Archaeological Science* 56: 1-20.

Cuthrell, Rob Q., Chuck Striplen, Mark Hylkema, and Kent G. Lightfoot

2012 A Land of Fire: Anthropogenic Burning on the Central Coast of California. In, *Contemporary Issues in California Archaeology*, edited by Terry L. Jones and Jennifer E. Perry, pp. 153-172. Routledge, New York.

d’Azevedo, Warren L. (editor)

1986 *Handbook of the North American Indians, Volume 11: Great Basin.* Smithsonian Institution, Washington, D.C.

Davis, Jonathan O., Robert Elston, and Gail Townsend

1973 *Preliminary Archaeological Reconnaissance of Fallen Leaf Lake*. Nevada Archaeology Survey. University of Nevada, Reno.

Downs, James F.

1966 *The Two Worlds of The Washo: An Indian Tribe of California and Nevada*. Holt, Rinehart and Winston, New York.

Duke, Daron, Sharon A. Waechter, and Albert Garner
2010 *Cultural Resources Inventory of NV Energy Transmission Line Assets on the Tahoe National Forest, Truckee and Sierraville Ranger Districts, California.* Far Western Anthropological Research Group, Inc., Davis, CA. On file at North Central Information Center, Sacramento, California.

Elston, Robert G.
1971 *A Contribution to Washo Archaeology.* Nevada Archaeological Survey, Research Paper Number 2. University of Nevada, Reno.
2013 Washo Archaeology: Clued In by Clelow. *Journal of California and Great Basin Anthropology* 33(2):127-131.

Elston, Robert G., and David W. Zeanah
2002 Thinking outside the box: a new perspective on diet breadth and sexual division of labor in the Prearchaic Great Basin. *World Archaeology* 34(1): 103-130.

Elston, Robert, Jonathan O. Davis, Alan Leventhal, and Cameron Covington
1977 *The Archaeology of the Tahoe Reach of the Truckee River.* Northern Division of the Nevada Archaeological Survey. University of Nevada, Reno.

Elston, Robert, Susan Stornetta, Daniel P. Dugas, and Peter Mires
1994 Beyond the Blue Roof: Archaeological Survey on Mt. Rose Fan and Northern Steamboat Hills. Intermountain Research, Silver City, NV. Submitted to the Toiyabe National Forest, Carson City, Nevada.

Esri
2023a How Generalized Linear Regression works. Electronic document, accessed January 31, 2023, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-glr-works.htm, accessed January 31, 2023.
2023b Forest-based Classification and Regression (Spatial Statistics). Electronic document, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatialstatistics/forestbasedclassification regression.htm.
2023c How Presence-only Prediction (MaxEnt) works. Electronic document, https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/how-presence-only-prediction-works.htm, accessed January 31, 2023.
2023d Presence-Only Prediction (MaxEnt) 101: Using GIS to model species distribution. Electronic document, https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/presence-only-prediction-maxent-101-using-gis-to-model-species-distribution/, accessed January 31, 2023.

Galletti, Christopher S., Elizabeth Ridder, Steven E. Falconer, and Patricia L. Fall
2013 Maxent modeling of ancient and modern agricultural terraces in the Troodos foothills, Cyprus. *Applied Geography* 29:46-56.

GeoEngineers
2009 Washington Statewide Archaeology Predictive Model Report. Electronic document, accessed January 6, 2023, https://dahp.wa.gov/archaeology/archaeological-survey-inventory.

Healy, A., C.D. Lippitt, D. Phillips, and M. Lane
2016 A comparison of suitability models to identify prehistoric agricultural fields in western New Mexico. *Journal of Archaeological Science* 11: 427-434.

Heizer, Robert F. and Albert B. Elasser
1953 *Some Archaeological Sites and Cultures of the Central Sierra Nevada.* The University of California Archaeological Survey, Berkeley.

Jochim, Michael A.
1976 *Hunter-Gatherer Subsistence and Settlement: A Predictive Model.* Academic Press, New York.

Kleam, Samantha
2019 *Projectile Points and Late Prehistoric Obsidian Procurement in the Central Sierra Nevada, California: A View from the Rim Fire Collection.* Master's Thesis, California State University Fullerton.

Lindström, Susan and Devin Blom
2021 *Schmaier Pier Project Cultural Resource Inventory.* Consulting Archaeologist, Truckee, CA. On file at North Central Information Center, Sacramento, California.

Lindström, Susan, Judith Marvin, and Devin Blom
2021 *Blakemore Pier Extension Project Cultural Resource Study.* Consulting Archaeologist, Truckee, CA. On file at North Central Information Center, Sacramento, California.

Mergen, Bernard
2014 *At Pyramid Lake.* University of Nevada Press, Reno, Las Vegas.

Nevada Today
2021 Washoe Tribe scientists discuss using traditional knowledge to mitigate wildfire risk and restore Tahoe's Meeks Meadow. Electronic resource, accessed February 27, 2023, https://www.unr.edu/nevadatoday/news/2021/lwf-meeks-meadow-podcast.

Pyne, Stephen J.
2021 *The Pyrocene: How We Created an Age of Fire, and What Happens Next*. University of California Press, Oakland.

Ringhoff, Mary and Edward Stoner
2011 *The River and the Railroad: An Archaeological History of Reno*. University of Nevada Press, Reno.

Ryan, Kevin C., Ann Trinkle Jones, Cassandra L. Koerner, Kristine M. Lee
2012 Wildland fire in ecosystems: effects of fire on cultural resources and archaeology. United States Forest Service, Rocky Mountain Research Station. Electronic resource, accessed February 8, 2023, https://www.fs.usda.gov/research/treesearch/40417#.

Thomas, David Hurst
2013 Great Basin Projectile Point Typology: Still Relevant? In *Journal of California and Great Basin Archaeology* 33(1): 133-152.
2020 *Alpine Archaeology of Alta Toquima and the Mt. Jefferson Tablelands (Nevada).* American Museum of Natural History, New York.

United States Department of Agriculture (USDA)
2022 Who We Are and What We Do. Electronic document, accessed December 22, 2022, https://www.fs.usda.gov/detail/ltbmu/?cid=fsm9_046755.
2022 Geospatial Data Gateway. Accessed January 4, 2023. https://gdg.sc.egov.usda.gov/GDGHome.aspx.
n.d. Description of Gridded Soil Survey Geographic (sSSURGO) Database. Electronic document, accessed January 4, 2023. https://www.nrcs.usda.gov/resources/data-and-reports/description-of-gridded-soil- survey-geographic-gssurgo-database.

United States Geological Service (USGS)
2004 Geologic Provinces of the United States: Basin and Range Province. Electronic document. Accessed December 22, 2022. https://web.archive.org/web/20090125163038/http:/geomaps.wr.usgs.gov /parks/province/basinrange.html.
2018 National Land Cover Database. Accessed January 5, 2023. https://www.usgs.gov/centers/eros/science/national-land-cover-database.
n.d. What is a digital elevation model (DEM?). Accessed January 4, 2023. https://www.usgs.gov/faqs/what-digital-elevation-model-dem.
n.d. National Hydrography Dataset. Accessed January 4, 2023. https://www.usgs.gov/national-hydrography/national-hydrography-dataset.

Verhagen, Philip, and Thomas G. Whitley
2011 Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene. *Journal of Archaeological Method and Theory* 19: 49-100.

Verhagen, Philip
2018 Predictive Modeling. In, *The Encyclopedia of Archaeological Sciences*, edited by Sandra L. López Varela, pp. 1-3. John Wiley & Sons, Inc, Hoboken.

Wachtel, Ido, Royi Zidon, Shimon Garti, and Gideon Shelach-Lavi
2018 Predictive modeling for archaeological site locations: Comparing logistic regression and maximal entropy in north Israel and north-east China. In *Journal of Archaeological Science* 92(2018):28-26.

Waecheter, Sharon A. and William Bloomer
2009 Tahoe Reach Revisited: The Late Pleistocene/Early Holocene in the Tahoe Sierra. In *Society for California Archaeology Proceedings* 2009(22): 1-8. Electronic document, accessed February 15, 2023.

Washoe Tribe
2023 Washoe History. Electronic document, accessed February 20, 2023, https://washoetribe.us/aboutpage/4-Page-washoe-history.

Willey, Gordon R., and Philip Phillips
1958 *Method and Theory in American Archaeology*. University of Alabama Press, Tuscaloosa.

Winterhalder, Bruce, and Eric Alden Smith
2000 Analyzing Adaptive Strategies: Human Behavioral Ecology at Twenty-Five. *Evolutionary Anthropology: Issues News and Reviews* 9(2):51–72.

Yan, Lijie, Peng Lu, Panpan Chen, Maria Danese, Xiang Li, Nicola Masini, Xia Wang, Lanbo Guo, and Dong Zhao
2021 Towards an Operative Predictive Model for the Songshan Area during the Yangshao Period. *International Journal of Geo-Information* 10(217):1-18.

Yaworsky, Peter M., Kenneth B. Vernon, Jerry D. Spangler, Simon C. Brewer, and Brian F. Codding
2020 Advancing predictive modeling in archaeology: An evaluation of regression and machine learning methods on the Grand Staircase-Escalante National Monument. *PLoS ONE* 15(10):1-22.