A Predictive Analysis of Classical Greek Settlements

by

Andrew Joseph Wheeler Turpin

A Thesis Presented to the
Faculty of the USC Dornsife College of Letters, Arts and Sciences
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)
May 2020

Ὅσον ζῇς φαίνου

μηδὲν ὅλως σὺ λυποῦ

πρὸς ὀλίγον ἔστι τὸ ζῆν

τὸ τέλος ὁ χρόνος ἀπαιτεῖ.

hóson zêis, phaínou

mēdèn hólōs sù lupoû

pròs olígon ésti tò zên

tò télos ho khrónos apaiteî.

While you live, shine

Have no grief at all

For life exists only a short while

And time demands its toll

-Epitaph of Seikilos

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank the entirety of my family and friends, especially my parents, whose support was everything in my accomplishing this goal. I would also like to thank my advisor and the members of the committee for all their advice and help in shaping this document. Finally, I would like to dedicate this to my grandparents; *Too ra loo ra loo ra*.

List of Abbreviations

ASCII	American Standard Code for Information Interchange
AUC	Area Under Curve
DEM	Digital Elevation Model
GIS	Geographic Information System
PC	Percent Contribution
PI	Permutation Importance
ROC	Receiver Operating Characteristic

Abstract

All professions work with limited resources, and archaeology is no different. Archaeologists were among the first adopters of GIS technologies, thanks to the field's focus on the spatial distribution of human populations and other location-based information. In recent years, advanced surveying and analysis techniques have boosted the capability of archaeologists to both discover and catalog newfound relics of ancient civilizations. Many agencies that claim jurisdiction over large tracts of land have begun utilizing GIS software to build predictive surfaces from previously gathered spatial data, streamlining their practices of searching for new archaeological sites.

This thesis project aims to quantify Greek settlement patterns using Maxent modeling and use the results to create predictive surfaces in the Peloponnese, a peninsula in Greece that is home to some of the greatest of the Ancient Greek civilization. I chose this area, and this civilization, as it marks one of the most influential in the context of Western development; its contributions range from mathematics to drama to politics. My goals were twofold. First, I identified important environmental variables in the establishment of Ancient Greek settlements, looking at proximity to water sources, proximity to the coast, land cover, underlying soils, elevation, slope, and aspect. Second, using the variables identified, I created a predictive surface for ancient settlements in this region. I expanded on the existing literature by including proximity variables as walking distances, which are more relevant to ancient civilizations than a simple Euclidean distance.

This study serves as a proof of concept of the validity of including walking-time variables in archaeological analysis. These could be utilized in other areas for similar diagnostic and predictive purpose.

Chapter 1 Introduction

Classical Greece was one of the most influential societies in history, especially in the context of what we now conceive of as western civilization. Greek mathematics, architecture, and philosophy were all prevalent enough to be the foundation of much of western society today. (Raaflaub, Ober and Wallace 2007). Greek myths become blockbuster movies, Greek philosophy laid the basis for much of western law, and Greek tragedies and comedies are performed in schools and stages across the country (Kruger 2012; Hartigan 1995).

Due to this massive influence on western culture and society, any archaeological discoveries could potentially have a massive impact on the field for which it is discovered. A lost play of Sophocles, or lost writings of one of the great philosophers, have the potential to truly revolutionize their fields.

For evidence of what a new archaeological discovery can do to enhance a field of study, one must look no further than the discovery of the Dead Sea Scrolls. Discovered by chance in the late 1940s by a group of Bedouin herders, these scrolls affirmed that important writings of the Jewish Torah had stayed remarkably consistent for a thousand years longer than was initially thought.

Another example lies in the Rosetta Stone. Discovered in Egypt during a French invasion, the eponymous stone contained a decree recorded in Greek, demotic Egyptian, and hieroglyphic Egyptian. Scholars realized upon translation that the first two contained the same message, providing the means to crack the long-held code that was hieroglyphics and so revolutionizing understanding of Ancient Egyptian civilization.

These two discoveries, though minor in physical size, were massive in the value they provided to archaeology and knowledge of history as a whole. As such, the goal of this study is to provide a method to empower archaeologists as much as possible, to place them in a position to be able to make these sorts of discoveries.

In the aim of defining a proper scope, this study was restricted across the Peloponnese, a peninsula that comprises about half of Greece and contains many of its important archaeological sites (Biers 1987). The Peloponnese is home to the ancient Mycenean civilization, the predecessors to the Golden Age of Greek classical antiquity and the first major civilization of Greece and Europe as a whole (Boardman, Griffin, and Murray 1991). It also is home to the well-known cities of Sparta and Corinth, as well as Olympia, the location of the famed Olympic Games. Though not ancient history, it is still worth noting, that the Peloponnese is where the Greek war of independence started (Clogg 2013). All these factors make the region a hotbed of archaeology. With Ancient Greece being a place of such profound impact, many of its most important locales have been recorded and we have definite ideas on the location of the most important ones. In addition, over the centuries, multiple excavations have revealed many more sites as well as information about them that could prove useful to this investigation.

Archaeology in Greece covers four basic periods: Neolithic Greece, the beginners of agriculture, Bronze Age Greece, which began the great civilizations of the Minoans and Mycenaeans, Iron Age Greece, which rebuilt society in the wake of the great and mysterious Bronze Age Collapse, and Classical Greece, the period of Greece with which we are all most familiar and during which contributed many of the great influences to modern society mentioned earlier (Lavas 1974). For the purposes of this study, I use a database of settlements from any of these periods, as they all denote settlement patterns of the early Greek peoples.

As the spatial sciences have been revolutionized by the advent of integrated geographic information systems (GIS) and increased processing power, archaeology has benefitted from the exact same advancements. Mapping and computational programs have been key tools in archaeologists' work for a long time now; in fact, archaeologists were some of the first adopters of GIS technology. Archaeology is in many ways the spatial study of human behavior over time, and archaeology inherently carries a spatial component. Archaeological data is rich in spatial components, and GIS is the best option for merging and processing large quantities of geographic data (Snodgrass 1992).

GIS is useful to archaeology in two more specific aspects. First is in surveying. Remote sensing has revolutionized the field, allowing for archaeologists to search far greater areas than ever before. It is also more accurate and easier to store in greater quantities (Fisher and Van Wees 1998). Second is in analysis. The ability to combine and manipulate multiple datasets enables complex analysis of the land and relevant human geographic features.

These advanced techniques have come with advanced methodologies. Most important is the ability to work with inherently incomplete, lacking, and changed data, a byproduct of history. As such, there is a great push to critically assess diverse, variable, and incoherent datasets and translate those for effective use. This is called archaeoinformation science, the ongoing process of gaining awareness of the specificities of utilizing GIS for archaeology (Whitley and Burns 2008). There is a deep process in determining what the exact strengths and weaknesses of GIS are currently as they apply to archaeology and working to fill the gaps and obtain a unique archaeological approach to the spatial sciences.

A new approach in archaeology called ethnoarchaeology serves as additional motivation for this project (Carrer 2013). Ethnoarchaeology is an approach to archaeology rooted in human behavior and culture, aiming at reconstructing ancient ways of life based on material remains. By including variables based around human interaction with the environment, I aim to conduct this study with an ethnoarchaeological approach. This approach promises to give a deeper understanding of the people who inhabited this area and why they chose the areas they did to place down their homes.

1.1 Research question and hypothesis

The idea of this study originated from a single research question; what factors most influenced Ancient Greek settlement patterns? After a realization that this was far too nebulous a goal, I broke this down into two main objectives. The first of which was to quantify which features most influence Greek settlement patterns, in the Peloponnese and its specific regions. A subpoint of this is confirming if the hierarchy matched a list generated from a comprehensive literature review, from highest to lowest 1. Distance from Rivers 2. Distance from Coast 3. Elevation 4. Slope and 5. Aspect (expressed with the variables Northness and Eastness). The second objective is to use this information to create predictive surfaces to aid in future archaeological searches.

The current hypothesis centers around the first part of this study, with a list of expected order of importance for the environmental variables. All settlements require freshwater to function, for many reasons. Transportation, irrigation, and of course drinking water are all imperative to the success of a settlement, all of which are contingent on the presence of freshwater. As the Greeks were a seafaring people, the ocean is predicted to be the second most important factor. The ocean brought greater food, trade, and protection from other city-states through open water transport (Elleman 2017). I then predict that elevation will be the next most important factor, then slope, finishing with aspect as the least important of these deciding factors. In summary, I predict the order of importance (from large to small) for all areas will be 1.

Distance from Rivers 2. Distance from Coast 3. Elevation 4. Slope and 5. Aspect (expressed as a combination of Northness and Eastness).

To answer the question of what environmental factors most influence the settlement patterns of Ancient Greeks in the Peloponnese, I will be using Maxent, an open-source modeling software, to test this hypothesis. Maxent (short for maximum entropy) utilizes observed occurrences to create a probability distribution. Its focus on presence only data and environmental variables makes this program especially well-suited to this study.

Maxent was developed in 2004, with the goal of modeling the occurrence of species within a given area utilizing environmental variables. Maxent is valuable to this project because of its ability to measure the importance of each environmental variable to the placement of the species. This is an important component of the study, to be able to quantify the settlement patterns and so assign value to what factors drove the choices of settlement location for these ancient people.

One of the largest changes I aim to incorporate into this thesis is the inclusion of data related to bodies of water. When dealing with settlements, position to fresh water was critical in establishing a town due to need for drinking water, sewage, transportation, irrigation, and the many other benefits that fresh water provided. It is no wonder that civilization as we conceive of it today began and flourished in the great river valleys of Mesopotamia, the Indus River Valley, and China.

Proximity to the coast was also vitally important. Ancient Greece was a seafaring culture, with many of the most powerful city-states building their power of the backs of their triremes (Hale 2009). Proximity to the coast allowed for trade, for fishing, for transportation to other cities, and much more (Tsetskhladze 1999). The importance of the sea was such that any study would be incomplete without it (Boardman 1980).

Partially due to the structure of Maxent, my review turned up no studies that utilized proximity to water as a portion of its study. To address this gap and accurately address this vitally important component of the settlement of ancient societies, I aim to develop a technique to include distance from both freshwater sources and the coast in the final study.

The project was based around three main stages. The first stage consisted of obtaining and cleaning all the data so that it could be effectively used in the Maxent study. The second stage consisted of running the analysis and identifying which environmental variables contribute most to the placement of the settlements in the study. And finally, the last stage consisted of utilizing this information to develop the final predictive surface for the study area, predicting where undiscovered settlements are likely to lie.

1.2 Thesis Document

The goal of this project is to determine percentage contributions of different environmental variables to the presence of ancient Greek settlement and use this information to create predictive surfaces for future study sites. Chapter 2 provides historical context and background on the use of geographic information science in archaeology, as well as providing the variables that are included in the study. Chapter 3 discusses the methods used in preparing the environmental variables for Maxent and the options selected for the modeling software. Chapter 4 provides and interprets the results from each Maxent run. Chapter 5 gives the final conclusions and limitations of the study and provides recommendations for future work.

Chapter 2: Background

The unique geographic and political features of Greece created an equally unique pattern of settlement for the ancient peoples that lived there. High urbanization and a seafaring culture were some of the most unique aspects, reflective of the numerous independent city-states that defined Greek civilization at this time. GIS and archaeology have a long history together in aiding with the excavation of these settlements and benefit greatly the creation of predictive surfaces.

2.1 Greek Settlement Patterns

Understanding Greek settlement patterns reveals more than meets the eye (Graham 2015). From these things we can understand “geographical knowledge, seafaring, state organization, economic conditions,” and much, much more.

Hellenistic settlements were prevalent through much of the Ancient World, stretching through Europe, many of the islands in the Mediterranean Sea, and in Asia Minor (Cohen 1995). Such a diversity of settlement provides a great deal of locations and potential to expand the study if need be to validate any of early discoveries or provide comparisons between areas.

The fundamental unit of Greek society was an entity known as the polis (Hansen 2006). In Greek, this translates to both ‘city’ and ‘state,’ so giving a greater insight into its full meaning as ‘city-state’. Ancient Greece had around 1,500 of these, lending great credence to the argument that the polis was truly the foundation of Ancient Greek society. This has led Greece to be defined as a ‘city-state culture’, which has been defined as consisting of four main components. First is a high degree of urbanization. Second is an economy based on trade and centered on the city’s market. Third is a political process wherein many decisions are made by a majority vote of assembly. Fourth is interactions between city-states which resulted in the rise of leagues and federal states.

The Ancient Greek settlement was a very specific and intentional creation, bounded by specific criteria (Lavas 1974). Its purpose was to fulfill defensive, residential, sociopolitical, and economic needs. The ideal size was between five to ten thousand inhabitants, within an area no greater than 180 hectares. Identifying a site as a previous settlement is a matter of identifying the presence of government buildings, a gymnasium, a theatre, an agora, and other such public use buildings. The inhabitants were mainly farmers or fishermen, bound together by a common currency, a self-sufficient economy, and an autonomous government (the previously mentioned polis or city-state).

Ancient Greek settlements shared many commonalities with other settlements of peoples around the same time (Lavas 1974). Buildings were constructed close together for defensive purposes. Walls typically ringed the settlements, for much the same purpose. Locations were chosen carefully, typically near some sort of natural resource, with fresh water being a universal concern.

However, Greek settlements had several distinct features that differentiated them from other peoples of the time period (Lavas 1974). First was the political and social structure of ancient Greeks, which separated them into the aforementioned city-state known as the polis. Greek society was very fractured, a feature that the geography of Greece undoubtedly played a large role in. High mountains, narrow passes, and jagged coastlines all combined to create cities that were fiercely independent and ferocious in defending that freedom (Bintliff 1977).

Despite this, the Hellenistic world was very aware of its common heritage (Cohen 1995). Greeks shared a common tongue, common religion, and common traditions. Pan-Greek competitions even existed in events such as the famous Olympic Games, in which the Greek states would lay aside their enmities for a time to share in athletic competition and brotherhood.

Despite these binding factors, the Greeks never seemed to entertain any concept of unification (Hammond 1959). When a polis was defeated by another, it was rarely conquered or ruled by the victor, usually just paying tribute and losing its political clout in the region. Greeks typically politically organized in leagues, such as the famed ones that threw back the Persian invasions or stood up to Philip of Macedon. However, membership in these was extremely fluid. Lost in a patriotic narrative is the fact that in the former, many Greek poleis stayed neutral during the Persian invasion, and the allied went right back to fighting each other as soon as the Persians had departed. And in the latter, more than a few poleis fought at the side of Philip of Macedon, as his soon to be famous son and heir Alexander ended that league only to substitute his own. Even for the King for whom the ends of the earth were not enough, he arranged the defeated Greeks in an allied League of Corinth, speaking to the fierce independence of the polis and the power of this institution.

The Greeks were also famous colonizers (Graham 2015). Around the Archaic Period (800-400 BC) the population of Greece grew an incredible amount, by some estimates increasing by a factor of ten. This would increase the population from 800,000 to 10 to 13 million. With very limited arable land, the Greeks began fanning out across the Aegean, settling colonies wherever they went. The coast of Asia Minor, Thrace, the island of Cyprus, and the coast of the Black Sea were first. Soon after, the Greeks expanded even further, reaching the coasts of Illyria, Southern Italy, Southern France, and even as far as Northeastern Spain (Graham 2015). In most cases, these colonies held little to no allegiance to their mother polis and rapidly established their own political destiny, carrying forward the Greek tradition of fierce independence.

2.2 Archaeology in Greece

Archaeology in Greece is special in its ability to cross disciplines in its impact (Snodgrass 1992). A great deal of that is due to the far-reaching accomplishments of the Greeks, and the society they created that still informs and impacts societies today. A new discovery in Greek archaeology could impact art history, philosophy, or a great many more varieties of disciplines and approaches. And in this vein, any archaeological study undertaken in Greece is inherently a multidisciplinary one. A rigorous review of historical and classical context is necessary to properly design any study as well as make proper sense of any eventual results.

The highly urbanized nature of Greek society has made archaeology very profitable in Greece. Rural sites only classify some twenty percent of classical Boeotia, and forty percent for Archaic-Classical Laconia. This indicates that a large portion of the population lived in cities during this time period, impressive considering the majority of the world consisted of something like ninety percent farmers during this time period, living in small farming villages or in isolated homesteads.

Archaeology in Greece covers four basic periods: Neolithic Greece, the beginners of agriculture, Bronze Age Greece, which began the great civilizations of the Minoans and Mycenaeans, Iron Age Greece, which rebuilt society in the wake of the great and mysterious Bronze Age Collapse, and Classical Greece, the form of Greece with which we are all most familiar and contributed many of the great influences to modern society mentioned earlier (Boardman, Griffin and Murray 1991).

The field of archaeology has changed as well. Previous methods have been enhanced by the addition of surveying, remote sensing, and advanced analytics (Fisher and Van Wees 1998). These discoveries have, in many cases, caused a rethinking of what we consider to be canon about the Ancient Greek civilization. As such, throughout this study I must strive to strike a balance between a rethinking of Greek society and what the results contribute to preexisting knowledge.

2.3 GIS and Archaeology

Advances in GIS have enabled a great deal of potential integration into archaeological workflows (Farinetti 2009). The vast availability of data has allowed for complex modeling operations to be undertaken, and for these operations to smoothly combine data from multiple sources. Most exciting for this study is the ability to integrate data both with material culture and the environment. Advances in GIS also allow for a critical comparison of micro-regions. This will be especially useful for a suitability surface and the predictability analysis that will accompany it.

These advances also come with advanced methodologies (Jones and Sarris 2000). Most important, especially when working with inherently incomplete historical and archaeological datasets is the ability to critically assess diverse, variegate and incoherent datasets and translate that for effective use in this analysis.

The ability of GIS to work with vast amounts of spatial data over very large areas has made it a boon and natural addition to the field of archaeology. Archaeologists have taken full advantage, utilizing these technologies in many a study (Lock and Stantic 1995). These technologies are now in use by many government agencies that manage large areas of land, such as the Bureau of Land Management or the Department of Transportation.

Density mapping, interpolation, and distributional analysis have all contributed vastly to archaeology, allowing for archaeologists to examine processes of change and the evolution of societies over time (Bevan 2013). This also allows them to attempt similar objectives to this study in examining unifying features among settlements of a variety of cultures and time periods.

2.4 Predictive Surfaces in Archaeology:

Predictive surfaces have been used in archaeology, with a variety of methods and as much variety of success. It is my hope with this study to devise a system that is able to take into account historical context and a deep variety of land features and, as a result, produce replicable results across the Hellenic world. One of the largest challenges with such a system is working with qualitative data to derive quantitative insights (Mehrer and Wescott 2005).

One such method utilizes Binary Logistic Regression to achieve its goals of predicting where ancient peoples may have chosen to place down their roots (Vaughn and Crawford 2009). Utilizing remotely sense data, this method includes one variable from a Landsat image representing vegetation patterns consistent with Maya settlement, two variables derived from a digital elevation model, and a hydrology map representing the availability of resources critical to the Maya.

Another such method utilizes fuzzy logic to make its predictions. This consists of reclassifying variables over a defined study area to identify locations with a high possibility of occurrence. The values are obtained by identifying the optimal conditions for a particular object to occur, then combining all rasters that contain that information to isolate the areas that fit all the criteria (Alexakis and Apostolos 2010).

The Georgia Coast Model converts costs and benefits into so-called ‘spatial currencies,’ in its example case to identify where settlers would prioritize in pursuit of maximum caloric intake (Verhagen and Whitley 2012). Such a model allows for small deviations, accounts for variation within a geographic area, and allows researchers to develop a sort of ‘cultural biography.’

Finally, an ethnoarchaeological predictive model was created focusing around Chi-Square and Monte Carlo techniques (Carrer 2013). This approach is inductive, focusing around a sample of archaeological sites whose characteristics are generalized in order to predict the locations of other such sites. Ethnoarchaeology deals with a gap in the literature in the anthropological idea of causality which truly links to human settlement patterns. Ethnoarchaeology deals with human behavior as well as material culture through inclusion of a great deal more variables than the typical study. It is this ethnographic approach that I feel holds the greatest potential for my own study.

2.5 Maxent

Maxent, short for Maximum Entropy Species Distribution Modeling, is a modeling software that was originally conceived of as a tool for use in ecology. Maxent functions by using information available from environmental layers (covariates) and points of presence of observed species and putting them through a series of transformations, or formulas. The tool utilizes the values of the covariates at the species’ points of presence to conduct a high number of formula combinations to find the best model fit (Kalinski 2019). The final output is a predictive surface of locations likely to hold that species, as well as the statistical explanation of how that conclusion was arrived at (Phillips 2005). The power and relative ease of use of Maxent has led to it being used beyond ecology, for modeling tasks in diverse fields including archaeology.

One of the advantages of Maxent is the ability to work with an absence of data. Maxent does not assume null values at its locations with absence of data, but instead interpolates these areas to create a continuous surface. It also has a great deal of underlying complexity that adds weight to its analysis, as well as an interface that makes it relatively easy to use.

These things that make Maxent such a powerful tool can also be drawbacks. Maxent's ability to handle a large range of variables can lead to over inclusion of covariates, each of which can significantly impact the final output if not chosen for good reason. A review of the literature has provided a comprehensive understanding of which inputs to select and for what reason. This will be covered further in the Methods section of this paper.

Maxent's deep underlying complexity and ease of use of its interface can lead to errors when using the default settings. To derive the most meaning from any models created, users must understand thoroughly which parameters they are selecting and for what reasons (Kalinski 2019). These parameters will be covered thoroughly in the Methods section of this paper.

Maxent's capabilities have seen it used in archaeological studies similar to this one. Maxent has shown its abilities to both predict archaeological site locations and describe the contributions each covariate made to the predictive model. Limitations across these studies focus primarily on availability of data, pointing to the importance of proper data collection and organization when undertaking a Maxent-heavy study.

Chapter 3 Methods

This chapter details the study area within the Peloponnese, and the 3 distinct geographic regions within that were tested. Data utilized within the various Maxent models are detailed, as well as how slope, aspect, distance from rivers, and distance from coast were calculated for inclusion in the models. Lastly, parameters for each Maxent model are discussed.

3.1 Study Area

As discussed earlier, the study area consists of the Peloponnese, a peninsula in Greece (depicted in Figure 1). This area is 8,320.3 square miles with over 1,300 archaeological sites of all types included in its area. The Peloponnese is a diverse region, with high mountains, lush valleys, and a deeply indented and jagged coastline. These variety of features lead to a great diversity of options in establishing settlements, each combination with their own set of advantages and disadvantages. This of course is the objective of this research, to identify the exact combinations that were of importance to the Ancient Greeks and utilize this information in a predictive manner.

Figure 1 Peloponnese Study Area

With such a large area, it is potentially necessary to reduce the study area into smaller sample areas in order to derive better conclusions from the study. In order to test this secondary hypothesis, I tested Maxent models on the entirety of the Peloponnese, as well as three administrative regions of Greece within the Peloponnese, each consisting of a different type of area. First, the island of Kefalonios, which has no river features. Second, the coastal area of Messines. And third, the inland region of Gortunias, which is devoid of coastal features. Each of these areas contains around seventy settlements, making them ideal areas to compare, and can be seen in Figure 2. The results of each of these areas was then compared to the Peloponnese as a whole to identify standout traits of each area and compare model performance.

Figure 2 Reduced Study Areas

3.2 Software

The Maxent program, as discussed earlier, is used to conduct the modeling. ESRI’s ArcMap 10.7.1 and the various tools contained within were also used in conjunction with Maxent, to observe and prepare the data and to represent the results.

3.3 Data Sources

The modeling requires an input of multiple environmental variables, all of which will be utilized and measured for impact in creation of a probability prediction surface. All data were obtained from the official open data portal maintained by the Greek government. These data were transformed for use in English and the Latin alphabet by the American University at Athens. All these variables, and their metadata, can be summarized in Table 1 below.

Table 1: Data Sources

Layer	Original Data Format	Source	Publishing Year	Spatial Resolution
Settlements	Vector – Point	Geodata.gov.gr	2015	1352 data points, 8,320 mi^2
Elevation	Raster – Tiff	NASA SRTM mission	2012	3 arc seconds, +/- 90 m
Slope	Raster – Tiff	NASA SRTM mission	2012	3 arc seconds, +/- 90 m
Eastness	Raster – Tiff	NASA SRTM mission	2012	3 arc seconds, +/- 90 m
Northness	Raster – Tiff	NASA SRTM mission	2012	3 arc seconds, +/- 90 m
River Lines	Vector – Line	Geodata.gov.gr	2015	8,320 mi^2
Coast Lines	Vector – Line	Geodata.gov.gr	2015	8,320 mi^2

I included the maximum number of environmental variables possible to attempt to gain as complete a picture of settlement patterns as possible. These will be proximity to freshwater, proximity to the ocean, elevation, slope, and aspect (expressed as Northness and Eastness).

Settlements

The settlements were taken from a point shapefile obtained from the American University of Classical Studies at Athens. The shapefile contains the names of each settlement in both the Greek and Latin alphabet, as well as the time periods each settlement was constructed in. This study was conducted using settlements from the same (Pre-Classical) era.

These settlements were prepared for analysis in Maxent by exporting the settlements from each region into a comma-separated values file (.csv) after calculating the latitude and longitude of each point. Extraneous information, such as the name of the settlement in the Greek alphabet, was deleted from the csv. Then, the file was given a column “species” and each set of coordinates was given the same name, so that Maxent would read the points all as part of the same “species.” This allowed for each set of Greek settlement points to be used in the Maxent analysis.

Distance from Coast

While many previous studies utilize Euclidean distance to gauge their distance variables, this does not provide a clear picture of the difficulty in reaching a particular area. The Euclidean Distance “as the crow flies” ignore terrain and remove the human elements from an inherently human-oriented study that aims to examine and explain human behavior.

This necessitates the creation of a walking-time layer, to most effectively measure the practical distance each point lies from the relevant waters. This can be accomplished by utilizing the ArcGIS tool Path Distance, which takes into account terrain features to determine the practical distance of each point from the relevant features. This tool functions by calculating the cost distance from the nearest source input feature, accounting for distance and surface factors. This differs from the similar Cost Distance tool, which utilizes a cost surface weighted and created by the user as opposed to the actual surface. A benefit of using Path Distance as opposed to Cost Distance is that it accounts for the intricacies of a physical surface, such as whether the movement is up or down a slope (ESRI, “Path Distance,” August 28, 2019).

The Path Distance tool only requires two inputs for these purposes. The first is an input which details the original feature, from which the distance will be calculated. In this case, this is a line shapefile of the coastline surrounding Greece. The other input is the cost surface, which in this case is a DEM of the administrative region. This was obtained from techniques detailed in the section on elevation data.

Utilizing this tool, as well as including elevation as a variable, raises issues with over including variables. Specifically, with including variables that are too closely related to each other, a phenomenon called collinearity. However, a study found that there was low difference between collinear and non-collinear datasets when used in methods such as Maxent (De Marco and Nóbrega 2018). In light of this discrepancy, it is best to see the performance of the models to evaluate the effect of multiple collinear variables. The Path Distance surface was calculated for the region as a whole and for each regional study site. The resultant distance surface for the region of Messines is shown in Figure 3.

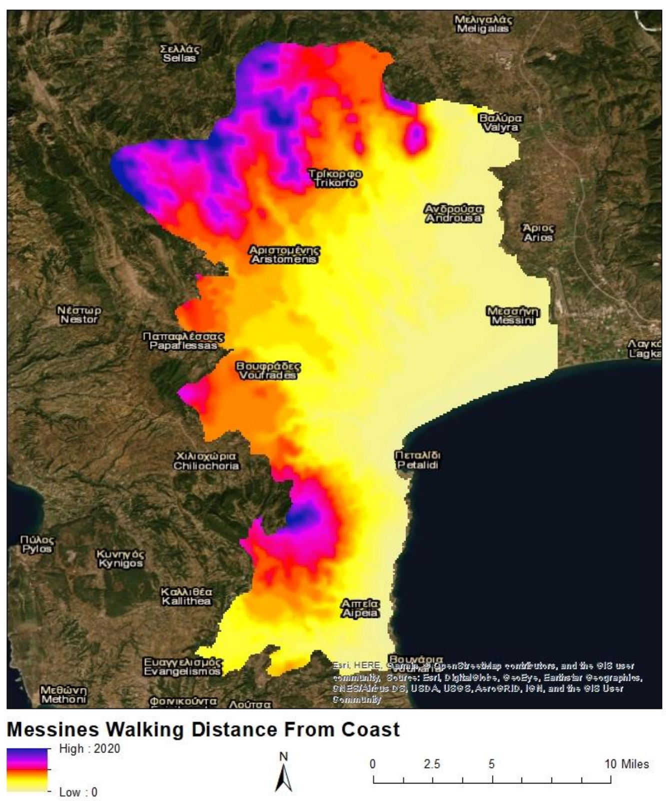

Figure 3 Messines Walking Distance from Coast

Distance from Rivers

The distance from rivers variable is extracted in much the same way as the distance from coast variable, using the ArcGIS Path Distance tool. As with the Distance from Coast variable, the cost surface was the DEM of the region and the vertical factor was set to linear. The line feature utilized was a rivers shapefile obtained from the American University at Athens. The resultant distance surface is shown in Figure 4.

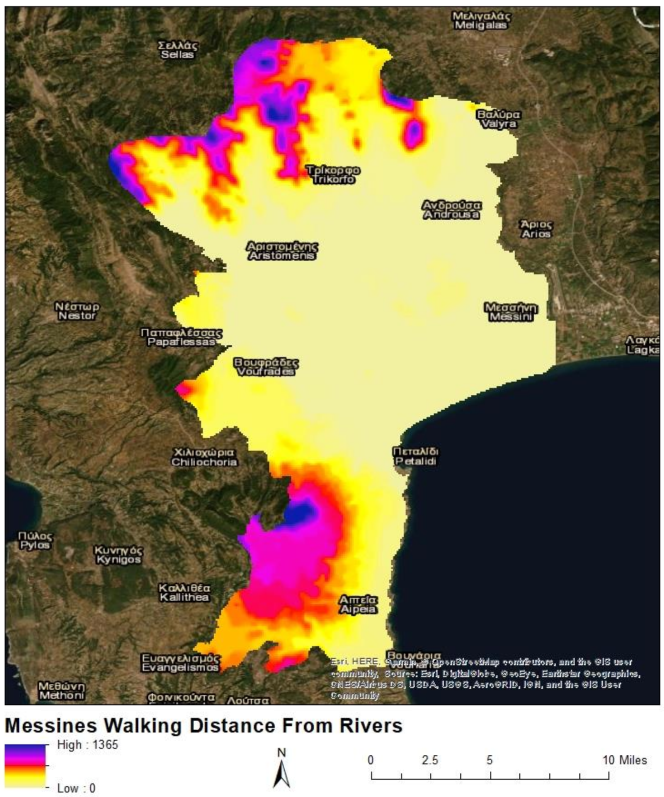

Figure 4 Messines Walking Distance from River

Elevation

Elevation data is a necessity, and also one of the more interesting datasets to explore conceptually. In a region wracked by constant war and conflict between neighbors, would settlers prefer to establish on top of nearby mountains and hills, for defensive purposes? Or would lower ground, nearer to rivers and fertile fields, be the preferred spots to place down their roots? Elevation data provides the insights for building the predictive model for the archaeological sites in the region.

To prepare this data for processing, I first looked for the DEM with the best spatial resolution for this region. DEMs of Greece were sparse, but I was able to obtain a DEM of the entirety of the country of Greece from the University of North Carolina, Chapel Hill's Ancient World Mapping Center. According to the University's site, this data was sourced from the NASA SRTM mission, in the year 2012. It has a spatial resolution of 3 arc-seconds, or +/- 90 meters. The surface was clipped to the study area using the Extract by Mask tool in ArcGIS, and an example is shown in Figure 5.

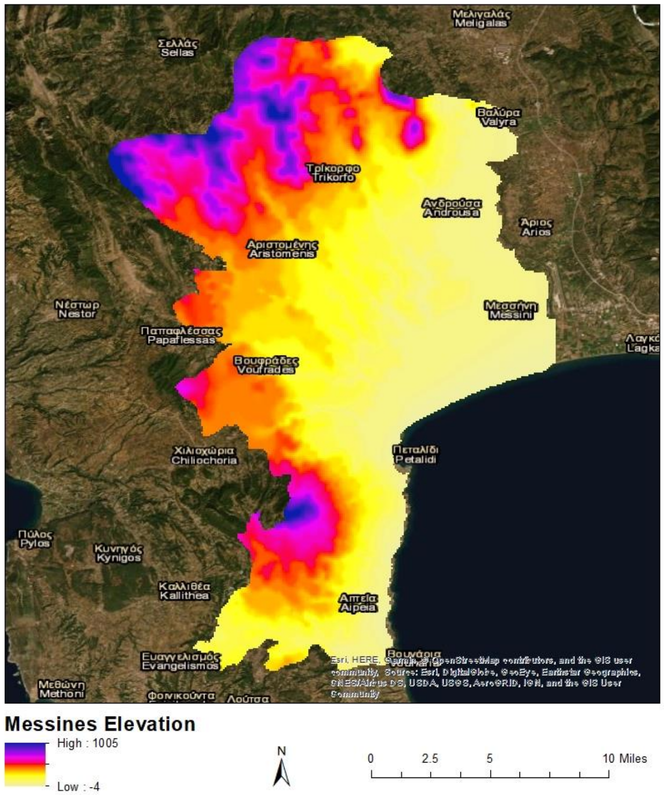

Figure 5 Messines Elevation

Slope

Slope data is an important environmental variable to consider, as those constructing ancient villages and towns would undoubtedly prefer to build their settlements on flat or gently sloped ground when it was available. It would be interesting to see if there is a greater slope than flat (around 0 degrees) that was preferred by ancient builders, for reasons possibly pertaining to ease of construction or defense. It is also extremely necessary to obtain this information in order to assist in the construction of the predictive surface. This slope data was obtained from the DEM using the Slope tool available through ArcGIS, and an example is depicted in Figure 6.

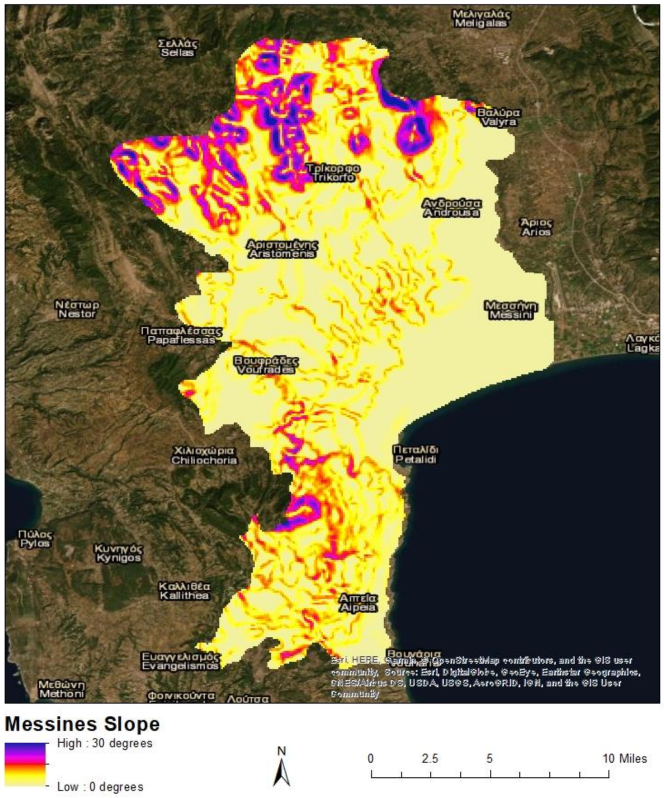

Figure 6 Messines Slope

Aspect

Aspect data would allow for insight as to if there was any particular direction favored. Aspect affects sunlight, temperature, and the kinds of vegetation that would grow in an area as well as the animals that sort of vegetation would attract to the region. This initial Aspect surface was created from the Slope surface derived from the DEM, utilizing the Aspect tool available through the ArcGIS toolbox.

However, aspect data is a circular variable which makes it unfit by itself for use in linear statistics such as this study. This data can be transformed for effective use in these types of studies. One such practice is the transformation of these into continuous variables, corresponding to the gradient of belonging to one of two directions. For this study, the data was transformed into two variables, 'Northness' and 'Eastness' (Lecours et. al 2016). This results in a value ranging from -1 to 1 for each variable. A 1 value represents total belonging to the variable's name, so in northness's case, a 1 value would be totally north. A -1 would be in the opposite direction, so in northness's case would indicate south. A 0 value, in the middle of these two, would indicate either east or west for the northness variable. This transformation into a linear scale allows aspect data to be effectively used in Maxent (Langston 2013).

The steps in doing this were to convert the aspect data, initially in degrees, into radians and take the Sine for Eastness and the Cosine for Northness. Luckily, each of these operations could be accomplished with a single formula using the Raster Calculator tool. The formulas for these are as follows:

$$\text{Eastness} = Con\ (\ [Aspect] == \ -1\ , 0\ , Sin\left([Aspect] * \frac{3.14159}{180}\right))$$

$$\text{Northness} = Con\ (\ [Aspect] == \ -1\ , 0\ , Cos\left([Aspect] * \frac{3.14159}{180}\right))$$

These were then converted to the ASCII data format using the tool Raster to ASCII and then from there used in Maxent. Examples of the Aspect, Eastness and Northness can all be seen below in Figures 7, 8 and 9 respectively.

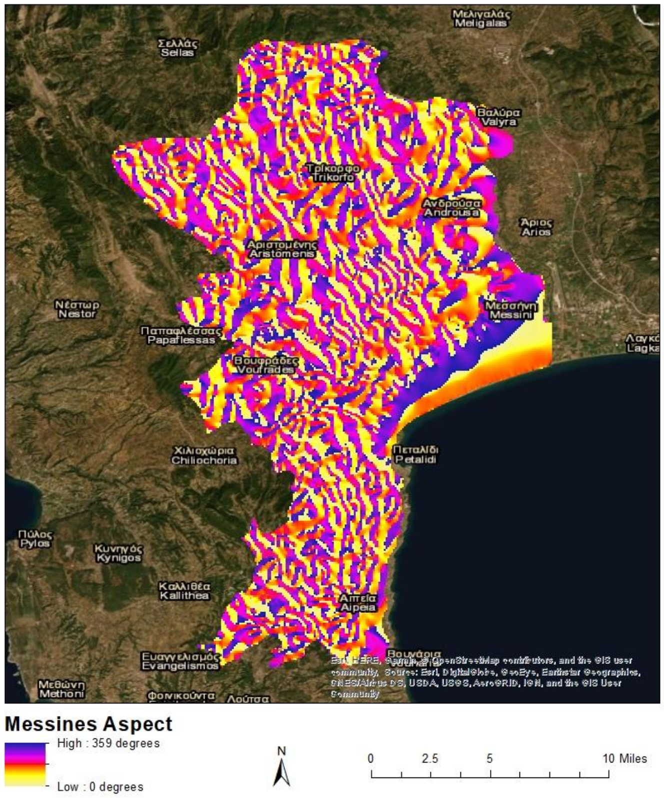

Figure 7 Messines Aspect

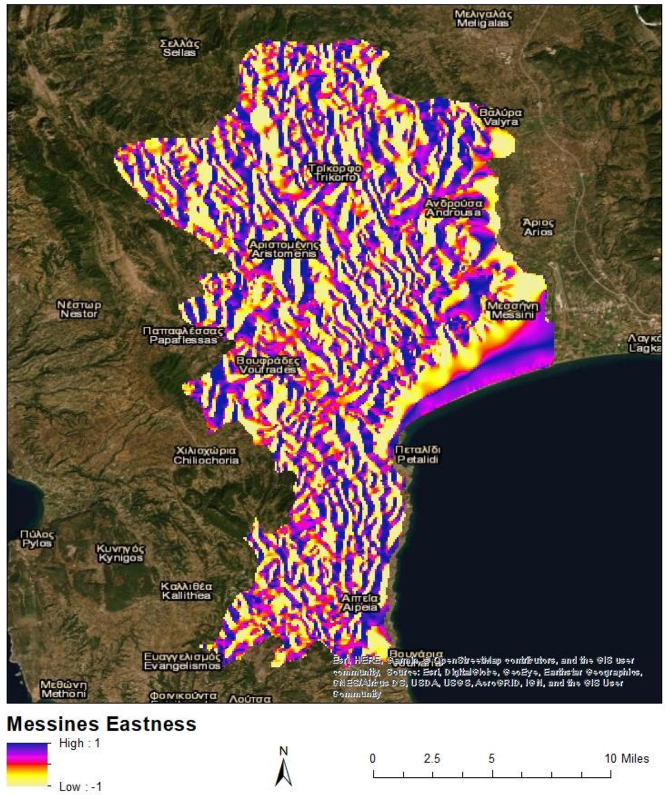

Figure 8 Messines ‘Eastness’

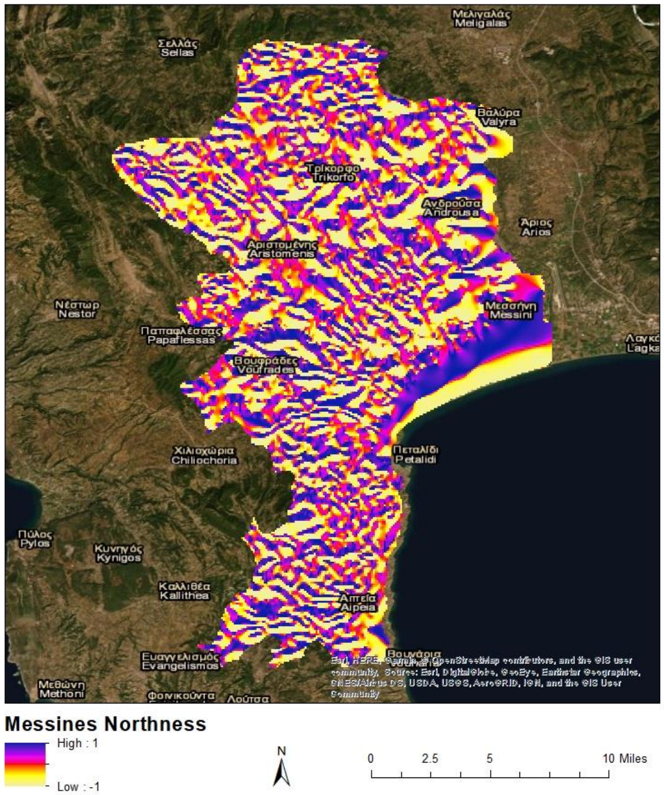

Figure 9 Messines ‘Northness’

3.4 Procedures/Analysis:

The first project stage consisted of the transformation of data for use in Maxent. This was accomplished through the use of ArcMap 10.7.1 and the tools contained within.

Maxent required for the points input to be arrayed in a specific format, with the first column of every point containing the same name. This is required for the points to be analyzed together, or else each will be analyzed separately. It also requires the next two columns to contain the longitude and latitude of each point.

The environmental surfaces used in the analysis must be in the .asc file format. After running the INT tool to convert the raster’s values to integers, as ASCs do not support floating point data types, the rasters can be run through the Raster to ASCII tool. This converts the rasters to a format that can be used by Maxent.

3.5 Maxent Modeling and Settlement Distribution

Maxent includes the option to conduct its analysis using linear, quadratic, product, threshold and hinge features. Product and Threshold features are only advisable to run if there are over 80 location sites available (Steven Phillips, Miro Dudik and Rob Schapire, "Maxent Help," December 2019), and since all three survey areas contain around 70 sites, these settings were turned off. Doing this also helped preserve uniformity across the models.

Maxent was run with the parameters to which generate a predictive surface that serves as the final product of this study. It also includes the jackknife test, which gives the importance of each included variable. This test was run on each of the variables and returned a range of gains for that variable. High values of gains indicate high correlations, and low gains indicates a low correlation. These gains symbolize how much better-than-random the model fit is. As such, high gains mean a higher predictive value. The output format was selected as logistic, allowing to create probabilities of occurrence for studied points. These model options can be seen below.

Maxent allows for the classification of covariates as either categorical or continuous data. Categorical contains a fixed number of categories to which the covariate belongs. An excellent example of this would be soil type, in which membership must be totally contained within the defined category. Continuous data, meanwhile, contains an infinite number of values bounded between two values. This second type, continuous, is indicative of the data used in this study and so was the option selected when uploading the environmental layers. These settings can be seen below in Figure 10.

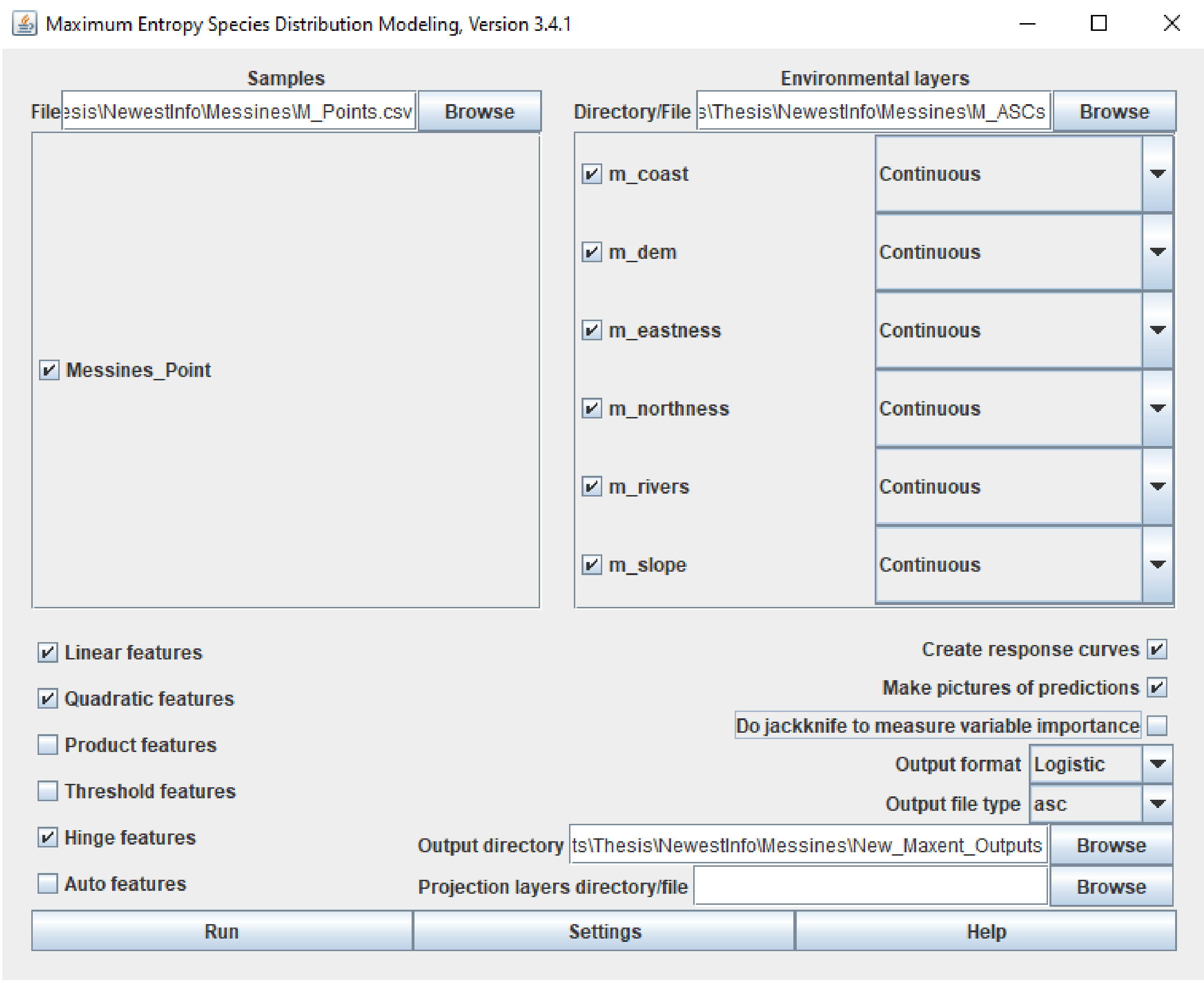

Figure 10 Maxent Settings Example

A setting was also included to return a value of 20% of the points tested as possible locations for other 'species,' utilizing a standard deviation of the data provided.

Maxent produces a raster of the results of the probability distribution, an average of all the results for the model runs of the 'species.' This final probability distribution serves as the predictive layer, placed into ArcGIS to provide a final picture side by side with environmental variables.

Chapter 4 Results

This chapter explains the results of the four models' runs. First, the chapter covers the outputs produced by Maxent and examines their significance in relation to this study. Then, each region's results are examined and measured against the hypothesis. Each region's section also includes the predictive surface generated through the Maxent modeling. Finally, the chapter concludes with an assessment of the overall performance of the models.

4.1 Explanation of Maxent Outputs

Area Under Curve, from here on expressed as AUC, is the area between the graph of y=f(x) and the x axis. This area is intended to provide the reliability of the model it measures, though it is not a perfect measure. AUC indicates how well the model can discriminate presence locations from background locations, with a higher AUC value indicating better discrimination between presence and background locations (Kalinski 2019). Any value under 0.7 means the model is lacking. Values from 0.7 to 0.9 mean the model is performing at an average level, and values above 0.9 indicate a high performing model (Franklin and Miller 2014).

Maxent returns two main outputs in the final table it generates for each model that is tested: percent contribution and permutation importance. Percent contribution is defined as the increase in gain, the penalized likelihood function, due to modifying the covariate's coefficient. This is then converted to a percentage for ease of understanding. There are issues with this as a form of measurement, as it is unreliable when the model contains multiple closely related variables. In addition to this, percent contribution is calculated from the code path of the model. It is entirely possible that a subsequent model could arrive at the same conclusions, but with a different code path return a different percent contribution value.

Permutation importance is obtained from randomly permuting values of each covariate against training points and measuring the resulting decrease in AUC. A high permutation importance value is indicative of a high model dependence on that variable. Due to the inclusion of multiple highly correlated variables in elevation, slope, and the walking time raster, permutation importance is a better measure of a covariate's significance than percent contribution for this study. As such, permutation importance will be the main measure by which I will gauge the impact of the environmental layers. All permutation importance values will be summarized in Table 2 at the end of the chapter. Percent contribution will be included as well but summarized in Table 3 at the end of the chapter.

Another important output will be the jackknife test, which provides more information on the importance of each covariate to the overall model performance. The jackknife shows each covariate with two bars depicting gain, a darker blue "with only variable" and a lighter green "without variable." These are all plotted with a red "with all variables" bar on the bottom of the graph. "With only variable" shows the gain of the model when it is run with that variable in isolation. This is primarily used to show which variable has the most unique information. "Without variable" shows the gain with the model run without that variable. The amount of gain lost here shows how much the model relies on that variable. This is used in conjunction with permutation importance to determine the impact of omission of variables on the model's performance.

Finally, the last important output is the response curves, which show the distribution of covariate values and the correlating gain. These values allow one to see the effect of each covariate coefficient on the Maxent prediction, as well as compare covariates to identify similar cutoff or convergence points.

4.2 The Peloponnese

The entirety of the Peloponnese model ran with 1,352 settlement test points, all of which were given the same 'species type' of P_Point. This study area included coast distance features, river distance features, elevation, slope, northness, and eastness features.

For the remainder of the results, a measure called ROC returning a mean of (AUC = value) will be used. Simply put, the ROC is a probability curve generated for the model and the AUC, the area under that curve, is a measure of how good the model is at discriminating between presence and background data. The ROC curve for the entirety of the Peloponnese returned a mean of (AUC = 0.623). This model's performance was below average, with low confidence in the performance of this model being better than random.

Of the variables included in this study, elevation had the greatest impact on the occurrence of settlement points with a permutation importance value of 54.9. The next most important variable was river distance, with a permutation importance of 24.4, lower but still impactful. The next variable, distance from coast, had a relatively low high permutation importance value of 10. The next variable, northness, had a low permutation importance value of 6.2, while the other aspect variable, eastness, also had a relatively low permutation importance value of 4.4. The last remaining variable, slope, had a low permutation importance level of 0, meaning that its contribution to this model was negligible. Overall, the model performed with a high amount of instability with an AUC of 0.632, thus making its results unreliable.

The prediction surfaces generated for the Peloponnese were erratic as well, with high probability areas scattered across the study area. Erratic as it is, it does point to clusters around coastal regions and river valleys. Low probability areas lie at high elevations and away from the major rivers in the area. This predictive surface can be seen below in Figure 11.

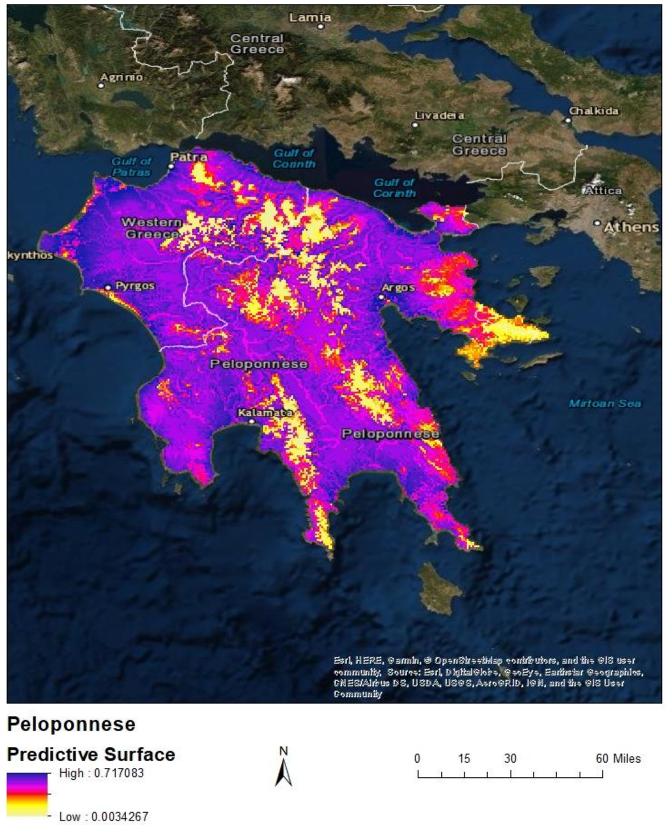

Figure 11 Peloponnese Predictive Surface

4.3 Messines

The model for the coastal region of Messines ran with 70 settlement test points, all of which were given the same ‘species type’ of Messines_Point. This study area included coast distance features, river distance features, elevation, slope, northness, and eastness features.

The ROC curve for Messines returns a mean of (AUC = 0.598). This model’s performance was below average, with low confidence in the performance of this model being better than random.

Of the variables included in this study, eastness had the greatest impact on the occurrence of settlement points with a permutation importance value of 39.5. The subsequent variable, elevation, had a relatively high permutation importance value of 26.5. The next variable, northness, had a permutation importance of 23, lower but still impactful. The next variable, slope, had a relatively low permutation importance value of 9.1. The next variable, distance from rivers, had a very low permutation importance value of 1.9 indicating a low contribution. The remaining variable, distance from the coast, had a low permutation importance level of 0 meaning that its contribution to this model was negligible. Overall, the overall instability of Messines with an AUC of 0.598 means that its results must be considered highly unreliable. The prediction surfaces generated for Messines were erratic as well with high probability areas scattered across the study area, as seen below in Figure 12.

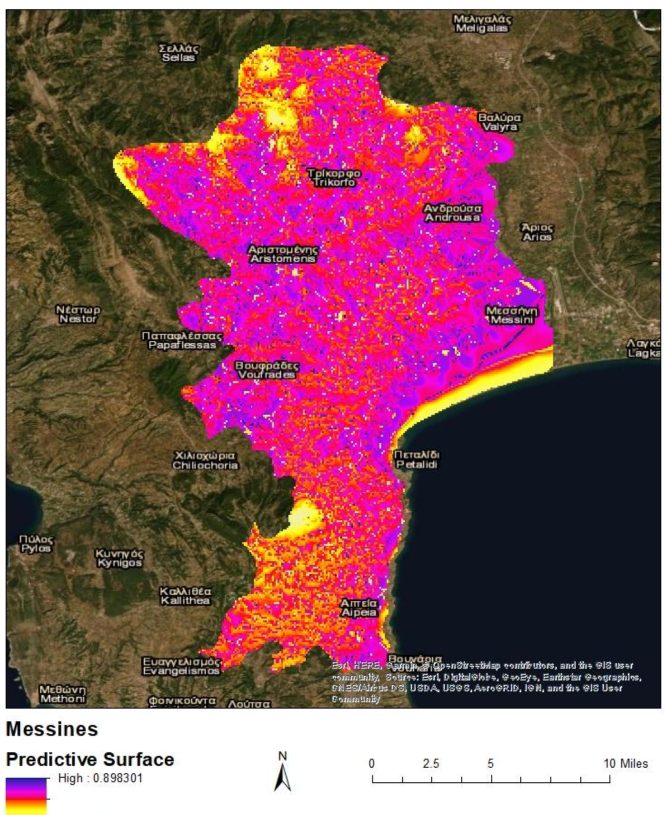

Figure 12 Messines Predictive Surface

4.4 Gortunias

The model of the inland region of Gortunias ran with 68 settlement test points, all of which were given the same 'species type' of Gortunias_Point. This study area included river distance features, elevation, slope, and aspect features.

The ROC curve for Gortunias returns a mean of (AUC = 0.691). This model's performance was about average, with moderate confidence in the performance of this model being better than random.

Of the variables included in this model, elevation had the greatest impact on the occurrence of settlement points with a high permutation importance value of 34.7. The next two variables, slope and northness, had high permutation importance values of 27.8 and 26.8, respectively. The other aspect variable, eastness returned a relatively low but still meaningful permutation importance value of 9.5. The last remaining variable, distance from rivers, held a low permutation importance value of 1.4, meaning that it contributed little to this model. Overall, the model was just outside of an AUC that would have rendered its performance average, so making its results unreliable.

The prediction surfaces generated for Gortunias were more consistent, with high probability areas scattered across contiguous areas. The highest probability areas were distributed around river valleys, with very low probability areas around high altitude areas far from freshwater sources. This can be seen below in Figure 13.

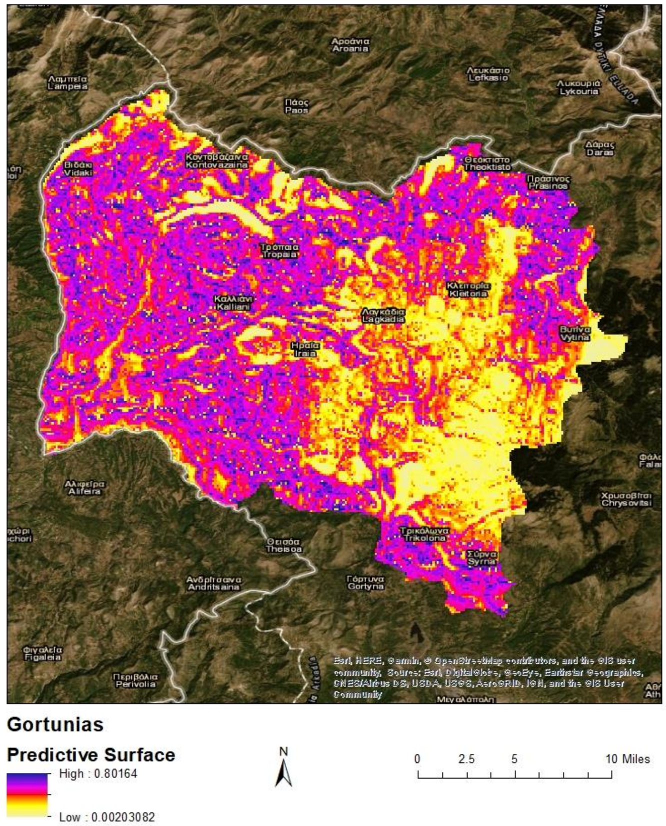

Figure 13 Gortunias Predictive Surface

4.5 Kefalonios

The model of the island of Kefalonios ran with 70 settlement test points, all of which were given the same ‘species type’ of Kefalonios_Point. This study area included coast distance features, elevation, slope, and aspect features.

The ROC curve for Kefalonios returns a mean of (AUC = 0.749). This model’s performance was above average, with high confidence in the performance of this model being better than random. Of the variables included in this study, distance from the coast had the greatest impact on the occurrence of settlement points with a high permutation importance value of 57.4. The next, elevation, had a relatively high permutation importance level at 19.5. The next variable, slope, had a medium permutation importance value of 16.5. The next two variables, eastness and northness, contained relatively low permutation importance values at 3.4 and 2.8, respectively. Overall, distance from the coast proved to be of highest importance, while slope and elevation maintained significant levels of contribution.

The prediction surfaces generated for Kefalonios were fairly consistent with high probability areas concentrated around lower elevation valleys closer to the coast, as seen below in Figure 14. Low probability areas were gathered around steep, high elevation areas far from the coast. The average performance of the model at AUC = 0.749 means these results can be treated as reliable.

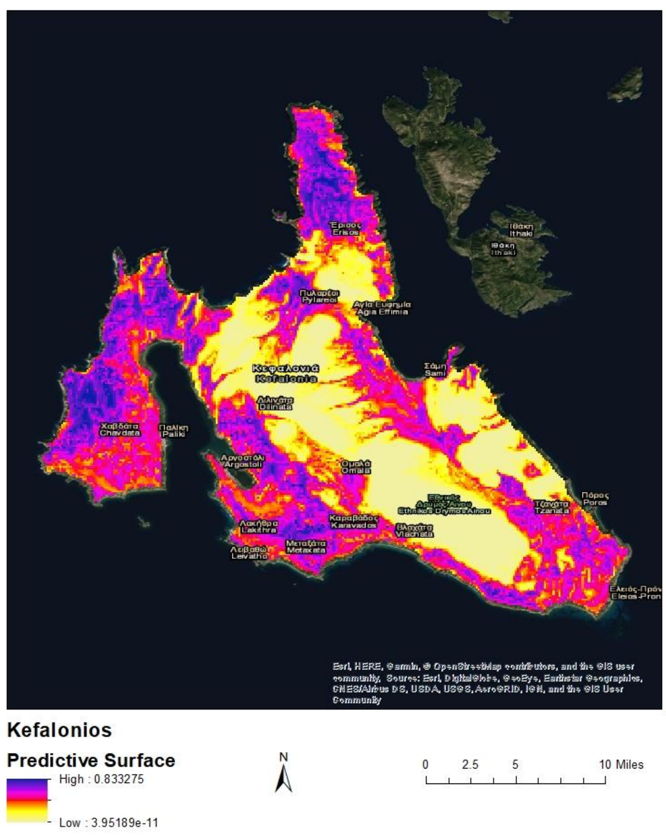

Figure 14 Kefalonios Predictive Surface

4.6 Evaluation of Models

There was a great variance in model successes when compared region to region. In this section, the fit of the models is evaluated in conjunction with the data and parameters provided for the models.

4.6.1 Data Fit

The data used in this study was the most comprehensive available and successfully drew conclusions, unstable though the overall models were. The model covering the entirety of the Peloponnese proved to be highly unstable, with a low AUC value that indicated high unreliability in the results. The analysis of the three administrative regions also provided data contrary to the hypothesis, in that each individual region had its own hierarchy of contributing variables, with few being reliable enough to establish a definitive ranking. As such, this portion of the study must be decidedly inconclusive.

This inconclusive nature must be carried forward to the predictive surfaces. However, these are less inconclusive, as these surfaces paint with broader strokes. The relative probability remains applicable, with Kefalonios performing with the highest AUC and most reliable results. Gortunias performed slightly below average, and Messines performed poorly.

4.6.2 Study Area Fit

As two of the models in more focused areas provided more reliable results with a higher AUC value than the model covering the entirety of the Peloponnese, it can be deduced that reducing the scale of the study area helped in creating a more useful model. The variance between results in different study areas implied a different fit of the model based on location. Generally low AUC scores, except for an average score in Kefalonios, showed general instability across the models. Though with less reliable results, overall the breaking of the study area into three smaller ones allowed for a comparison between regions that otherwise would not have been possible. A future study would account for this and work to find the optimal study area to ensure maximum model stability and insight.

Table 2: Summary of Highest Contributing Variables (Permutation Importance)

Region	Highest contributing variable	Second highest contributing variable	Third highest contributing variable
Peloponnese	Elevation (PI 54.9)	Distance from River (PI 24.4)	Distance from Coast (PI 10)
Messines	Eastness (PI 39.5)	Elevation (PI 26.5)	Northness (PI 23)
Gortunias	Elevation (PI 34.7)	Slope (PI 27.8)	Northness (PI 26.8)
Kefalonios	Distance from Coast (PI 57.4)	Elevation (PI 19.9)	Slope (PI 16.5)

Table 3: Summary of Highest Contributing Variables (Percent Contribution)

Region	Highest contributing variable	Second highest contributing variable	Third highest contributing variable
Peloponnese	Elevation (PC 60.4)	Distance from River (PC 27.7)	Distance from Coast (PC 6)
Messines	Eastness (PC 53.8)	Northness (PC 19.7)	Distance from Rivers (PC 10.2)
Gortunias	Elevation (PC 38.8)	Slope (PC 30.9)	Northness (PC 21.5)
Kefalonios	Distance from Coast (PC 45.7)	Slope (PC 35.7)	Elevation (PC 9.5)

Chapter 5 Conclusions

The results of the study were inconclusive, but there is a use to be derived from its results. An ethnoarchaeological approach focused on human interaction with the environment as variables retains validity. This chapter covers the limitations of this study, as well as recommendations for future improvements, then concludes with final thoughts.

5.1 Initial Conclusions

The results of this study made it difficult to draw definitive conclusions about settlements as a whole, especially as there remained significant variation between regions. The low AUC values meant it is hard to state any general findings with certainty. Each model needs to take into account the varying geographies, and so models may need to be slightly varying as well to match these different features. These generally low AUC values are proof that one model does not fit all locations.

The first hypothesis stated that the predicted order of covariates, (from large to small) for all areas would be 1. Distance from Rivers 2. Distance from Coast 3. Elevation 4. Slope and 5. Aspect (expressed as a combination of Northness and Eastness). The high level of instability in the models reflected by low AUC, as well as the variance between regions, means that these results do not support the hypothesis.

When examining the three sub-regions, distance from rivers remained a minor variable despite being tested in three unique terrain features: a rugged, mountainous region, a broad coastal plain, and an island. Distance from coast, meanwhile, played a large role only on the island; a reasonable occurrence in conjunction with strong maritime traditions and necessity on the sea for resources, survival and trade. The creation of these variables, as path distance with terrain features integrated, displays how variances in elevation - and the resultant impact on people's ability to reach rivers or the coast - are contained within the variables and so within the final results of the models.

The model of the coastal region of Messines showed that the aspect rasters were the most important along with elevation, though the high instability of this model makes drawing definitive conclusions difficult. As Messines is a broad coastal river plain, it can be inferred that these aspect rasters had such high impact because of their effect on what vegetation and crops grew in the area. The people settling in Gortunias, an inland region, valued elevation and slope, indicating a rugged region. From the relative importance of elevation and slope, it can be inferred that plots of land with favorable elevation and slope characteristics were of relative scarcity in this region. Building at a certain slope and elevation, at a higher vantage point for example, would also allow for greater defensibility of the settlement. Finally, it appears that the people settling in the island region of Kelafonios valued proximity to the coast given that the distance to coast was the most important variable included in the model. This is a reasonable inference as it would also allow for contact with the outside world and access to maritime industries.

5.2 Limitations

The largest issue in the construction of the models was the triple counting of variables all related to elevation. Derivation of slope from the DEM weighted the models heavily towards these variables. This was further exacerbated by the inclusion of the DEM as the surface which was utilized in the Path Distance tool to create the walking time from rivers and walking time from coast variables, essentially triple weighting elevation in all the models. Despite this discrepancy, some studies have shown that there is low difference between collinear and non-collinear datasets for methods such as Maxent. The poor performance of the models, though, indicates that the high degree of collinearity in this study negatively impact the reliability of the results. Further research will be required to determine the validity of these studies and derive an effective method to include multiple related variables while maintaining model accuracy.

Finally, a large limitation on the results was the inconclusive nature of the study. Interpreting the results of this study must be tempered with the low AUC values indicating low model reliability. Despite these low AUC values, the comparison between different types of regions, as well as between scales of study areas, gives any future work a framework to work from to refine these models.

5.3 Recommendations

This study brought a new approach to archaeological analysis by utilizing a human-centered approach that used walking time as a variable rather than simple distance. Future work can continue to build on this human centered element by consulting with historians and archaeologists to construct future models. These professionals could contribute a more refined understanding of the variables that influence settlement. This would allow for an inclusion of these variables in the Maxent modeling, so narrowing the scope and potentially increasing the stability of the model.

When comparing the results of the smaller study areas to the results of the entirety of the Peloponnese, the smaller areas performed better, except for the highly unstable Messines. In this instance, a larger study area with more data points did not sharpen the results. Focusing the study area proved more beneficial both for obtaining more stable results and so more reliable predictive surfaces.

With this in mind, the models would be sharpened with the inclusion of more environmental variables. Data such as soils, land cover, and data such as temperature and rainfall would certainly improve the performance of the models. Anyone undertaking this study and utilizing this information would have to be sure to consult with historians and archaeologists to fully understand the context of these variables in order to properly utilize them. They would also have to effectively manage this data across a small area with a low data range.

Future work can be continued with this human-centered approach, and improved variables of human interactions to the environment, such as walking time from a river rather than distance, can be adopted in the new model. This would ensure that this fundamentally anthropological study would remain grounded in the people it aims to understand.

5.4 Conclusions

This study aimed to both quantify the contributions of environmental variables of the settlement of Ancient Greeks and to create predictive surfaces for archaeologists to identify areas with the highest probability of settlement. The first of these goals was inconclusive, as the models were mostly unreliable. The evidence proved contrary to the hypothesis, as each different region had different variables of importance. The models of Kefalonios, the island region, prioritized distance from the coast and slope. Island dwellers would settle in ways that would allow them access to the outside world and maritime industries. The models of the inland region of Gortunias prioritized elevation and slope. The settlers of this rugged region would prioritize areas that offered high defensibility. And the coastal region prioritized aspect, as aspect affected the vegetation that could grow on a large coastal plain.

The second goal was a limited success, providing areas that archaeologists could use to identify areas of high probability though the underlying models that produced them were unreliable. A concerted effort by archaeologists and historians could greatly contribute to future work by providing more data and updated historically relevant variables to enhance the validity of the model. This future work would broaden societal understanding of an important civilization and increase chances of a discovery that could revolutionize the way we think about the past and the origins of contemporary societies.

References

Alexakis, Dimitrios, and Apostolos Sarris. 2010. "Environmental and human risk assessment of the prehistoric and historic archaeological sites of Western Crete (Greece) with the use of GIS, remote sensing, fuzzy logic and neural networks." In *Euro-Mediterranean Conference*, pp. 332-342. Berlin, Germany: Springer.

Bevan, Andrew, and James Conolly. 2004. "GIS, archaeological survey, and landscape archaeology on the island of Kythera, Greece." *Journal of Field Archaeology* 29, (no. 1-2): 123-138.

Biers, William R. 1987. *The archaeology of Greece: an introduction*. Ithaca, New York: Cornell University Press.

Bintliff, John L. 1977. *Natural environment and human settlement in prehistoric Greece: based on original fieldwork*. Vol. 28. Oxford: British Archaeological Reports.

Boardman, John. 1980. The Greeks overseas: their early colonies and trade. London: Thames and Hudson.

Boardman, John, Jasper Griffin, and Oswyn Murray, eds. 1991. The Oxford History of Greece and the Hellenistic World. Vol. 1. Oxford Paperbacks.

Carrer, Francesco. 2013. "An ethnoarchaeological inductive model for predicting archaeological site location: A case-study of pastoral settlement patterns in the Val di Fiemme and Val di Sole (Trentino, Italian Alps)." *Journal of Anthropological Archaeology* 32, (no. 1): 54-62.

Clogg, Richard. 2013. A concise history of Greece. Cambridge University Press.

Cohen, Getzel M. 1995. The Hellenistic settlements in Europe, the islands, and Asia Minor. Vol. 17. Univ of California Press.

De Marco Júnior, Paulo, and Caroline Corrêa Nóbrega. "Evaluating collinearity effects on species distribution models: An approach based on virtual species simulation." PloS one 13, no. 9 (2018): e0202403.

Elleman, Bruce A. 2017. "The Sea and Civilization: A Maritime History of the World by Lincoln Paine." Journal of World History 28, no. 2: 306-309.

Farinetti, Emeri. 2009 "Boeotian landscapes. A GIS-based study for the reconstruction and interpretation of the archaeological datasets of ancient Boeotia." PhD diss., Faculty of Archaeology, Leiden University.

Franklin, Janet, and Jennifer A. Miller. 2014. “Mapping species distributions: Spatial inference and prediction.” Cambridge: Cambridge University Press.

Fisher, Nick, and Hans Van Wees, eds. 1998 *Archaic Greece: new approaches and new evidence*. ISD LLC.

Graham, A. J. 1971. “Patterns in Early Greek Colonisation.” The Journal of Hellenic Studies 91: 35–47. doi:10.2307/631368.

Hale, John R. 2009. Lords of the sea: The epic story of the Athenian navy and the birth of democracy. Penguin.

Hammond, Nicholas Geoffrey Lemprière. 1959. “A History of Greece to 322 BC.” Oxford: Clarendon Press.

Hansen, Mogens Herman. 2006. Polis: an introduction to the ancient Greek city-state. Oxford University Press.

Hartigan, Karelisa. 1995. Greek tragedy on the American stage: ancient drama in the commercial theater, 1882-1994. No. 60. Greenwood Publishing Group.

Jones, R. E., and A. Sarris. 2000. "Geophysical and related techniques applied to archaeological survey in the Mediterranean: a review." *Journal of Mediterranean Archaeology* 13, (no. 1): 3-75.

Kalinski, Cass E. 2019. “Building Better Species Distribution Models with Machine Learning: Assessing the Role of Covariate Scale and Tuning in Maxent Models.” Master’s thesis, University of Southern California.

Kruger, Erin. 2012 "An Introduction to Ancient Greece and the Birth of Drama."

Langston, Lucinda M. 2013. "Site location modeling and prehistoric rock shelter selection on the upper Cumberland plateau of Tennessee."

Lavas, G. P. 1974. "Settlements in ancient Greece." Ekistics: 330-335.

Lecours, Vincent, Craig J. Brown, Rodolphe Devillers, Vanessa L. Lucieer, and Evan N. Edinger. 2016. "Comparing selections of environmental variables for ecological studies: A focus on terrain attributes." PloS one 11, no. 12: e0167128.

Lock, Gary R., and G. Stancic, eds. 1995. *Archaeology and geographic information systems: A European perspective*. CRC Press.

Mehrer, Mark W., and Konnie L. Wescott, eds. 2005. *GIS and archaeological site location modeling*. crc Press.

Phillips, Steven J. 2005. "A brief tutorial on Maxent." AT&T Research.

Raaflaub, Kurt A., Josiah Ober, and Robert Wallace. 2007. Origins of democracy in ancient Greece. Univ of California Press.

Snodgrass, Anthony M. 1992. *An Archaeology of Greece: the present state and future scope of a discipline*. No. 53. Univ of California Press.

Tsetskhladze, Gocha R., ed. 1999. Ancient Greeks west and east. Brill.

Vaughn, Sallie, and Tom Crawford. 2009. "A predictive model of archaeological potential: An example from northwestern Belize." *Applied Geography* 29, (no. 4): 542-555.

Verhagen, Philip, and Thomas G. Whitley. 2012. "Integrating archaeological theory and predictive modeling: a live report from the scene." *Journal of Archaeological Method and Theory* 19, (no. 1): 49-100.

Whitley, Thomas G., and Gwendolyn Burns. 2008. "Conditional GIS surfaces and their potential for archaeological predictive modelling.": 292-298.