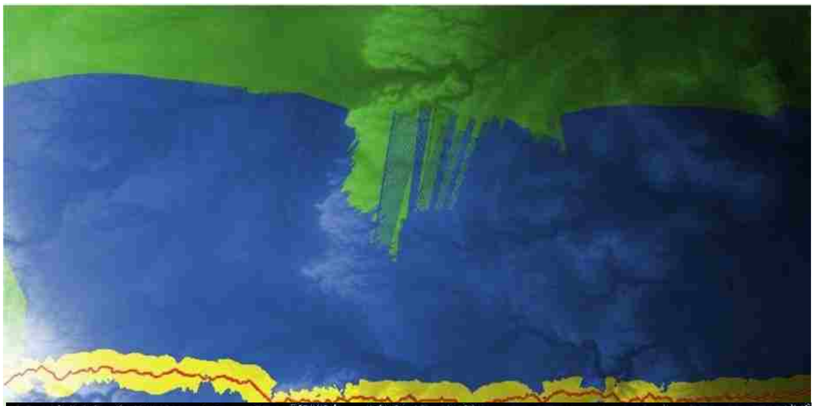
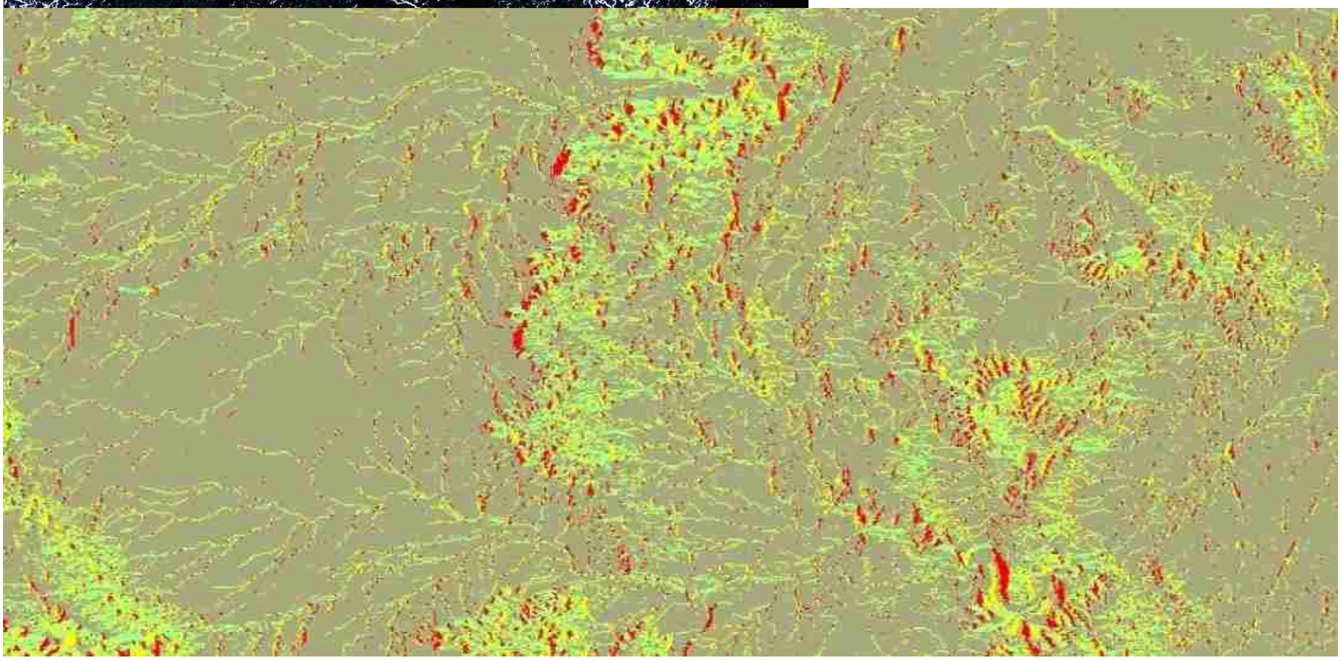


Agent Based Predictive Models in Archaeology

Thesis for PhD at University of Edinburgh.
Corrections Submitted 2014.

Doug Rocks-Macqueen



Abstract

For over 40 years archaeologists have been using predictive modelling to locate archaeological sites. While great strides have been made in the theory and methods of site predictive modelling there are still unresolved issues like a lack of theory, poor data, biased datasets and poor accuracy and precision in the models. This thesis attempts to address the problems of poor model performance and lack of theory driven models through the development of a new method for predictive modelling, agent based modelling. Applying GIS and agent based modelling tools to a project area in southeaster New Mexico this new methodology explored possible behaviours that resulted in site formation such as access to water resources, travel routes and resource exploitation. The results in regards to improved accuracy over traditional methods were inconclusive as a data error was found in the previously created predictive models for the area that were to be used as a comparison. But, the project was more successful in providing explanatory reasons for site placement based on the models created. This work has the potential to open up predictive modelling to wider archaeology audiences, such as those based at universities. Additional findings also impacted other areas of archaeological investigation outside of predictive modelling, such as least cost path analyses and resource gathering analyses.

Font Matter and Statement of Work

This thesis was composed by myself – Doug Rocks-Macqueen – as an individual project. This work has not been submitted for any other degree or professional qualification except for this PhD in Archaeology at the University of Edinburgh.

Doug Rocks-Macqueen

Table of Contents

Table of Figures

Table of Tables

Table of Equations

Agent Based Predictive Models in Archaeology Introduction and Outline

This thesis has its roots in a 14-hour drive to the 'Bootheel' area of New Mexico, USA, and some of my experiences while working in Commercial Archaeology in United Kingdom and Cultural Resource Management Archaeology (CRM) in the USA. In 2008, I was working in CRM Archaeology in New Mexico after finishing my undergraduate degree. At one of the excavation projects that I worked on, the landowners of the property asked for the artefacts to be given to them and due to the nature of the project that request was granted. So in the Spring of 2009, my colleague Brian and I got in a truck and drove seven hours down to the 'Bootheel' of New Mexico from our company office in Albuquerque, and seven hours back, all in the same day.

During that 14-hour drive many conversations were had, but one conversation in particular contributed to my undertaking of this thesis – Brian said that I should enjoy our work now, because in 30 years all of the archaeological sites would have been found. This statement struck me as something that did not seem right. 'Surely we have barely scratched the surface of archaeology', was what I thought. In the truck we did not have the resources to confirm or refute this hypothesis and so it was left at that.

Yet, that conversation had an impact on me because later that year, when I began my Masters studies at Newcastle University in the United Kingdom, I remembered it when I was considering topics to research for my dissertation. 'Had we found most of the archaeology in New Mexico?' was the question that initially drove my research. To me the simple solution was to take the known archaeological sites and compare that against the estimated undiscovered sites to see how much we still had to find and then to look at the number of sites found per year to determine the pace of discovery. With that information I could determine if we were quickly running out of archaeology. The known sites were the easy part, as they were stored in a database; it was attempting to find estimates of the unknown sites that was difficult as no one had made such estimates. Eventually my research led me to site predictive modelling, computer models that estimate where undiscovered archaeological resources are, as a solution to my problem. Predictive modelling would form the basis of my Masters' dissertation, but that dissertation focused on a different question as I moved on to more complex issues than how much is left to find.

A comment on a long drive had brought me to site predictive modelling but it would be my experience with CRM in the United States and Commercial Archaeology, as CRM is known as in the UK, which would provide the drive behind this thesis. I have always had an eye on a career outside academia. When I undertook my Masters I did so because a Masters is usually needed to get a permanent CRM position in the United States. I knew at that time that academic positions were few and far between, but it would be years later that I would discover exactly how hard they are to get (Rocks-Macqueen 2016), so I did not intend to undertake a PhD as it seemed unnecessary to me at that time. Instead of focusing on a career in academia I found opportunities to broaden my CRM experience. One of these experiences was working with a Council Archaeologist in the UK. I learnt about the planning process in the UK and was shocked by some of the practices I observed. Planning conditions were placed on development projects by first looking at the Historic Environment Record and if archaeological resources had been found previously in the general area, as estimated by eyeballing the map, then a planning condition would be recommended for archaeology. If nothing had been found in the area, then no planning condition would be recommended. The fact that areas could have nothing because no archaeological work had been conducted did not factor in the decision. As I saw it, potentially lots of archaeology was simply being bulldozed because there was no better system in place. Given my experience with site predictive modelling, I saw the potential for this method to help solve these issues. While not perfect, predictive modelling, in my opinion, could give better results than eyeballing a map.

Further influential experiences occurred through my CRM work in the United States. During one project we surveyed the routes for several access roads and found archaeological resources. Later, it was deemed too expensive to excavate what we had found. Instead, a proposal was put forward to survey the land around the proposed routes and then redirect the access roads around the archaeological sites to avoid a costly excavation. The company had done this before by sending out crews with an engineer to survey a route until they found a site, then reroute around the site and continue, basically blindly surveying until they could find a route without sites. This process takes up resources and time. Again, site predictive modelling seemed well posed to help solve this by helping plan out potential routes without archaeological resources before the survey began.

Another experience that highlighted the advantages of predictive modelling was project planning for surveys in the western United States. Projects were typically planned by taking the number of acres to be surveyed and then dividing that by an estimated number of sites that would be encountered per acres surveyed. This would give roughly the number of sites planners could expect to encounter and have to record, helping them get an estimate of time/costs. The inevitable problem with such a method is that if more sites were discovered than anticipated the project would overrun its budget. Conversely, if a project over-estimated sites and budgeted too much they might lose out in any competitive tendering as the bid would be too high. Site predictive modelling in my mind had the potential to better estimate the archaeological resources encountered and thus be able to improve project cost planning in cases with the need for inflexible budgets.

Once I had been exposed to site predictive modelling I could see a range of potential applications in CRM archaeology where it was not currently being utilised. However, my research found that predictive modelling had been hyped in the past and failed to deliver its promised returns (Chapter 1). I also knew there were numerous problems with predictive modelling (reviewed in Chapter 2) and that these problems had kept site predictive modelling from having wider use. With my experience and background knowledge, though limited as it was in 2010, when I only had a few years of archaeological experience, I came to the conclusion that research was needed to see if predictive models could be improved enough to be of greater use for archaeologists.

Thesis Aims, Objectives and Activities

Given the background that led to this project the following aim was set on the outset of this project — **improve the performance of site predictive modelling for CRM archaeology uses**. To reach that aim the following objective was set for this project – **increase the explanatory abilities of predictive modelling**. When researching how to improve performance, I found that adding more theory/explanatory aspects to predictive models, which have typically been correlative and theory devoid, improved the model's performance. I saw this as a route worth exploring for this research project. To achieve that objective a new methodology for creating predictive models, agent-based modelling (ABM), was tested. It was also realised, by myself and others, that such a methodology could attract more academic interest in predictive modelling (Chapters 1 & 2), a bonus benefit but one which was not the primary focus of this project.

To achieve that research aim the following activities were planned for at the beginning of the project:

1. research causes of poor model performance and find cause to address;
2. create a methodology to solve the problem(s) that lead to poor model performance;
3. test proposed solution(s);
4. compare solution results against independent models to determine their effectiveness.

Thesis Outline

This thesis reviews the process undertaken to meet that goal and objective. What follows is the outline of this thesis, which in turns mirrors the process that was undertaken for this project. As will be demonstrated, initial assumptions turned out to be wrong and changes were made throughout the project that altered the process. The thesis is broken down into chapters and each chapter considers a distinct subject which contributes to the whole project. While certain chapters may be read out of order, for example, skipping the background and moving on to model testing, the thesis is intended to be read in sequential order. It lays out in order the process of research undertaken for this project.

The first chapter starts with definitions and key concepts so that readers will have a clear understanding of what is meant by certain terms. After that it reviews the historical development of site predictive modelling in archaeology. It follows the early development of the theory underpinning predictive modelling into the post-modernist critique of such work and up to current developments. At the same time, developments in methodology and technology are reviewed to create a holistic view of the expansion and waning use of predictive modelling over the last few decades. This provides the historical context in which this project began.

The problems that this project will tackle are outlined in Chapter 2. This outlines criticisms of predictive modelling, from theoretical, methodological, technological or cultural standpoints, and then examines known solutions to these problems. It will show that several problems remain outstanding in predictive modelling. The chapter ends with a review of the decisions made, and their justifications, as to why this project concentrated on certain problems and not others. Specifically, the lack of contribution most predictive models made to our understanding of the past. As will be shown, the literature indicated that improving models explanatory abilities would increase its performance.

Chapter 3 introduces the proposed solution of using a new methodology: Agent-based modelling. Firstly, the precursors to the use of ABM in predictive modelling are reviewed and how they have indicated that ABM could present a viable solution to the issues. An examination of recent developments by other researchers in predictive modelling that occurred during the project is also conducted. These developments present an even stronger case for the use of ABM in predictive modelling.

Chapter 4 examines a range of ABM programmes that could have been used for the project. This chapter lays out the rationale used in deciding which features were most important in an ABM programme for use in this project. The list of required features is then used to find the best ABM to use for this project, in this case, NetLogo. This chapter also briefly scrutinises some of the other programmes that met the required criteria but were not chosen, and the reasons why they were not chosen for this project.

The fifth chapter is an examination of the project area in southeastern New Mexico, USA, used to test the project’s methodology. This chapter outlines why the project area was chosen. It then provides relevant background information on the project and surrounding areas, ranging from the cultural history to the environment of the area. It provides a general review of factors that were key components required for the creation of the model.

This is followed by three chapters, 6, 7 and 8, which discuss the application of ABM in different contexts. Chapter 6 looks at site distribution as a result of travel routes across the project area and indicates that least-cost-path analysis is deeply flawed. Furthermore, this chapter will show that, in the project area, site location does not appear to be related to long-distance travel routes. The results of this chapter are not just applicable to predictive modelling but to other projects that attempt to look at human mobility.

Chapter 7 reviews the hydrology of the project area and concludes that site location is not based on access to water. It also provides significant information about how past peoples interacted with the landscape and what activities they could have undertaken in the project area. Essentially, it will show that a lack of perennial water sources limited how long people could stay in the project area. It will also demonstrate a link between water availability and the activities available in the project area that may have attracted people to it, as discussed in the next chapter.

Chapter 8 examines different resources that could attract people to the project area, such as lithic quarries, plant resources and shelter. This chapter will show that these different resources could have acted to draw past peoples to the project area. The modelling will provide both valuable information about site locations and potential site locations as well as information about the cultural history of the area.

The final chapter summarises and discusses the results of the preceding three chapters' investigation. It then draws together these results into the solutions sought by this project and attempts to assess the overall effectiveness of ABM as a solution to some of the problems facing predictive modelling.

Appendices present various datasets and the code used for the models.

Confidential appendix

The archaeological site data used for this project is from the Archaeological Records Management Section (ARMS) in New Mexico and comes with the stipulation that:

> 'All information obtained pursuant to the user's access to the HPD/ARMS system shall remain confidential unless subject to public disclosure pursuant to the New Mexico Cultural Properties Act of 1978 [NMSA 1978, §§18–6-1 to 18–6-17], and the Archaeological Resources Protection Act of 1979 [16 U.S.C. §470aa *et. seq.*], and anyone misusing such information may be subject to prosecution under the Federal Computer Security Act of 1987 [15 U.S.C. §§271, 272, 278g-4, and 278h].'

Consequently, limits are imposed on the information that can be shared publicly, in particular site locations. These restrictions are intended to prevent the looting of archaeological sites in New Mexico. For this study all site locations have been removed from any datasets found throughout this thesis, including any reference to sites on maps. Many figures have purposely been made vague to ensure that known sites or potential site locations, as this project deals with site predictive modelling, are not discoverable.

A confidential appendix has been created for this project containing more detailed images and sensitive information. This appendix will not be published with this thesis but researchers can obtain this appendix by contacting the author and demonstrating they will be using the information for reputable purposes. Furthermore, should the author be uncontactable the full list of sites by LA number, the reference number used by ARMS to identify sites, used in this project has been included so that the exact site datasets may be obtained from ARMS. The locations of where all other datasets can be obtained are given and the code for the ABM models included, which makes it possible to recreate the results of this project once one has obtained the site data from ARMS.

Chapter 1: Site predictive models: background and historical development

The purpose of this chapter is to introduce the concept of site predictive modelling, discuss its origins in archaeology and the direction of its development when this project started in 2010. The defining of site predictive modelling and its associated concepts ensures clarity when using the terms throughout this work. The history and trajectory of site predictive modelling in archaeology will give context to the problems facing predictive modelling when this project began in 2010. It will show the historic situation in which this project was developed – low usage and scepticism about its benefits. Moreover, some of the issues raised with predictive models, and discussed further in Chapter 2, are related to the history of its development.

Defining Predictive Modelling

The way one interprets or defines site predictive modelling, as with many different terms, can be highly subjective, e.g. a 'long road' could be 1km, 10km, 100km, etc. In the interest of clarity, it is best to define what the term 'site predictive modelling' means for this project. There have been a plethora of definitions of what a site predictive modelling is or should be:

> '... a predictive archaeological locational model may simply be regarded as an assignment procedure, or rule, that correctly indicates an archaeological event outcome at a land parcel location with greater probability than that attributable to chance.' (Kvamme 1990 p. 261)

> '… simplified set of testable hypotheses, based either on behavioural assumptions or on empirical correlations, which at a minimum attempts to predict the loci of past human activities resulting in the deposition of artefacts or alteration of the landscape.' (Kohler 1988 p. 33)

> 'Models that are deductively derived and attempt to predict how particular patterns of human land use will be reflected in the archaeological record' or 'identify and quantify relationships between archaeological site locations and environmental variables.' (Judge and Sebastian 1988 p. 4)

> 'Predictive models are tools for projecting known patterns or relationships into unknown times or places.' (Asch and Warren 2000 p. 6)

These are but a few of the many ways archaeologists have defined site predictive models. A definition used by several archaeologists (Ebert 2004, Kamermans, van Leusen et al. 2009b) is Kohler and Parker's:

> 'Predictive modelling is a technique to predict, at a minimum, the location of archaeological sites or materials in a region, based either on the observed pattern in a sample or on assumptions about human behaviour.' (Kohler and Parker 1986 p. 400)

Essentially, all of these definitions say almost exactly the same thing— that a predictive model's purpose is to predict the location of unknown archaeological resources, however one defines archaeological resources, e.g. sites, landscapes, tombs, monuments, Roman villas, etc. This is how this project will define site predictive modelling as a process for determining the potential locations of undiscovered archaeological resources.

Site Predictive Model or Archaeological Resource Predictive Model or Predictive Model?

Most literature on archaeological predictive modelling refers to the models as 'site predictive models'. However, as pointed out by some predictive modellers, 'site' is an arbitrary term that changes definition from project to project, or location to location, or even author to author (Ebert 2000). A more accurate term would have been 'archaeological resource predictive model', as this does not explicitly define the unit of measurement of the archaeological resources that are being modelled. However, given the historic use of the term by other authors, the term 'site predictive model' will be used here to ensure clarity and continuity of terminology. Moreover, throughout this thesis 'site predictive modelling' is shortened to just 'predictive modelling'.

While I agree with much of the criticism of the concept of 'sites' – that it entails arbitrary boundaries and does not capture the full range of past human behaviour – this is how the archaeological data for the project area is provided. As such, this thesis uses 'sites' as a unit of the archaeological resources that are tested for in this project.

Contextualising the History of Predictive Modelling

The historical development of predictive modelling has followed the 'Gartner Hype Cycle' (Figure 1).

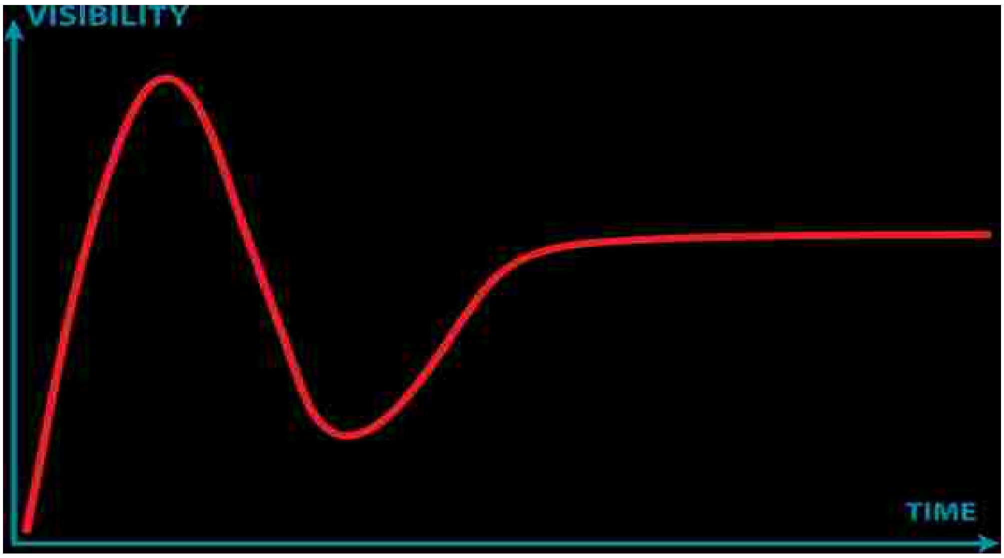


Figure 1: Graphical representation of the Gartner Hype Cycle (Wikipedia: http://en.wikipedia.org/wiki/File:Gartner_Hype_Cycle.svg Creative Commons Attribution-Share Alike 3.0. Author Jeremykemp)

The Gartner Hype Cycle, named after research company Gartner (Gartner 2008), is a cycle that many emerging technologies or new processes go through. It consists of five phases:

1. 'Technology Trigger' — the first phase of the hype cycle is the 'technology trigger'. It is a breakthrough, new application or event that generates significant interest in the technology/technique/process.
2. 'Peak of Inflated Expectations' — the second phase is significant publicity about the new development in question. This 'hype' generates over-enthusiasm and unrealistic expectations of the new development. During this time there may be some successful applications of a technology, but there are usually more failures.
3. 'Trough of Disillusionment' — this phase occurs when the new development fails to meet the inflated expectations that occurred in the 'Peak of Inflated Expectations'. This usually results in the development quickly becoming unfashionable.
4. 'Slope of Enlightenment' — even though many people will have stopped working with the technology, some people or organisations will continue to work on the subject through the 'slope of enlightenment'. They typically experiment in attempts to understand the benefits and practical application of the technology.
5. 'Plateau of Productivity' — this is the phase where the benefits of it become widely demonstrated and accepted. The technology becomes increasingly stable and evolves in second and third generations. The final height of the plateau varies according to whether the technology is broadly applicable or benefits only a niche market.

The development of predictive modelling mirrored this cycle, as will be demonstrated in the rest of this chapter. However, it should be noted that the Gartner Hype Cycle is not reflective of all possible routes. As Newman and Lee (2012) have discussed, the Gartner Hype Cycle should actually have an additional drop at the end when a technology, method or process eventually becomes obsolete (Figure 2). There are also technologies that will never reach certain stages of the cycle.

Figure 2**:** Gartner Hype Cycle with Newman and Lee 'slope into obsolescence'.

The First Landscape Archaeology

Site predictive modelling evolved from the culmination of decades of work by various individuals and organisations. The first step in the creation of site predictive modelling was for archaeologists to identify wide-scale patterns between site locations and characteristics of the landscape, either cultural or environmental. Most modern predictive modellers attribute the start of predictive model theory (Kohler 1988, Dalla Bona 1994, van Leusen 2002, Altschul, Heidelberg et al. 2004, Altschul, Hayden et al. 2005, Canning 2005, Kamermans and van Leusen 2005, Kvamme 2006, Mount and Schwarz 2006, Verhagen 2007d) to Willey's work in the Virù Valley of Peru (Willey 1953, Willey 1956), though some mention is given to Julian Stewart's work as an inspiration. Willey undertook large-scale landscape surveys, and established that there were correlations between site locations and environmental features (Willey 1953, Willey 1956).

While much of the literature credits Willey, in the 1950s, as the first to recognise patterns in site placement, similar work identifying patterns to site placement was accomplished in the United Kingdom decades before (Crawford 1912, Fleure 1916, Crawford 1921, Fox 1923). These earlier endeavours are not mentioned in any predictive modelling literature. This could be because the early development of predictive modelling was undertaken primarily by North American archaeologists, as will be discussed further in this chapter (Bradt, Groenewoudt et al. 1992 p. 268; Garcia Sanjuan and Wheatley 1999 p. 215; Kamermans 2007 p. 72), who may not have been aware of these earlier developments outside of their geographic area of work when they created the first history of predictive modelling.

Environmental Determined Predictive Models

Willey's 'first discovery' started what has been termed an 'ecological approach' to site placement (Verhagen 2007d). This term refers to the idea that the environment influences the location of sites. This approach stresses the environmental characteristics of a landscape, e.g. rivers, hills, etc., which determine where archaeological sites are situated. In the last few decades this perspective has come under scrutiny as both limiting the potential of predictive models and not fitting into the wider interests of archaeologists (Brandon, Burgett et al. 2000, Ebert 2000, Wheatley 2004, Harris and Lock 2006).

After Willey first made the connection between site locations and environmental factors several theories and analytical techniques developed this idea further (Kohler 1988, Dalla Bona 1994, van Leusen 2002, Altschul et al. 2004, Altschul et al. 2005, Canning 2005, Kamermans and van Leusen 2005, Verhagen 2007d). One such concept was Chisholm's (1962) geographical location theory, which adopted many of the concepts laid out by Isard (1956). This would be followed by site catchment theory (Higgs and Vita-Finzi 1972), which tried to establish the rules that determine human spatial behaviour, advanced from the perspective of subsistence economy. Appearing a little later was Jochim's publication, *Hunter-Gatherer Subsistence and Settlement: A Predictive Model* (Jochim 1976), which looked at the location of hunter-gatherer sites. All these theories led predictive modelling to be described in terms of environmental determinism:

> 'Most archaeological predictive models rest on ... [the] fundamental assumption[s] ... that the settlement choices made by ancient peoples were strongly influenced or conditioned by characteristics of the natural environment ...' (Asch and Warren 2000 p. 8)

Landscape Studies

During the 1970s several settlement pattern projects were undertaken, similar to Willey's project. A leading group in the United States concerned with this research was the Southwest Anthropological Research Group (SARG). The goal of SARG was to investigate where archaeological sites were located. They concentrated on environmental factors as the source of site placement and tried to create 'objective measurements' of environmental variables and their relationships to site placement (Hill and Plog 1971). By the 1970s and 1980s this work, and that of others (Bettinger and Thomas 1976, Bettinger 1980, Shermer and Tiffany 1985, Kellogg 1987), had established that, as Willey had found, it was possible to identify that there are significant correlations between environmental factors and site location.

Development of Methodologies

With the knowledge that site locations could be correlated landscape components, such as closeness to water, slope, etc., a method was needed to take these observations and convert them into a predictive model. It was at this time, the 1970s, that the three oldest and most widespread methodologies for predictive modelling were created: Boolean, Weighted and Regression Algorithms. The following sub-section of this chapter will briefly describe these methods for readers who may be new to predictive modelling.

Boolean Models

This method works by first making a hypothesis about where a site/archaeological resource will be located (Gillings and Wheatley 2002). The hypothesis is decided by the designer and some are instinctive/judgement choices (Dalla Bona 2000, Gillings and Wheatley 2002, Canning 2005, Cole, Madry et al. 2006, Kamermans 2006, Kuiper, Leveson et al. 2006, Verhagen 2006, Kay and Witcher 2009) while others use decision rules (Verhagen 2006). Attempts to make a more rigorous protocol for the assignment of values to attributes have resulted in the use of the ratio of expected sites to observed sites (Brandt, Groenewoudt et al. 1992, Johnson 1996, Berger and Verhagen 2007) or even using statistics to gather correlations to guide their choices (Gazenbeek and Verhagen 2007). Other modellers have used hybrid models of both judgement and observed patterns (Verhagen 2006, Clarke, Ford et al. 2009). Regardless of how these hypotheses are made, Boolean mathematics is used to assign one of two numbers, 0 or 1, to each of the two possible outcomes. In the case of site predictive modelling, those outcomes are site or archaeological resources either present (1) or absent (0) (Gillings and Wheatley 2002). More than one hypothesis can be made at a time and can also be combined by multiplying together the assigned numbers (Altschul et al. 2004). Accordingly, the results indicate one of the two possible outcomes (0 or 1) (Equation 1).

$$((x_1, x_2, x_3, \ldots) > 0) = M \quad \text{ex. } (1 \times 1 \times 1 \times 1) = 1 = M$$
$$((x_1, x_2, x_3, \ldots) \leq 0) = M' \quad (1 \times 0 \times 1 \times 0) = 0 = M'$$

Equation 1: Boolean maths for site predictive models.

In practice, this method is fairly straightforward. The model creator makes statements about where sites will be using any of the above-mentioned methods, e.g. sites will be located: within 2km of water; within 1km of a Roman road; only on flood plains. After these statements are made the different data is assigned either a 0 or a 1. Using this example, if sites will be found within 2km of water, a modeller would then say all locations within 2km will be assigned 1. The reverse of this statement then must be assigned 0. So any area that was more than 2km from water should be assigned a 0. This is done with every statement and then applied to each section of the study area (Equation 2).

$$((x_1, x_2, x_3, \ldots) > 0) = M \quad (1 \times 1 \times 1) = 1 = M = \text{site presence}$$
$$((x_1, x_2, x_3, \ldots) \leq 0) = M' \quad (0 \times 0 \times 0) = 0 = M' = \text{no site present}$$
$$(0 \times 1 \times x\ 1) = 0 = M' = \text{no site present}$$

(distance to water ≤ 2km = 1) × (distance to Roman road ≤ 1km = 1) × (land type is floodplain = 1)

(distance to water > 2km = 0) × (distance to Roman road >1km = 0) × (land type is not floodplain = 0)

Equation 2: Boolean maths in site predictive models.

A result of this method is that those areas on a map that meet *all* of the conditions will be labelled 1, indicating the likelihood of a site being present. Those that *do not* meet all of the conditions will be labelled 0, defining the absence of a site. The results of this method can be seen in Figure 3. A definition 0 would follow even if a location were to meet 99 out of 100 conditions. These results are called a ‘multivariate discrimination function’, because it produces a binary, mutually exclusive outcome, making it impossible to have any sort of gradation of where sites might be (Gillings and Wheatley 2002 pp. 152–153). These binary results make it very hard to scale these models to include multiple factors, as the more factors that are added, the more likely false exclusion will occur.

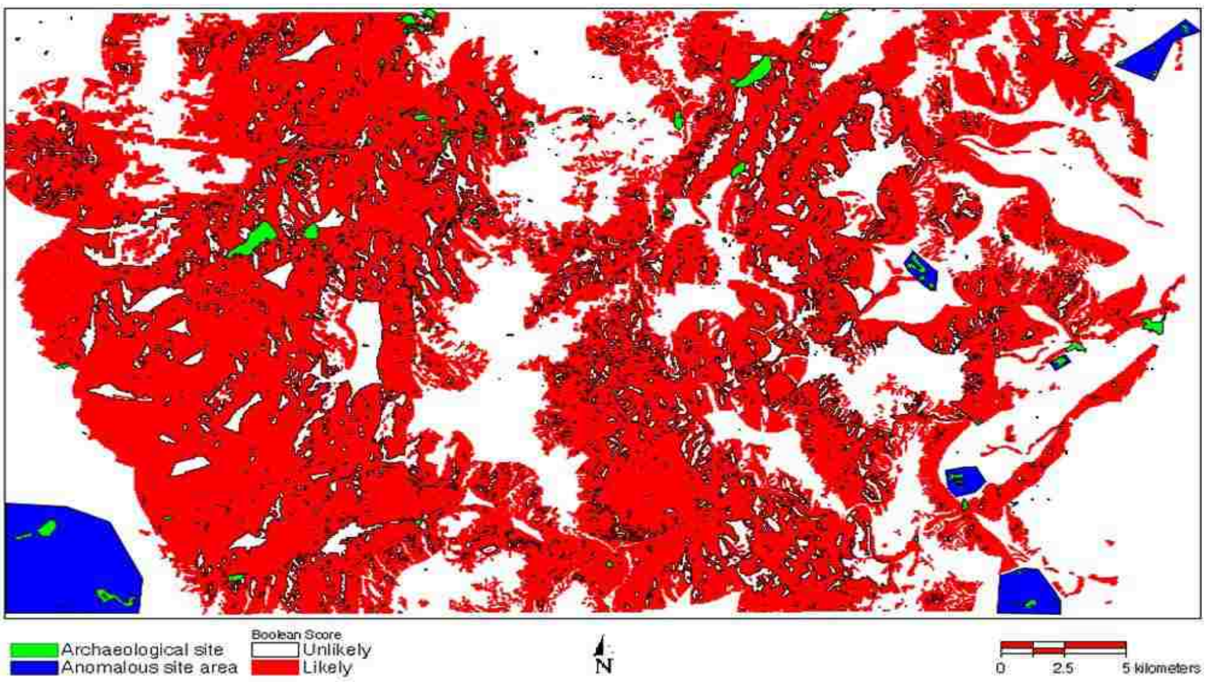


Figure 3**:** Boolean model of Azoto Mesa (Altschul et al. 2005 p. 90 Figure 6.10).

Weighted Models

In this method, the variables used to construct the model (distance to water, slope, etc.) are given weights (ratings) on how likely they are to indicate the location of archaeological resources (Brandt et al. 1992, Dalla Bona 2000, Gillings and Wheatley 2002, Altschul et al. 2004, Kamermans 2006). These weights are then added together to create models with multiple outcomes. Weighted models do not reject a location as suitable if it meets 99 out of 100 of the criteria as a Boolean model would, and because of this are usually considered a better method (Dalla Bona 2000).

How one assigns weights, as with the Boolean method, varies. Some weights are made through instinctive/judgement choices (Dalla Bona 2000, Gillings and Wheatley 2002, Canning 2005, Cole et al. 2006, Kamermans 2006, Kuiper et al. 2006, Verhagen 2006, Kay and Witcher 2009), by decision rules (Verhagen 2006), the ratio of expected sites to observed sites (Brandt et al. 1992, Johnson 1996, Berger and Verhagen 2007) or even using statistics to gather correlations to guide weighting (Gazenbeek and Verhagen 2007). Finally, in some instances, hybrid models of both judgements and correlative patterns have been advanced (Verhagen 2006, Clarke et al. 2009). In these instances, a variety of weighting systems can be used depending on the preference of the model creator. However, once the weighting is done, all the weights for the variables are then added together to obtain a result for the location (Altschul 1990, Dalla Bona 2000, Kay and Witcher 2009).

Location 1= (x1 + x2 + x3) = (.1 + .2 + .5) = .8 = more likely to have sites

Location 2= (x4 + x5 + x6) = (.01 + 0 + .01) = .02 = less likely to have sites

Equation 3: Example of how weighted method works.

As with the Boolean models, this simple addition can be completed using either raster or vector data in a GIS program. Those areas with higher numbers are then labelled as being more likely to have archaeological resources and vice versa for lower probability areas. Figure 4 shows a graphical representation of this, with a grading of red areas being more likely to have sites to green areas where there is very little chance of finding archaeological sites.

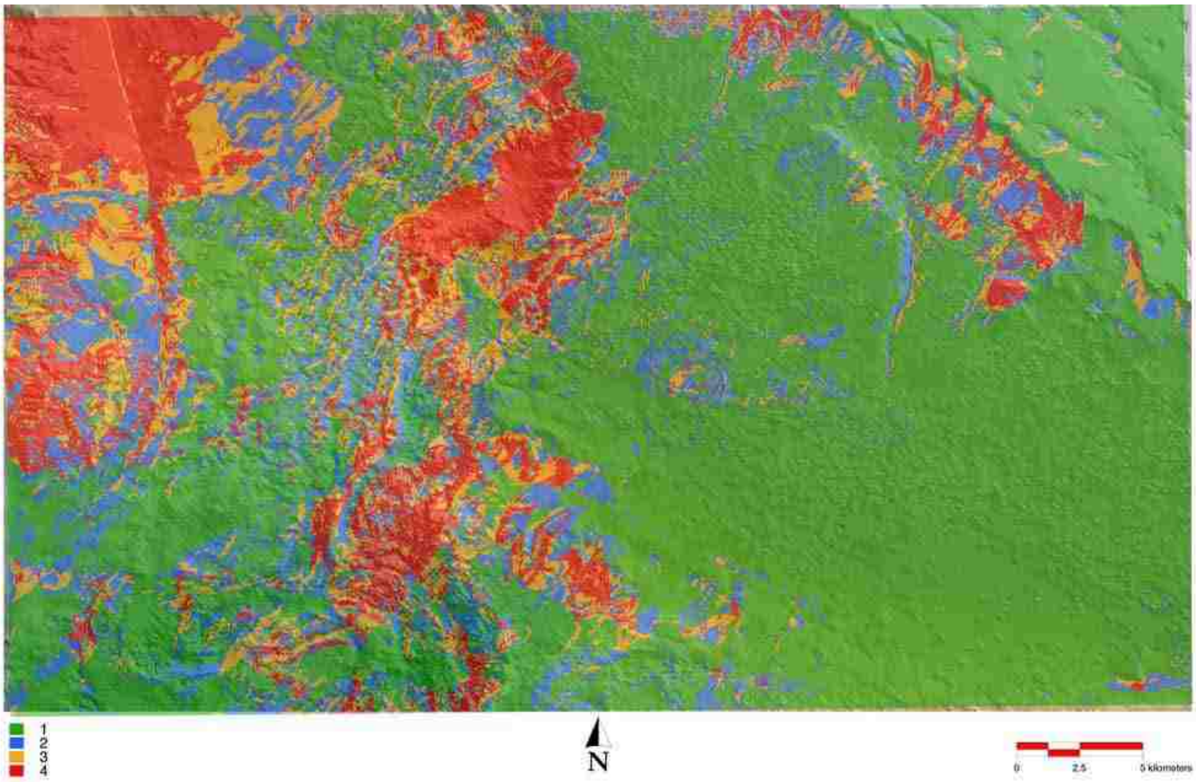


Figure 4: Loco Hills weighted site predictive model with four classes (Altschul et al. 2005 p. 59 Figure 5.10). Class 1 are areas least likely to have sites and class 4 are areas most likely to have sites.

Regression Algorithms

While environmental determinism became one of the persuasive theories of the 1960s and 1970s archaeologists, mainly in North America (Trigger 2006), quantitative approaches became a driving force behind the methodologies of that era (Kohler 1988, Kamermans 2006). Several volumes on spatial analysis (Hodder and Orton 1976, Clarke 1977) were published during this time and heavily influenced the methods of predictive modelling, especially towards the use of statistics (Kohler 1988, Gaffney, Stancic et al. 1996, Kamermans 2006). In the early 1970s, Green was the first to use multiple linear regression on Mayan sites in Belize to create a predictive model (Green 1973).

Since then regression algorithms have been one of the most popular methods for creating predictive models, especially regression algorithms (Cole, Gould et al. 2006, Hatzinikolaou 2006). This method calculates the correlations between **known** sites and different aspects of a landscape (Figure 5), e.g. distance to water, political boundaries (Warren 1990b, Gillings and Wheatley 2002, Cole et al. 2006, Hatzinikolaou 2006). These correlations are then projected onto areas not previously surveyed (Warren 1990b, Asch and Warren 2000, Beckman and Duncan 2000, Gillings and Wheatley 2002, Altschul et al. 2004). Like other methods, the results can be represented graphically in an easy to understand map form (Figure 6). Unlike Boolean or weighted methods, regression algorithms are not easily accomplished by hand and require the use of GIS software (Warren 1990b).

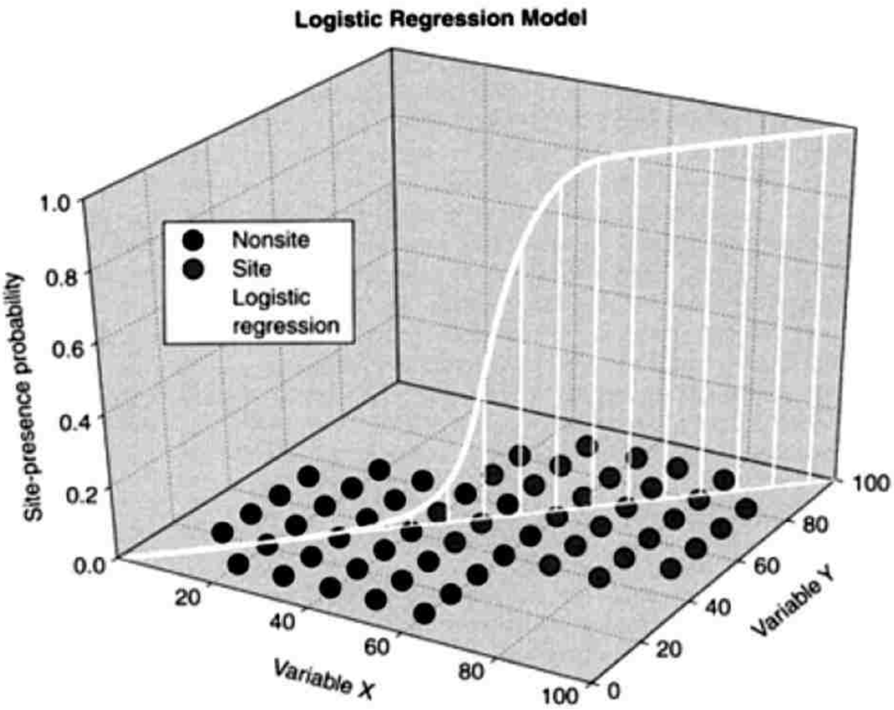


Figure 5**:** Idealised logistic regression of two groups of objects (sites and nonsites) across two independent variables (X and Y) (after Warren 1990a). The line running lengthwise through the horizontal scatter of points is the axis that best discriminates sites from nonsites. The vertical plane is defined by an S-shaped logistic regression line. This line shows an increase in site-presence probability from left to right along the axis of discrimination (Asch and Warren 2000 p. 11 Figure 2.2).

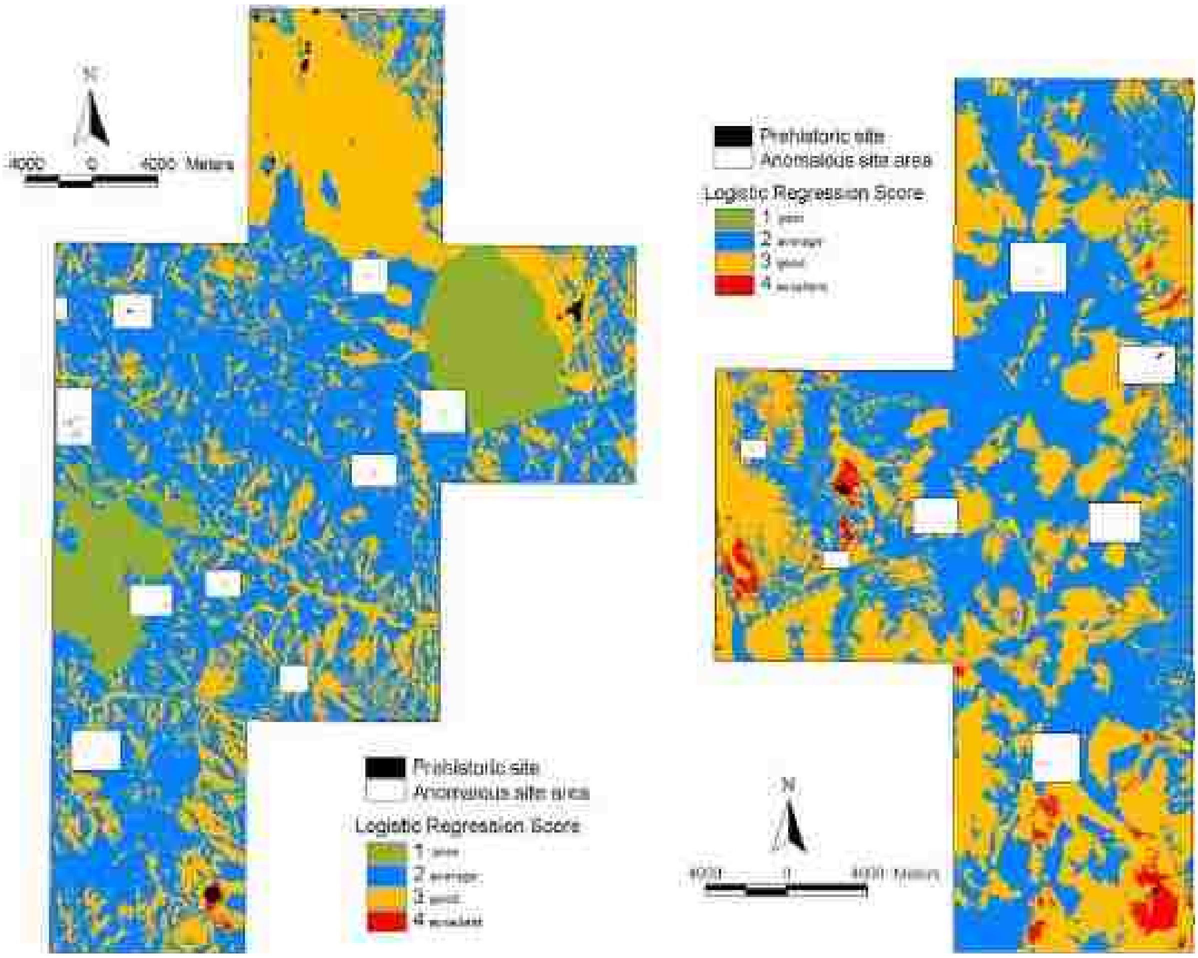


Figure 6**:** Logistic regression model Otero Mesa (Altschul et al. 2005 p. 138 Figures 7.12 & 7.13)

There are two commonly used regression algorithms for site predictive models: linear multiple regression (Equation 4) and logistic multiple regression (Equation 5) (Gillings and Wheatley 2002 pp. 154–156). The results produced by each algorithm are different and have advantages and disadvantages. While linear multiple regression produces estimates on the absolute value of a variable, logistic multiple regression produces results of the probability of a particular outcome (Warren 1990b, Gillings and Wheatley 2002, Legg and Taylor 2005, Cole et al. 2006).

$$\mathbf{y = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_kx_k}$$

Equation 4: Linear multiple regression; where a is a constant and $b_1 \ldots b_k$ are all the variables, and y is the dependent variable, in this case, archaeological resources.

$$\mathbf{L = a + b_1x_1 + b_2x_2 + b_3x_3 + \ldots + b_kx_k}$$

Equation 5: Logistic multiple regression.

There are multiple linear regression and logistic multiple formulae but they all take the same forms (Equation 4 and Equation 5). Unlike linear multiple regression, logistic multiple regression algorithms usually undergo a further step to calculate the value of L (Gillings and Wheatley 2002 p. 155) using the following formula:

$$\mathbf{P = (e^{L} / 1 + e^{L}) = 1 / 1 + e^{(1-L)}}$$

Equation 6: Cumulative logistic distribution function.

There is also binary logistic regression, which has been used on occasion for modelling (Crawford and Vaughn 2009). Binary logistic regression does not assume normality of the variables, but does require non-collinearity between independent variables (Crawford and Vaughn 2009 p. 548):

$$\mathbf{Probability\ (y) = 1 / (1 + (Exp - b_0 + b_1X_1 + b_2X_2 \ldots b_pX_p))}$$

Equation 7: Binary logistic regression function.

The maths for these formulae is not simple and cannot be given justice in this short section, but for more information on regression approaches in site predictive modelling see Warren (1990b) or Asch and Warren (2000) or Kvamme (1988b). For more on the statistics it is best to examine Altschul and Rose (1998).

Demand for Predictive Modelling (Gartner Hype Trigger and Peak Expectations)

The development of these methods was created as a result of the perceived demand for predictive models in the United States of America because of the passage of the National Historic Preservation Act (NHPA) in 1966, the National Environmental Policy Act (NEPA) in 1969 and the Archeological Resources Protection Act of 1979 (ARPA). These Acts mandated different actions on the part of federal government agencies with regard to the cultural resources that were, and still are, under their jurisdiction. One of the actions required was the identification of significant cultural resources that might be disturbed or destroyed (Berry 1984, Altschul et al. 2004, Dore and Wandsnider 2006, Naunapper 2006, Wescott 2006, Verhagen 2007d, Cushman and Sebastian 2008). Suddenly, hundreds of millions of acres of government land needed to be surveyed or required pre-development investigations. At the same time, the introduction of new state laws also required similar work on state land (Kaufmann 2006).

The principle was established that predictive models could locate sites and that CRM managers could demonstrate they had undertaken their due diligence to locate the archaeological resources without having to carry out a physical survey (Berry 1984, Brandon et al. 2000, Altschul et al. 2004, Dore and Wandsnider 2006). Not having to physically survey millions of acres of land would save significant amounts of time and money. This concept of 'predictive locating' can be seen in some of the first models created. For examples see Dincauze and Meyer (1976), Fuller, Gregonis et al. (1976), Grady, McCarth et al. (1978), Hackenberger (1978), Robertson and Robertson (1978), Barber and Roberts (1979), Holmer (1979), Burgess, Kvamme et al. (1980), DesJeans, Feiss et al. (1980), Senour (1980). This was part of the 'Trigger Event' on the Gartner Hype Cycle for predictive modelling, the other trigger was GIS.

The Rise of GIS

A rising demand for predictive models in archaeology also coincided with the appearance of a new tool, the computer. The late 1970s brought the first use of affordable computers; more accurately, 'affordable for archaeologists', to speed up the process of predictive model creation (Kvamme 1983a, Kvamme 1983b, Kvamme 2006). Chadwick was probably the first to use computers for predictive modelling in his study of the location of Early to Middle Helladic settlements in the Messenian region of Greece (Chadwick 1978). However, computers could only speed up the process of predictive modelling so much without the right software.

What had the most profound effect on modelling was the development of Geographic Information Systems (GIS) software programs for computers in the 1980s (Kohler and Kvamme 1988, Kvamme 1989, Kvamme 1999, Verhagen 2007d). GIS allows the user to manipulate and calculate large amounts spatial data (Savage 1990, Kvamme 1999, Beckman and Duncan 2000, Ebert 2000, Wescott 2000). Like word processing programs, there are many versions of GIS programs but they all do comparatively similar things, with some notable differences (Zubrow 1990, Hunter and Steiniger 2010). As stressed by many predictive modellers, GIS is a tool, nothing more (Allen, Green et al. 1990, Gaffney and van Leusen 1995, Harris and Lock 1995, Neustupny 1995, Berman, Camilli et al. 1996, Church and Ruggles 1996a, Church and Ruggles 1996b, Verhart and Wansleeben 1997, Beckman and Duncan 2000, Ebert 2000, Altschul et al. 2004, Verhagen 2007a, Wilcox 2009). For a general discussion on archaeology and GIS see Gillings and Wheatley (2002).

Before GIS, the tools involved in the creation of site predictive models were simple grid maps, pencils and either a calculator or scrap paper (Church and Ruggles 1996a, Kvamme 1999, Wescott 2000, Kvamme 2006a). Probabilities were calculated by hand and then sketched onto grid maps. This was, and still is, a time-consuming endeavour that severely limited the size and speed at which models could be created (Wescott 2000). By adopting GIS programs archaeologists could complete these steps in a matter of hours instead of days, saving effort, time and money (Brandon et al. 2000). If it had not been for GIS, the creation of models would have 'required the services of a small army to … create, test, and display the … predictive model' (Asch and Warren 2000). Further advantages of GIS were:

> 'The standards inherent in the design of the system assure that the data structure remains constant. In this way, analyses are cumulative, regardless of current research questions, researchers, or biases. In addition, the intuitive, graphical platform of GIS provides an interface through which ideas may easily be communicated with non-specialists.' (Altschul et al. 2004 p. 2)

It was the increasing ease of calculations and usability of GIS programs, combined with affordable computers to run the programs, that made the wide-scale use of site predictive modelling possible in the 1980s (Kohler and Kvamme 1988, Kvamme 1989, Kvamme 1999, Beckman and Duncan 2000, Brandon et al. 2000, Wescott 2000, Harris and Lock 2006). This, with the increased demand because of new laws in the United States, was the 'Trigger Event' on the Gartner Hype Cycle for predictive modelling.

Peak of Inflated Expectations

During the late 1970s and early 1980s there was an explosion in predictive modelling in the United States for CRM purposes (Kohler and Parker 1986, Thomas 1988). To support the growth of this work a US government department, the Bureau of Land Management, commissioned a comprehensive study of, and guide to, best practices for the creation of site predictive models. This was published as *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling* (Judge and Sebastian 1988). The late 1980s also saw that first development of predictive modelling in Canada (Dalla Bona 2000). This was the 'Peak of Inflated Expectations' for predictive modelling, the creation of a guide in anticipation of significantly more models being created and use expanding to other countries.

Clouds on the Horizon: Problems with Theory ('Trough of Disillusionment')

The focus on the environment as the most significant factor influencing site location in the 1960s, 1970s and 1980s arose because of the prevailing theoretical trends in North America at that time (Canning 2005). 'New Archaeologists' or 'Processual Archaeologists' prevailed as the dominant group and their theories and methodologies were firmly rooted in objective measurements and environmental determinism (Aldenderfer and Maschner 1996, Harris and Lock 2006, Trigger 2006, Ladefoged and McCoy 2009). Subsequent studies of archaeological theories have shown that the specific theory an archaeologist follows is predominantly determined by the intellectual climate when they began working, often approximately defined by the decade (Zeder 1997). This caused many archaeologists to perceive predictive modelling as an essentially Processual endeavour (Kamermans et al. 2009b).

While site predictive modelling was flourishing at the beginning of the 1980s, academic archaeologists were going through a re-examination of their methods and theories. From the late 1970s and through most of the 1990s Processual archaeology was the focus of criticism by Post-Processual archaeologists (Verhagen 2007d, Verhagen and Whitley 2011). Predictive modelling ties closely with the idea of ecological determinism, and statistical analysis came under criticism as part of a the wider critique of Processual archaeology put forth by the Post-Processual archaeologists (Harris and Lock 2006, Verhagen 2007d, Verhagen and Whitley 2011). Even GIS, a tool, came under scrutiny for being too ecologically deterministic (Gaffney, Stancic et al. 1996, Harris and Lock 2006). These criticisms did not just come from Post-Processual archaeologists. Some who would consider themselves Processualist were also critical of predictive modelling.

The term 'environmental determinism' was levelled at predictive models both as a descriptor and as a derogatory term. The ecological traits that had been used to recreate patterns of site locations throughout the 1960s to the 1980s were criticised for not representing the full range of potential factors determining human choices and decisions for site selection, such as religion or trade (Ridges 2006, Verhagen 2007d, Kay and Witcher 2009). This critique created a debate in site predictive modelling about what actually determines site location, a debate that is still ongoing today (Verhagen and Whitley 2011).

Furthermore, many Post-Processual archaeologists firmly refute the idea that human behaviour can be patterned (Shanks and Tilley 1987, Harris and Lock 2006, Deeben, Hallewas et al. 2007). Such an approach is hard to reconcile with predictive modelling, which is predicated on the idea that there are patterns to human behaviour:

> 'Predictive modelling is a technique to predict, at a minimum, the location of archaeological sites or materials in a region, based either on the observed pattern in a sample or on assumptions about human behaviour.' (Kohler and Parker 1986 p. 400)

Clouds on the Horizon: Problems with Accuracy, Precision and Legality ('Trough of Disillusionment')

At the same time that the theory behind the models was being criticised, concern was raised about models being used as a substitute for field surveys to inventory heritage resources (Berry 1984, Gaffney and van Leusen 1995, Harris and Lock 1995, Ebert 2000, Wheatley 2004, Dore and Wandsnider 2006, Kamermans 2007, Verhagen 2007a, Kamermans et al. 2009b). The central argument against using site predictive models to replace field surveys to meet the demands of the new laws was that they were simply not accurate enough. The models did not capture all of the previously known sites in a study area and as such their ability to capture unknown sites was questioned (Kvamme 1990, Brandon et al. 2000, Wheatley 2004, Kamermans 2007). Ebert observed that:

> ‘For some reason, and I have yet to determine just why this may be, the reported accuracies of inductive modelling seem to hover in the 60-70% range. Perhaps no one wants to report “success” rates only minimally higher than 50% ... sixty to seventy percent is not really bad but it is not very good either – certainly not good enough to justify spending a lot of money.’ (Ebert 2000 p. 133)

Others pointed out (Verhagen 2007c, Verhagen 2009) that the reason accuracy is so low has to do with balancing it against precision. In predictive modelling, accuracy is the measurement of how many locations are correctly labelled as containing archaeological remains (Altschul et al. 2004, Whitley 2004b, Kvamme 2006, Verhagen 2007c). However, if we label 100% of a project area as likely containing sites then the model, technically, would be 100% accurate as the model has identified all the archaeological resources correctly (Figure 7). This sort of labelling, full coverage of a subject area, is useless as it does not tell us where sites are going to be located, other than everywhere. To limit the area archaeologists have to look for archaeological remains, they have to use precision in their models. Precision in archaeological predictive modelling is defined as labelling the least amount of land as likely to contain archaeological resources (Altschul et al. 2004, Whitley 2004b, Verhagen 2007c).

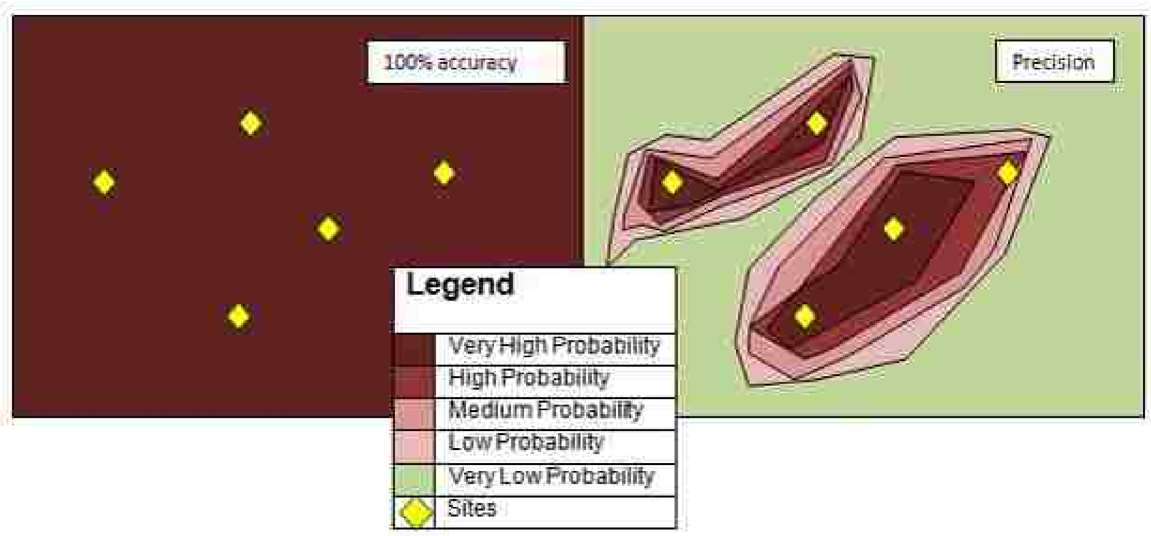


Figure 7**:** Accuracy vs. precision in site predictive modelling.

A modeller wants accuracy because missing a site creates the possibility for ‘gross error’. That is when a model indicates an area of land is devoid of archaeological resources when in fact it contains archaeological sites (Altschul 1988). Missed sites can result in two outcomes, neither of which are desirable for archaeologists. One is that archaeological resources are simply bulldozed without any further investigation, because the model indicated no archaeological resources were in a location (see Verhagen (2007c) for an example of this scenario almost happening). This is an incalculable loss of information, a scenario that archaeologists, heritage managers and the concerned public find unacceptable (Gaffney and van Leusen 1995, Harris and Lock 1995, Ebert 2000, Wheatley 2004, Dore and Wandsnider 2006, Kamermans 2007, Verhagen 2007a, Kamermans, van Leusen et al. 2009a). The other possible outcome is that an agency, such as government organisations, construction companies or developers, would have invested heavily in predictive modelling so they could avoid destroying archaeological resources when building but then still find archaeological resources in their development sites (Verhagen 2007c). They would then still have to undertake expensive surveys and rescue excavations. Not only would they have excavation costs, but delays in construction would ensue, which would be even costlier as they would also have to pay for idle equipment. This, in effect, makes it risky for developers to use predictive models. Precision reduces the scope required for archaeological prospection in advance of major developments, saving both time and money.

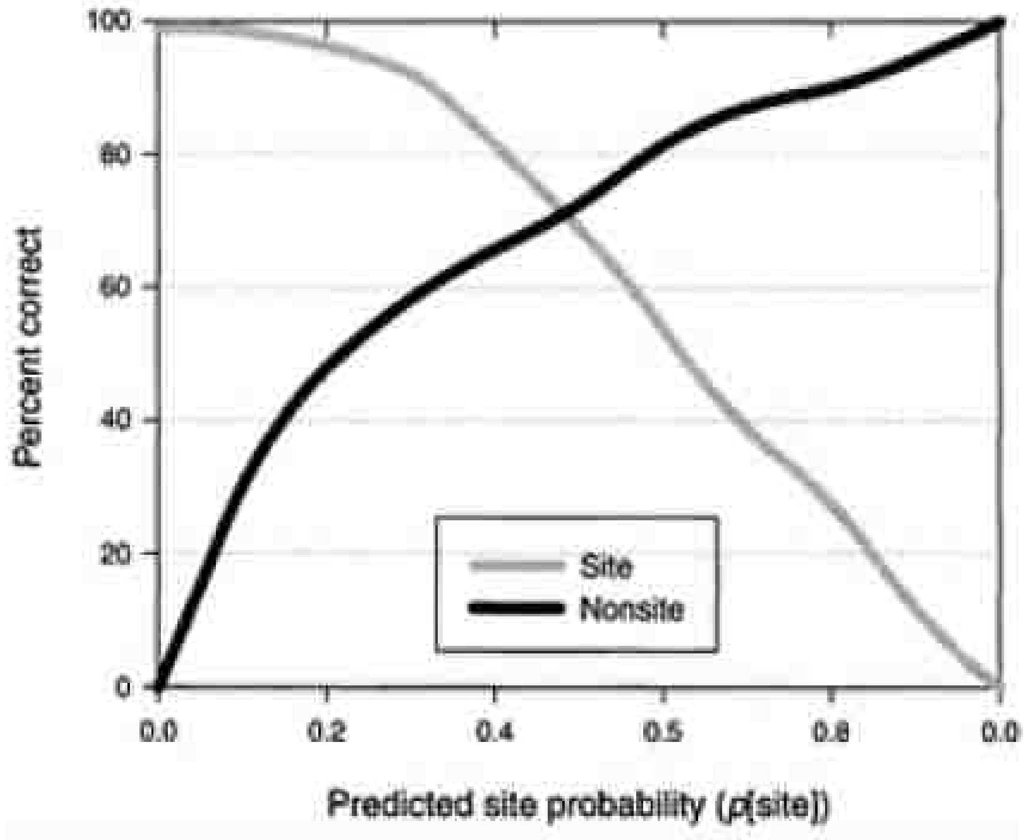


Figure 8**:** Accuracy of the Montgomery logistic regression model. The curves are percentages of correct predictions along a gradient of predicted site probability (Asch and Warren 2000 p. 25 Figure 2.7). Note there appears to be an error in the original figure – the x-axis scale should go to 1.0 and not 0.0.

As precision increases, accuracy decreases or vice versa (Figure 8). As models get more precise, labelling less of a study area 'high probability', accuracy goes down as fewer sites are captured. This is not perfectly linear but, as seen in Figure 8, rarely can significant gains be made in one area without some sort of diminishing returns in the other. Thus 30% or 40% of known sites were observed being missed in many models and it was concluded that this was an unacceptable error rate for Cultural Resource Managers (Custer, Eveleigh et al. 1986, Ebert 2000, Wheatley 2004).

In addition to these accuracy problems, CRM archaeologists in North America, which at this time was the primary location of predictive models (Brandt et al. 1992, Garcia Sanjuan and Wheatley 1999, Gillings and Wheatley 2002, Stancic and Veljanovski 2006, Kamermans 2007), realised that these models would not help them to fulfil all the legal requirements of the National Historic Preservation Act: Sections 106 and The National Environmental Policy Act. This legislation does not just require the identification of locations but also requires a determination on the significance of the cultural resource for our understanding of the past (Altschul et al. 2004, Altschul, Klein et al. 2005, Dore and Wandsnider 2006, Cushman and Sebastian 2008). In effect, this turns inventories from simple exercises in site location to locating sites and determining whether the sites can tell us something about the past. Early attempts to use predictive modelling to replace field surveys were thus ruled insufficient in the eyes of the law.

While several proposals have been put forward to create predictive models that could determine potential site significance, these proposals have never been implemented (Altschul et al. 2005, Cushman and Sebastian 2008). As such, predictive models are currently not judged to be able determine the significance of a site. Most US archaeologists now conclude that 'Although a model can identify areas of high potential for sites, it in no way substitutes for or eliminates the need for intensive archaeological survey.' (Kuiper and Wescott 2000 p. 74.)

Further Problems

To compound the issues faced by predictive models, a slew of technical problems appeared at about the same time. These details will be discussed in further detail in the next chapter but a sample of them from Kvamme's (2006 p. 6) list include:

- 'GIS data did not have sufficient resolution and poorly represent the real world';
- 'GIS data were inaccurate';
- 'known site distributions in extant government files and databases are biased because of (a) the haphazard way in which many were discovered and (b) variations in obtrusiveness, visibility, and preservation';
- 'many known sites are inaccurately located on maps and in databases'.

The Doldrums

During the 1980s and 1990s these technical issues, combined with a heavy critique of methodologies and purpose by the academic sector of archaeology, together with a reduction of interest from CRM archaeologists, resulted in a significant reduction in interest in predictive modelling. It has been said that at the beginning of the 1990s predictive modelling 'entered a period of doldrums' (Brandon et al. 2000, Brandon and Wescott 2000, Kvamme 2006) that some say persists to the present, or at least till recently (Altschul et al. 2004, Deeben et al. 2007). This was predictive modelling's 'Trough of Disillusionment' on the Gartner Hype Cycle.

A Silver Lining

Yet this period was not without its positive developments. Predictive modelling became less North America-centric and the application of predictive models began to spread around the world, with European archaeologists beginning to take an interest in predictive modelling (Brandt et al. 1992, Lock and Stançic 1995, Kvamme 1999, Verhagen 2007d). Archaeologists in the Netherlands were especially keen adopters (Kamermans and van Leusen 2005, Kamermans et al. 2009a). Since then, Dutch archaeologists have been the driving force behind some of the most recent innovations in predictive modelling (Verhagen 2007d), as will be discussed later in this chapter. At the same time, some CRM managers in the United States still saw use for them, especially those involved in CRM for the military (Altschul et al. 2004).

New Demand

An important development in site predictive modelling during the 1990s was the creation of the Minnesota Department of Transportation (MDOT) State-wide Site Predictive Model. The Minnesota model was created to help with the planning of new developments, in what is called the 'flag and avoid' method of use (MDOT 2009). 'Flag and avoid' is when locations that are likely to contain archaeological resources are highlighted by a site predictive model and then development projects are moved around those locations (Figure 9) (Warren 1990a, Altschul et al. 2004, Canning 2005, Cole et al. 2006, Naunapper 2006, Stancic and Veljanovski 2006, Wescott 2006, Cushman and Sebastian 2008, Bailey, Grossardt et al. 2009, MDOT 2009, Kamermans et al. 2009a). Surveys and mitigations are still carried out but the surveys take less time, because they are able to avoid the majority of sites (MDOT 2009). If there are fewer sites affected, that also means less time and money is spent on rescue excavations or further testing.

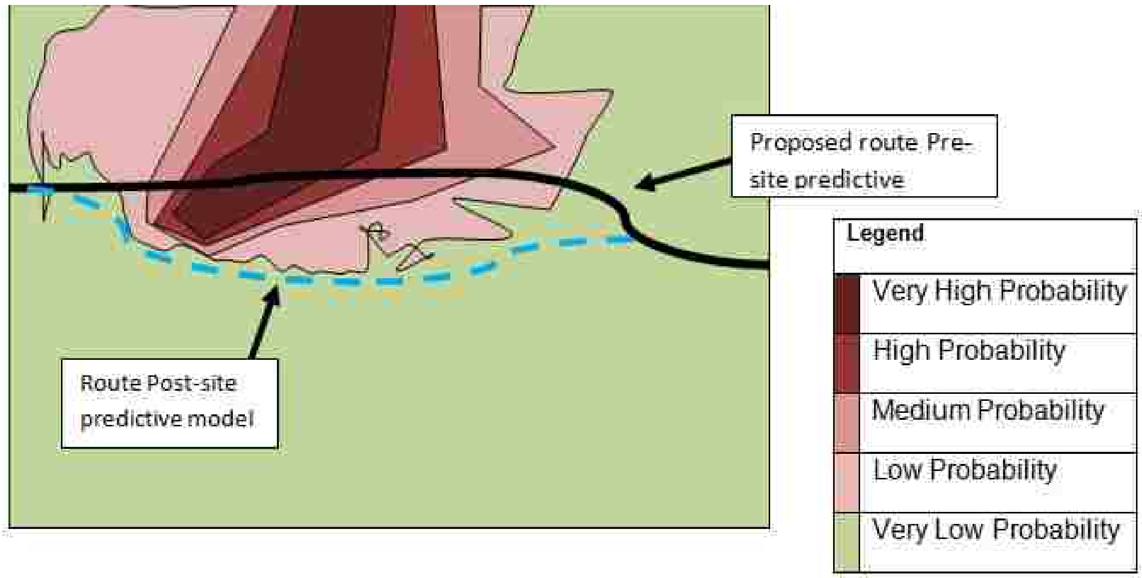


Figure 9: Hypothetical example of flag and avoid use in predictive modelling for a new road.

Using 'flag and avoid' had dramatic results in Minnesota. The MDOT estimates that it was able to handle an increase of 35% more projects without increasing its archaeological staff. At the same time, the application of this approach lowered the need for mitigations by 80% and increased their turnaround on projects by 30%. This resulted in savings of around $12 million over the first four-year period of the model's use and it was estimated that it paid for itself within the first year and a half of use (MDOT 2009). These numbers do not take into account other organisations or individuals who also used the model to support their planning needs.

This project was followed by another in North Carolina (Cole et al. 2006, Cole et al. 2006). Furthermore, using 'flag and avoid' with predictive modelling has also become a more widely accepted tool among a variety of construction organisations (Cambridge Systematics 2009). It has led to an increased interest in predictive modelling and a revival in its use.

The Predictive Decision Matrix

Flag and avoid was not the only new use of predictive modelling for CRM purposes. In Canada, managers began to expand the use of predictive models by combining them with decision matrixes. These were tools to tell the users what types of archaeological resources they will likely encounter, what sort of developments will disturb them and the appropriate action to take (Dalla Bona and Larcombe 1996, Dalla Bona 2000, Clement, Kloot et al. 2001). This simplified a complex process for non-archaeologists into a graphical presentation. It was found to provide a useful planning tool for non-archaeologist project managers, who were usually unsure about what to do (Dalla Bona 2000). These predictive model and decision matrix combinations have been proposed for use in the USA (Altschul et al. 2004, Altschul et al. 2005), for determining the significance of a site

(Altschul et al. 2004, Altschul et al. 2005, Cushman and Sebastian 2008, Goudswaard, Isarin et al. 2009) or rarity of the site type (Verhagen 2007c).

Predictive Erosion

At the same time, predictive models were also combined with erosion models and remote sensing techniques to produce maps of unsurveyed areas where potential sites were threatened by erosion (Figure 10) (Ebert and Singer 2004, Dore and Wandsnider 2006). As many CRM managers do not have the resources to examine all of the land under their care, this application of predictive modelling highlights those locations in most critical need of survey. There have been suggestions of expanding this technique to counteract additional problems such as looting, recreational damage, military training, etc. (Dore and Wandsnider 2006.)

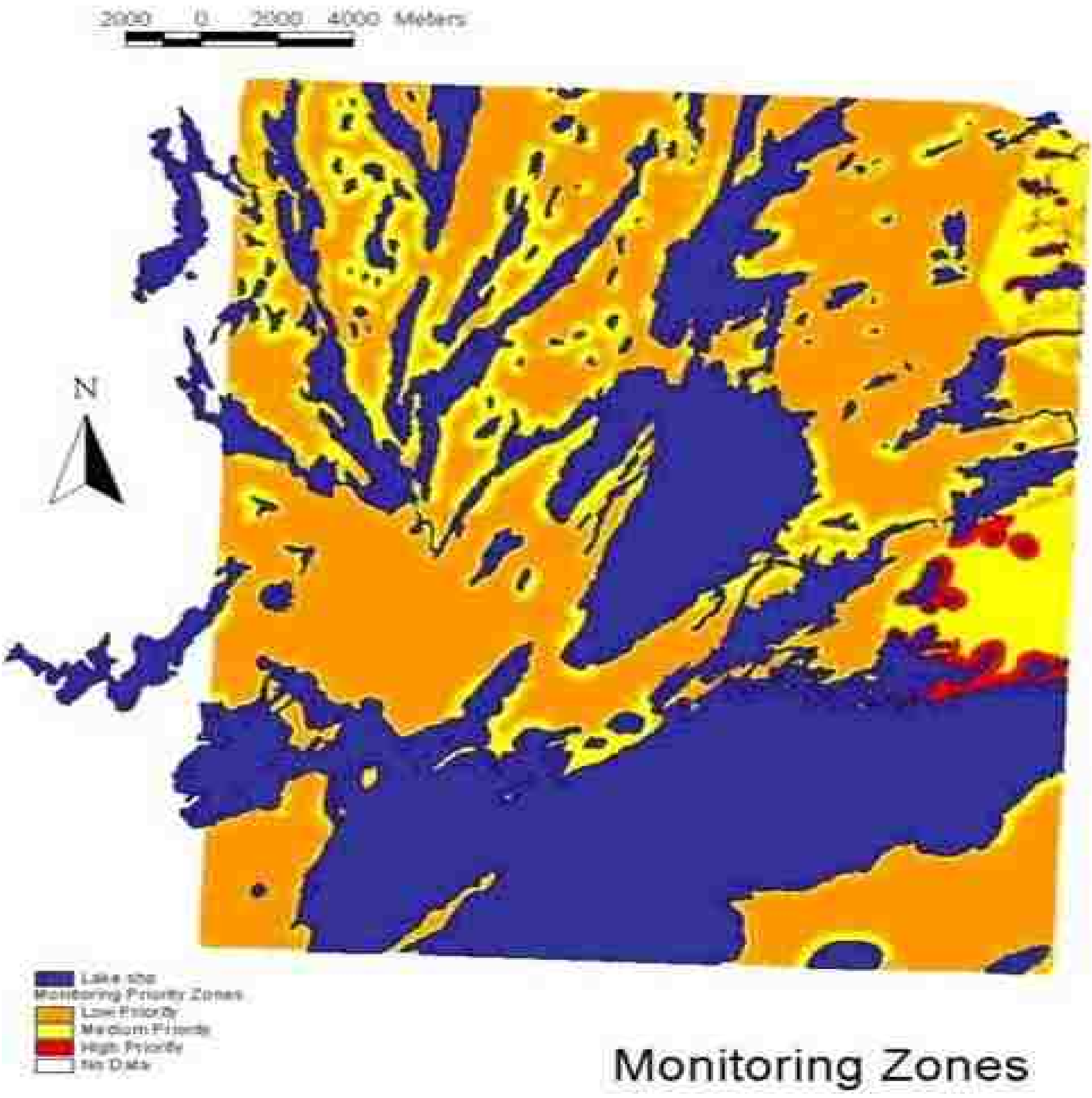


Figure 10: Example of erosion of site predictive model (Ebert and Singer 2004 Figure 4).

Light at the End of the Tunnel ('Slope of Enlightenment')

Even during the 'doldrums' a small group of archaeologists continued to publish on the topic, but with the renewed demand for predictive models because of new uses like 'flag and avoid', the turn of the millennium saw a significant increase in research. There were two major books from conferences, *Practice Applications of GIS for Archaeologists: A Predictive Modeling Kit* (Brandon and Wescott 2000) and *GIS and Archaeological Site Location Modelling* (Mehrer and Wescott 2006). In the Netherlands archaeologists ran a series of projects and created several publications on the subject of improving site predictive models (Kamermans and van Leusen 2005, Verhagen 2007a, Kamermans et al. 2009a). This has led to the development and testing of new methods such as Bayesian and Dempster-Shafer statistics (see Chapter 2 for discussion on these methods). The first of these publication was 'Predictive Modelling for Archaeological Heritage Management: A Research Agenda' (Kamermans and van Leusen, eds, 2005), which addressed the issues of methodology and model improvement. The second was *Case Studies in Archaeological Predictive Modelling* (Verhagen 2007a), which tackled some of the cultural resource issues of predictive models. These were followed by *Archaeological Prediction and Risk Management* (Kamermans et al. 2009a), which combined elements of both with additional projects to create an integrated overview of predictive modelling.

Furthermore, a series of individual journal articles and reports (Whitley 2006, Podobnikar and Šprajc 2007, Goel, Jackson et al. 2009, Verhagen and Whitley 2011) have also been published in the last few years that have added greatly to the development of predictive modelling. For example, Podobnikar and Šprajc's (2007) predictive model looked at the cultural concept of visualisation as a predictor; a practice proposed a few years ago (Harris and Lock 2006), while others have tested out newer methodologies, like Fuzzy Logic (Hatzinikolaou 2006, Bailey et al. 2009), in hopes of improving predictive modelling. Several recent models have looked at cultural factors as influencing features in predictive modelling, addressing some concerns of environmental determinism (Maschner 1996, Harris and Lock 2006, Ridges 2006, Stancic and Veljanovski 2006, Podobnikar and Šprajc 2007, Crawford and Vaughn 2009).

Where is Predictive Modelling in Its Development Now?

The late 1970s and early 1980s witnessed a strong need for predictive modelling to meet the needs of CRM archaeologists in the USA. This also coincided with the application of GIS in predictive modelling, which greatly aided in the speedy construction of predictive models. These combined became the 'trigger' point for predictive modelling, with strong demand combined with easier access. All of which led to a 'Peak of Inflated Expectations', when the BLM created a 'how to' guide for modelling in anticipation of great demand, *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeological Predictive Modelling* (Judge and Sebastian 1988).

The late 1980s and the 1990s, the period typified as the 'doldrums' (Brandon et al. 2000, Brandon and Wescott 2000, Kvamme 2006), saw the inflated hype surrounding predictive modelling dissipate. In the case of the United States, it was realised that predictive modelling cannot replace physical survey and still meet the requirements of the law. Furthermore, predictive modelling in general ran into problems with the rise of Post-Processualism as one of the dominant theoretical perspectives. This development coincided with the realisation that there are many technical problems with predictive models.

Finally, aspects of the 'Slope of Enlightenment' were seen, with archaeologists proposing new methodologies to help solve some of the problems facing predictive modelling. Models began to take into account non-environmental factors. Most importantly, with the development of the MDOT Minnesota predictive model, a new use was found for predictive modelling that did not attempt to substitute for physical surveys. All this began in the mid-1990s with the MDOT model and was continuing, arguably, to when this project began. although as some have mentioned, predictive models were, and possibly still are, in a period of depressed use (Altschul et al. 2004, Deeben et al. 2007).

Market for Predictive Models

Even with developments to address issues in predictive modelling there was still great potential for predictive modelling to expand beyond its current niche use. In the area of CRM, where it sees its strongest use, predictive models are still not of interest to the majority of cultural resource management organisations (Kaufmann 2006). Only one country has full site predictive model coverage, The Netherlands, and only one state in the United States, Minnesota. This represents a little less than .17% of the world's land, leaving well over 99% of the land without site predictive models. Even when including other smaller models, the percentage of world coverage was unlikely to break one or two percentage points.

This is the historical context in which this project began— uncertainty about whether predictive modelling had reached its peak performance and usefulness for archaeology, finding a small niche use within archaeology or if it could be improved still further and be expanded in use. There certainly was the potential for growth given the low usage, which meant there was still scope to add something significant to the development of predictive modelling. As the next chapter will review there were, and still are, many problems that can be addressed and significant areas in which this project can contribute to the development of predictive modelling.

Chapter 2: Current Problems Facing Archaeological Predictive Modelling

The previous chapter presented a history of predictive modelling that matches the Gartner Hype Cycle and indicated that predictive modelling was on the 'Slope of Enlightenment', as it was in 2010 when this research project began. Predictive models still had problems that needed to be solved in order to improve the process and expand its use beyond a small niche of archaeologists.

What follows is the presentation of the major problem that this project aims to solve: poor model performance. As discussed in the previous chapter, there were many issues raised with predictive modelling in the past and the justification for choosing poor performance as the topic is discussed. This is then followed by exploration of the potential causes of poor performance. In doing so this chapter covers the first planned-for project activity: 1. research causes of poor model performance and find cause to address. The chapter ends with a discussion of the causes of poor performance that this project would attempt to solve.

Overarching Problem

> 'the reported accuracies of inductive modelling seem to hover in the 60–70% range... sixty to seventy per cent is not really bad but it is not very good either – certainly not good enough to justify spending a lot of money ...' (Ebert 2000 p. 142).

There are problems with predictive modelling but Ebert's criticism cuts to the heart of the reason 99% of the world is without predictive modelling. As discussed in the review of predictive modelling, if a model does not capture enough sites it can lead to 'gross error' (Figure 11). A solution would then be to increase accuracy to eliminate this 'gross error'. The problem with this solution is that the 'flag and avoid' method requires a high level of precision in models. Most developers have their own constraints that require them to use only certain parts of the landscape, e.g. you build on land you own, wind turbines need to be in locations that have wind. This limits the amount of land that is suitable for development *and* where cultural resources are absent, thus the need for high precision. Fundamentally, this means that models need to have both high accuracy and precision, but that rarely happens.

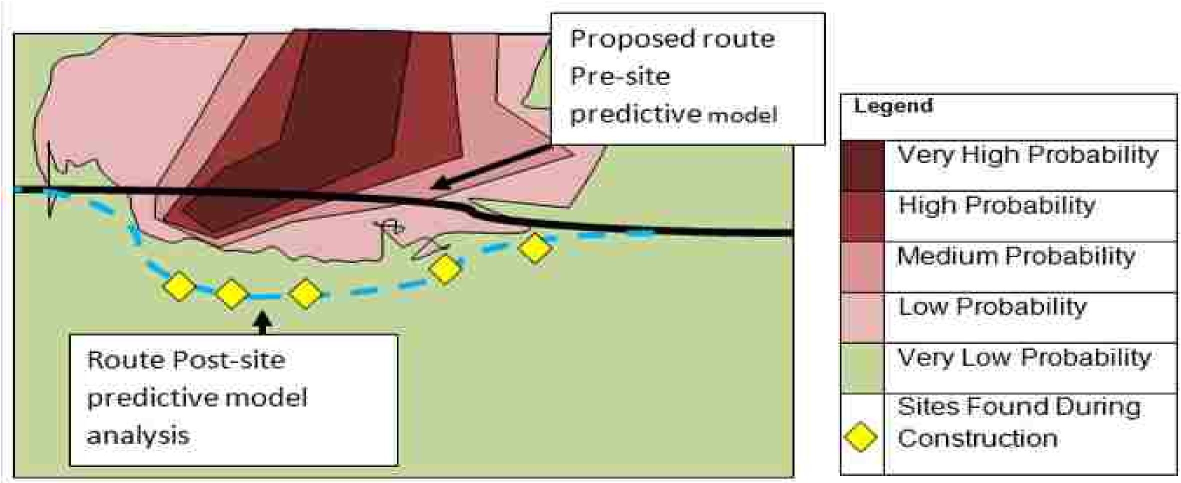


Figure 11: Example of gross error sites found during construction. Predictive model indicates that an area is devoid of sites and the project is moved to those areas. Sites are found in the very low probability areas.

One estimate is that for models to be truly useful they need to capture 70% of sites in the high probability zone while only covering 10% of the land (Gibson, 2005). The Minnesota Model (MnModel) discussed in Chapter 1 captures 85.5 % of the sites while covering 23% of the land and that has a proven record of use and utility (MDOT 2009). To measure both accuracy and precision Kvamme has developed a simple statistic known as Kvamme's Gain Statistic. This takes the total area covered by the model (precision) and the total number of sites captured, either with old data or new, to create a simple number that represents both accuracy and precision. Kvamme's Gain Statistic is the most commonly used method of measuring quality of models (Verhagen 2007c, Ducke, Millard et al. 2009).

Gain = 1 – (% of total area covered by model / % of total sites within model area)

Equation 8**:** Kvamme's Gain Statistic, the closer to one the better the model.

Using Kvamme's gain equation it was possible to compare these numbers to see what goal models should be obtaining in their predictive model as far as accuracy and precision is concerned.

The results were that the MnModel had a gain value of .73 and Gibson's proposal was .86. This indicates that a gain value of at least .7 to .85, or above, is needed to have significant precision and accuracy to reduce the possibility of gross error and to be able work in most CRM contexts. Otherwise models end up in drawers or buried on hard drives, and are never used again (Altschul et al. 2004).

An examination of previously reported gain values or calculated gain values revealed that many models do not reach this level of performance (Table 1 below). These models only represent those projects with published results; the actual number of predictive models created in the world is much higher. It was seen that across a plethora of locations, using a variety of methods, many models struggle to reach the desired level of performance.

This is a clear, unresolved problem for predictive modelling. Why does this matter? Because, as Ebert said, performance like that, '… is not really bad but it is not very good either – certainly not good enough to justify spending a lot of money on doing this sort of predictive modelling' (Ebert 2000 p. 142). Predictive models are not always cheap to produce; the MN model cost $6 million and the North Carolina model cost the same but did not cover the whole state (Cole et al. 2006, MDOT 2009). The original models for the subject area of this project were part of a $1 million project. These costs occur for different reasons, much of which has to do with the cost of obtaining data, and can vary from project to project. Still, this is a significant amount of funding to invest in a project that may not return the desired performance. It is not surprising that heritage managers are not more interested in predictive modelling. As such the aim of this research project was created – **improve the performance of site predictive modelling for CRM archaeology uses**.

Accuracy	Precision	Gain Statistic	Source	Method
82%	24.6%	.7	(Clarke et al. 2009)	Bayesian
59%			(Harris and Lock 2006)	Boolean
84%	22%	.74	(Hasenstab and Resnick 1990)	Boolean/correlative
1996: 36.1%; 1997: 67.6%; 1998: 70.3%	1996: 13.9%; 1997: 16%; 1998: 14.8%	.79	(Gazenbeek and Verhagen 2007)	Correlative
89.8%; 79.4%	61.3%; 46.7%	.32; .41	(Verhagen 2007c)	Correlative
89.19%; 83.78%	12.48%; 16.25%	.86; .58	(Dalla Bona 2000)	Deductive/weighted
		.45-.73	(Ducke et al. 2009)	Dempster-Shafer
		.9242; .9316; .9412; 9297; .9218; .9333	(Bailey et al. 2009)	Fuzzy Logic
42 of 46 high (91%); 4 of 46 medium (9%)	(19.2%) high; (29%) medium	Gain .52-.79	(Kuiper and Wescott 2000)	Regression
77%	30%	.61	(Asch and Warren 2000)	Regression
79%			(Beckman and Duncan 2000)	Regression
95%; 85%	8% high; 11% medium	.92 and .87	(Custer et al. 1986)	Regression
87.6% 1st; 50.7% 2nd			(Legg and Taylor 2005)	Regression
80%		24.6% and 30.6%	(Wilcox 2009)	Regression
65.5% high; 30.8% medium	26% high; 47% medium	.63	(Cole et al. 2006) St. John's County	Regression
84% high	29% high	.65	(Cole et al. 2006) Duval County	Regression
57.89%	42.82%	.26	(Crawford and Vaughn 2009)	Regression
90% 1st; 90% 2nd	22.4% 1st; 39% 2nd	.75; .57	(Devitt, Hill et al. 2009)	Regression
86.97%			(Maschner 1996)	Regression
68%			(Warren 1990a)	Regression
75%			(Carmichael 1990)	Regression
71% 1st; 75% 2nd			(Ridges 2006)	Regression
70.8%	20%	.72	(Cole et al. 2006)	Regression
86% 1st; 84.5% 2nd	18% 1st; 23% 2nd	.79; .73	(Graves 2010)	Regression
24% high; 57% middle	6%; 36%	.75; .37	(Brandt et al. 1992)	Weighted
20% highest; 53% with medium	6%; 26%	.7; .49	(Kay and Witcher 2009)	Weighted
38.75% 1st; 67.38% 2nd	25% 1st; 44.94% 2nd	.35; .33	(Altschul et al. 2004) FortStewart	Weighted
60%	40%	.333	(Altschul et al. 2004) Eglin AFB	Weighted
1: 71.43% & 53.33%; 2: 71.42% & 68.33%	46.18% 1; 51.3% 2	.75-.81	(Mink, Pollack et al. 2006)	Weighted
60%	46%	.23	(Brewster and Reddy 1999)	Weighted

Table 1: Selection of performance results from site predictive models. Kvamme's Gain Statistic column (Black text) when listed (Red text) estimated based on accuracy and precision values.

Performance Vs Other Problems

Poor model performance is not the only reason that predictive modelling is not more widely used. Chapter 1 reviewed the issues with the lack of explanatory power of predictive modelling, and its reliance on Processual theories and methods made it unappealing to non-Processual archaeologists. However, the initial goal of this project was to explore predictive modelling for CRM archaeology purposes. So initially the other concerns with predictive modelling were set aside and performance was focused on, but as will be shown in this chapter the issues of performance and lack of theory are not mutually exclusive.

Specific Causes of Performance Issues

A failure to obtain high accuracy and precision results is the problem, but what are the causes? When this project was started it was suspected that there might be several causes that contributed to this issue. There have been a few reviews of the detailed problems surrounding predictive modelling (Ebert 2000, Wheatley 2004, Kvamme 2006, Kamermans 2007). Since those lists were created work has been conducted to solve some or all of the issues listed. This section of the chapter presents the review of the listed issues conducted at the beginning of this project to determine which problems were still relevant to site predictive modelling, and whose solving might have contributed to solving model performance issues.

Kvamme's list (2006 p. 6) (Table 2 below) will serve as a baseline to start with in the review. Several additional problems not addressed in Kvamme's list will be examined as well. This should provide a clear idea of what are, and are not, still problems for predictive modelling.

	Archaeological
1	'Many archaeological sites are buried, and we cannot model them because we do not and cannot know their distribution.'
2	'Known site distributions in existing government files and databases are biased because of (a) the haphazard way in which many were discovered and (b) variations in obtrusiveness, visibility, and preservation.'
3	'Many known sites are inaccurately located on maps and in databases.'
4	'One cannot model archaeological site distributions because "site" is a meaningless concept; human behaviour did not occur in discrete bounded areas but formed a continuum over the landscape.'
5	'Functional, temporal, or cultural site types cannot be readily determined for most sites in an archaeological database, yet profound locational differences must exist between the types.'
6	'We must be able to model and understand the archaeological formation process, both natural and cultural, before we can model where sites might be found.'
	Environmental
7	'We do not know the locations of resources important in past times, such as water sources, springs, edible-species distributions, lithic raw material sources, and the like.'
8	'Past environments were very different from present ones, so we cannot model the past based on the present.'
9	'Models based on landscape variables are meaningless.'
	Technical

10	'Models based on site presence-absence criteria are mis-specified because one cannot assume site absence.'
11	'Blue-line features on topographic maps are frequently arbitrary and unreliable indicators of water.'
12	'Modern soil types are meaningless because they are changed from the past and, in any case, are frequently irrelevant to past farming practices.'
13	'GIS data have insufficient resolution and poorly represent the real world.'
14	'GIS data are inaccurate.'
15	'Linear distances computable in GIS are meaningless.'
16	'Models based on statistics cannot meet random-sampling assumptions because most extant data were not obtained by random sampling.'
17	'Models derived from random cluster sampling are mis-specified because they do not adjust for underestimated variances.'
18	'Grouping sites of many types into a single, site-present class creates too much variability to model.'
	Behavioural
19	'Environmental variables shown to be important to site locations may only be proxies for variables that were actually important.'
20	'Human behaviour is too idiosyncratic to be modelled; one cannot model the unique.'
21	'One must understand and model complete behavioural systems before archaeological models can be built.'
22	'Site location is more a function of unknown (and frequently unknowable) social environments representing dimensions that we cannot map.'
23	'The most interesting sites are the (idiosyncratic) ones that do not fit the pattern.

Table 2**:** Table created to include Kvamme's list of problems with predictive modelling (Kvamme 2006 p. 6).

Kvamme's list, while comprehensive, catalogues many issues as separate problems when they could have been combined as a single issue. Much of this can be viewed as semantics, whether or not a person is a 'splitter' or a 'lumper' when it comes to lists. To minimise repetition for this dissertation I take a lumping approach and try to address several similar problems at once, where applicable. which means not all of the problems are addressed in the order of the list.

To start:

3. 'Many known sites are inaccurately located on maps and in databases.' (Kvamme 2006 p. 6)

In many predictive models, site location data comes from existing databases (Kamermans 2007 p. 74) and is usually the result of the compilation of many different projects and surveys (Carmichael 1990). This results in a lack of standardisation, clerical mistakes, and incomplete records that lead to errors in the data (Marozas and Zack 1990, Garcia Sanjuan and Wheatley 1999, Beckman and Duncan 2000, Kuiper and Wescott 2000, Altschul et al. 2004, Kvamme 2006).

It is thought that inaccurate data will lead to poor model outcomes and work has determined there are problems with site records used for predictive models. A project in Nebraska, USA, found that sites were anywhere from a few hundred metres to kilometres away from their correct locations in the database. To account for that they had to add buffer rings using GIS to site locations of between 353m to 1km to produce a 90% confidence in site locations (Dore and

Wandsnider 1999, Dore and Wandsnider 2006). An examination of site locations held in Bureau of Land Management regional offices in the United States found that close to 10% of sites are located more than 100m, and some close to 400m, from their supposed locations (Kvamme 1988a). However, neither of these projects investigated whether this would cause problems with predictive models.

My Masters project did spot checks of sites. Errors were found to exist in site locations in the subject area but these errors had no effect on the results for three different methods of predictive model creation: Boolean, weighted and regression algorithms (Rocks-Macqueen 2010). This was because the number of misplaced sites was fairly small, not enough to change correlations or cause false results, and these incorrect locations actually had the same characteristics as the correct location (Rocks-Macqueen 2010). For example, one site was incorrectly located in the database by 100m but was still next to a drainage channel on flat land (Figure 12), an area that would have been considered high probability.

This leaves this problem as one that had potential to be investigated. There were known issues with data accuracy in predictive modelling. My Masters work found that it was not an issue, but that was a single case study. There was the possibility that this could be a major issue in predictive modelling and further testing would be needed to confirm that and find some possible solutions.

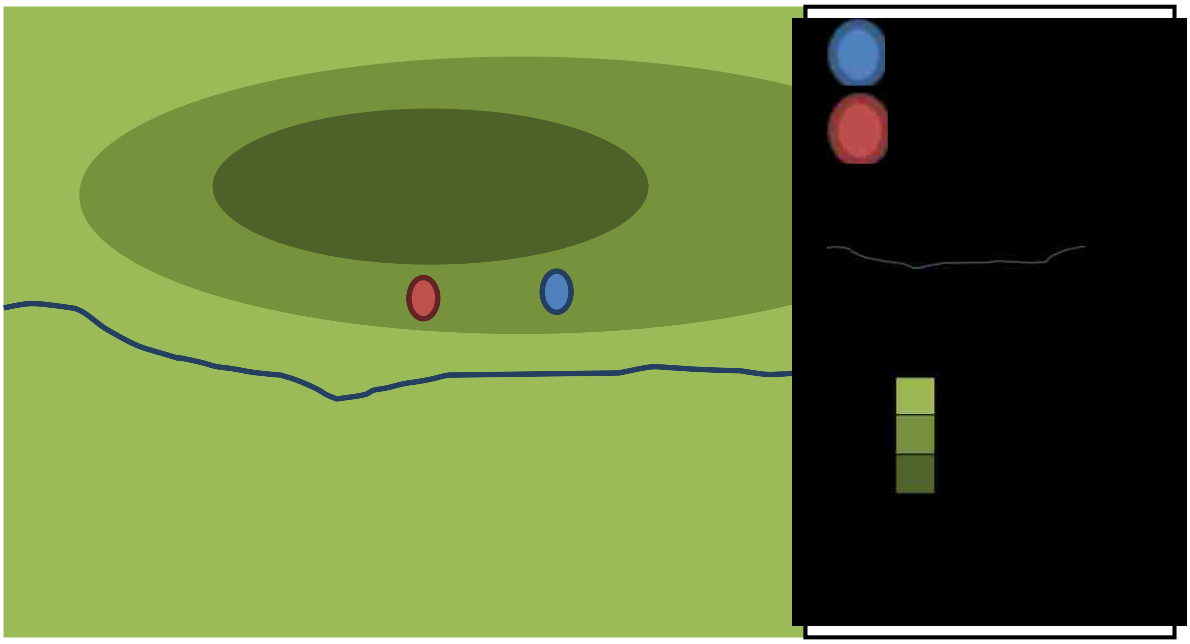


Figure 12: Hypothetical example of mis-located site not changing the correlation with environmental factors.

> 4. ‘One cannot model archaeological site distributions because “site” is a meaningless concept; human behaviour did not occur in discrete bounded areas but formed a continuum over the landscape.’ (Kvamme 2006 p. 6).

It has been recognised that sites are arbitrary boundaries drawn around locations.(Gaffney and van Leusen 1995, Berman et al. 1996, Ebert 2000, Dore and Wandsnider 2006, Harris and Lock 2006, Whitley 2006). This was required because CRM laws and regulations force some sort of limit on what is considered an archaeological resource (van Leusen 1996, Brewster and Reddy 1999, Dalla Bona 2000, Verhagen 2007b). This limit can range from any two different artefacts in close proximity (Verhagen 2007b) or a historically significant location to a group or larger society (Altschul et al. 2004, Altschul et al. 2005, Dore and Wandsnider 2006, Cushman and Sebastian 2008). For CRM purposes a site is a very real concept that must be modelled, though defined differently across countries.

If there is concern with modelling a ‘continuum over the landscape’ of cultural heritage, then one could use ‘siteless’ data. In practice, while not widespread, several ‘siteless’ surveys have been undertaken (Dunnell and Dancey 1983, Ebert, Larralde et al. 1984, Ebert and Kohler 1988, Banning 2002, Caraher, Nakassis et al. 2006). However, while ‘siteless’ surveys have been created there has been almost no interest in creating a ‘siteless’ predictive model. Still, the option is available should a model maker feel that it is in their best interest to do so.

The concept of the site and its relationship with archaeology is complex and affects whole swathes of archaeological research, theory and work, not just predictive modelling. But, given the goal of focusing on CRM which requires ‘sites’ this aspect was deemed too broad for a project that was to focus on predictive modelling and not particularly relevant to the project at hand.

> 5. ‘Functional, temporal, or cultural site types cannot be readily determined for most sites in an archaeological database, yet profound locational differences must exist between the types.’ (Kvamme 2006 p. 6)

Which is roughly the same concern as:

> 18. ‘Grouping sites of many types into a single, site-present class creates too much variability to model.’ (Kvamme 2006 p. 6)

Due to sample sizes and gaps in the information known about sites in databases, some modellers lump together all site types regardless of cultural, utility, or temporal makeup in a single class (Brandon et al. 2000, Kuiper and Wescott 2000, Dore and Wandsnider 2006, Kamermans 2007). It is said that models that employ such data are of low statistical value when trying to infer differential patterns of human behaviour (Hasenstab and Resnick 1990, Dore and Wandsnider 2006). Several models have been quite successful in separating out different site types and creating effective models from those (Hasenstab 1996, Maschner 1996, Kuiper and Wescott 2000, Stancic and Veljanovski 2000, Gazenbeek and Verhagen 2006, Kamermans 2006, Vermeulen 2006, Whitley 2006, Berger and Verhagen 2007, Gili, Mico et al. 2007, Ducke et al. 2009).

The variety of methods used in different models means that lumping together site types only effects some models and not others. Furthermore, the assumption that it ruined the model has yet to be tested, meaning this area did have the need for more research.

> 6. 'We must be able to model and understand the archaeological formation process, both natural and cultural, before we can model where sites might be found.' (Kvamme 2006 p. 6)

The cultural aspect of that statement encompasses this issue as well:

> 22. 'One must understand and model complete behavioural systems before archaeological models can be built.' (Kvamme 2006 p. 6)

The reason Kvamme lists this as a problem is not clear. Looking at the published literature, no reference is found that makes the statement, 'We must be able to model and understand the archaeological formation process, both natural and cultural, before we can model where sites might be found', or any related statements. It is possible that Kvamme is referencing unpublished work, but he does not cite where these problems come from so it is impossible to follow up on this point. No other predictive modelling references this as being an issue. Given the lack of discussion around this topic in predictive modelling it was not considered a strong candidate for further research for this thesis.

All of the following issues address the problem of past environments:

> 7. 'We do not know the locations of resources important in the past times, such as water sources, springs, edible-species distributions, lithic raw material sources, and the like.'
>
> 8. 'Past environments were very different from present ones, so we cannot model the past based on the present.'
>
> 12. 'Modern soil types are meaningless because they are changed from the past and, in any case, are frequently irrelevant to past farming practices.'
>
> (Kvamme 2006 p. 6)

Archaeologists have questioned the effectiveness of using present environmental data as a basis for past environments in predictive modelling (Canning 2005, Verhagen 2007e, Beckman and Duncan 2000, Brandon et al. 2000, Ebert 2000). At the risk of repeating the same statement over and over again, no research has been provided to back up these statements. For example, several modellers have discussed the changes in soil type (Allen 2000, Brandon et al. 2000, Ebert 2000) but as yet this is untested in regards to negative consequences for predictive modelling. Whitley's caloric based predictive models (Goel et al. 2009, Verhagen and Whitley 2011), which are created using soil patterns to estimate the caloric capacity of a landscape for his predictive models, appear to suffer no ill effects from using modern soil data. This leaves these problems as unsubstantiated but possible candidates for investigation.

> 11. 'Blue-line features on topographic maps are frequently arbitrary and unreliable indicators of water.' (Kvamme 2006 p. 6)

This problem is identified by several archaeologists (Brandon et al. 2000, Ebert 2000). At the beginning of this project it was assumed that this question was not an issue, due to past models working using modern hydrology data and some models recreating waterways. An example of the recreated waterways was the PUMP III predictive model, discussed in chapter 5, which used GIS tools to reconstruct waterways. However, as this project would discover after first dismissing this issue, it was actually a very significant problem, at least for this project's case study area.

> 20. 'Environmental variables shown to be important to site locations may only be proxies for variables that were actually important.' (Kvamme 2006 p. 6)

In essence, correlation does not equal causation in predictive modelling (Kvamme 1985, Maschner 1996, Brandon et al. 2000, Dalla Bona 2000, Harris and Lock 2006). This is a good point and one that can be applied to all of archaeology and other disciplines. It is too general a problem to be a site predictive modelling problem only, and was considered better placed as a critique of archaeology in general.

> 9. 'Models based on landscape variables are meaningless.' (Kvamme 2006 p. 6)

This issue is discussed in the section in Chapter 2 about environmental determinism and the Post-Processual critique of models based on environmental variables, concepts that have been hotly debated in site predictive modelling literature (Ebert and Kohler 1988, Wheatley 1993, Gaffney and van Leusen 1995, Wheatley 1995, Gaffney et al. 1996, van Leusen 1996, Wheatley 1996, Asch and Warren 2000, Beckman and Duncan 2000, Brandon et al. 2000, Dalla Bona 2000, Kuiper and Wescott 2000, Wheatley 2004, Legg and Taylor 2005, Harris and Lock 2006, Ridges 2006, Deeben et al. 2007, Kay and Witcher 2009). The crux of the argument is that, by focusing on environmental factors, modellers were potentially excluding important cultural factors (e.g. trade, religion) which could influence site locations.

It is unlikely that some archaeologists will ever accept the use of environmental factors from a theoretical standpoint. Furthermore, many archaeologists do not know how to include these non-environmental factors into models (Kohler 1988, Brewster and Reddy 1999). In addition, representing cultural factors using GIS is time-consuming and difficult, putting off other model makers from the idea (Judge and Sebastian 1988, Kvamme 1988b, Gaffney and van Leusen 1995, Joolen 2003, Kvamme 2006, Deeben et al. 2007) while environmental data is the easiest to obtain (Kvamme 1990, Kvamme 2006) and cheaper (Dore and Wandsnider 2006). Nevertheless, many recent models have branched out into cultural or semi-cultural based predictive modelling (Lock and Harris 2006; Ridges 2006; Verhagen, Kamermans et al. 2007) to challenge this assumption that predictive models are based solely on environmental factors. This was considered a potential topic of investigation.

13. 'GIS data have insufficient resolution and poorly represent the real world.'

14. 'GIS data are inaccurate.'

(Kvamme 2006 p. 6)

This is a problem that has been widely discussed in the literature (Altschul 1990, Carmichael 1990, Kvamme 1990, Savage 1990, Warren 1990a, Allen 2000, Brandon et al. 2000, Ebert 2000, Dore and Wandsnider 2006, Stancic and Veljanovski 2006, Kamermans 2007, Cushman and Sebastian 2008). Furthermore, how GIS programs interpret raw data can cause errors, as is the case when using some DEM datasets and different interpolation programs (Marozas and Zack 1990, Warren 1990a, Bennett and Hageman 2000). These problems, along with such issues as sample size, software used, datum of the data, data entry, etc. (Kaufmann 2006) makes the accuracy of GIS data a major concern for archaeologists.

Even though this is a real concern of modellers, the effects of such problems are diminishing with time. For example, it took the United States Geological Survey (USGS) from 1935 to 1992 to create 1:24,000 scale topographic maps of the whole United States of America (USGS 2012a). Now this work is done by satellites in a matter of months. It is now possible to obtain LiDAR data with 50cm resolution of most of the globe for predictive modelling (Mesterházy, Padányi-Gulyás et al. 2012). Data correction techniques are now catching many of the errors that can occur in the data. With these advances, not just in predictive modelling but in GIS and data collection, there was little room to contribute to this topic.

23. 'The most interesting sites are the (idiosyncratic) ones that do not fit the pattern.' (Kvamme 2006 p. 6)

Idiosyncratic sites are those sites described as 'red flag site' in Chapter 1. An argument has been made that these sites are more important because they are not predicted (Altschul 1990, Hasenstab and Resnick 1990, Altschul et al. 2004). The criticism is that because these sites are not in high-priority areas, according to the predictive model, they will be ignored or, worse, destroyed by development, which is not the purpose of a predictive model (Berry 1984, Gaffney and van Leusen 1995, Harris and Lock 1995, Ebert 2000, Wheatley 2004, Dore and Wandsnider 2006, Kamermans 2007, Verhagen 2007a, Kamermans et al. 2009b).

> 'Once anomalies, or red flags, are identified they become the subject of additional research. As patterns are found, many anomalies become predictable. Those sites whose locations remain anomalous grow in importance. Archaeologists want to know about these sites to further our insight into the past. Managers want to know the locations of these sites so that they can be included early in project plans.' (Altschul 1990 p. 288)

This is the issue with performance of models discussed at the beginning of this chapter. During my Masters research I examined some of these 'red flag' sites and found that they were the result of errors in site coordinates (Rocks-Macqueen 2010). The reason the model did not predict sites being at those locations was because they were, in fact, not actually there. Mislocation of sites in databases could explain these 'red flag' sites for some models. Some sites may be missed because of the resolution of the data used (Padányi-Gulyás, Stibrányi et al. 2012) or how the data were incorporated into a GIS program (Marozas and Zack 1990, Warren 1990, Hageman and Bennett 2000). This would indicate that 'red flag' sites are actually data errors and model performance may be related to that.

> 1. 'Many archaeological sites are buried, and we cannot model them because we do not and cannot know their distribution.' (Kvamme 2006 p. 6)

The potential issue is that buried sites might be missed during surveys or testing and excluded from the dataset used in the model creation or in model testing. This exclusion could result in false correlations or false assumptions; basically, models that are wrong (Savage 1990, Asch and Warren 2000, Dore and Wandsnider 2006, Berger and Verhagen 2007, Verhagen 2007c, Verbruggen 2009). This same line of thinking can be applied to other issues, such as a site being destroyed either through natural occurrences, like erosion, or man-made damage, such as looting or ploughing (Dore and Wandsnider 2006). Which is essentially this issue:

> 10. 'Models based on site presence-absence criteria are mis-specified because one cannot assume site absence.' (Kvamme 2006 p. 6)

Basically, both issues can be summarised as: one cannot assume that because a location has been reported as not containing a site that no site exists there, or did not exist at some point in time.

Is this still a problem for site predictive modelling? Surprisingly, no one had tested this assumption that missed sites would affect the outcome of a specific model, or attempted to quantify how many sites are supposedly missed. Still, methods have been developed to combat this issue:

- Fuzzy logic methods allow the incorporation of expert opinion on the likelihood of such problems (Bailey et al. 2009).
- With the Dempster-Shafer methods it is possible to conduct surveys or sub-level tests looking for buried deposits and then incorporate this into the current or previous predictive models (Ducke et al. 2009).
- The Minnesota State-wide model took into account locations that would have deep buried deposits, thus not found in most site databases (MDOT 2009).
- In a case from northeastern Oklahoma, archaeologists modelled where sites would be buried due to a soil-geomorphic model and then proceeded to find ones that had been missed by previous surveys (Artz and Reid 1983).
- Verhagen and Berger's model in France specifically used buried archaeological sites to create the predictive model (Berger and Verhagen 2007).
- There has been discussion on how to undertake a sampling process to look for such buried sites (Tol and Verhagen 2007).

These are just examples dealing with buried sites. Advanced models have recently been created to map not just site density but also the likely preservation of sites (Verhagen 2006). Kamermans has called this process 'land evaluation' and has used it with the evaluation of site locations in the Netherlands, see Kamermans (2000) for more details. This is a method advocated by other archaeologists (Brandon et al. 2000).

To say this issue is solved for all predictive models would be misleading, as there is potential for it to occur. However, there has been significant work undertaken to create methods to address this problem, as listed above. This has made the exploration of this problem for this project less appealing, as it was hard to see how one could improve upon the methods that already address the issue.

> 2. 'Known site distributions in existing government files and databases are biased because of (a) the haphazard way in which many were discovered and (b) variations in obtrusiveness, visibility, and preservation.' (Kvamme 2006 p. 6)

Part (b) is addressed in the previous section but part (a), survey bias, is still a problem. Several solutions to survey bias have been put forth. It has been suggested that to avoid potential biased data a project can collect its own data in a less haphazard way, which is well discussed in the volume *Quantifying the Present and Predicting the Past,* in the chapter on data collection (Altschul and Nagel 1988). This data collection process has been both accomplished (see Asch and Warren (2000)) in the past and has a detailed methodology laid out (Tol and Verhagen 2007). One does not have to rely on data gathered by others to create a predictive model.

However, in some cases collecting one's own data may not be feasible, as data collection costs money and projects may be working on a small budget. Moreover, just because one collects one's own data does not mean it is not biased in some way. As such, recently more methods have been created to address the issue of biased data. The use of Dempster-Shafer or Bayesian methodologies can be used to offset potentially biased data (Verhagen 2006, Wescott 2006). Briefly, those methods are as follows:

Bayesian

> '… it is impossible to separate opinions (prior beliefs), data and decisions/actions. In the "classical" approach, our opinions influence our procedures in all sorts of subtle and little-understood ways, for example in choosing the significance level of a hypothesis test. It's better to be as explicit as we can about our prior beliefs, and let the theory take care of how they interact with data to produce posterior beliefs, rather than to let them lurk at the backs of our minds and cloud a supposedly "objective" belief. This way the Bayesian approach can be more than just a nice piece of mathematics." (Orton 2003)

The basic concept of Bayesian statistics is:

posterior belief = conditional belief * prior belief

Equation 9: Formula for Bayesian statistics

In predictive models, prior belief would be what we know about current site locations. Conditional belief is new information used to adjust the traditional predictions to get the correct result, i.e. posterior belief. A hypothetical example of this would be an area where site density is one site per every 10 acres; the conditional belief. If you were to survey 20 acres you would expect to find two sites. But you know that when you survey land next to rivers you find one site per every one acre surveyed. In this hypothetical example, if you were to survey 20 acres but 10 of those acres were next to a river you would expect to find 11 sites (10 sites in the 10 acres next to the river and one site in the other ten acres). Bayesian statistics are as simple as taking into account new information to adjust previously held data.

For the use of Bayesian statistics in predictive modelling the prior belief has usually been set as the current site density per area unit for the study area. For some projects, previous predictive models have been used to set the prior belief (Millard 2005, Verhagen 2006, Wescott 2006). For other projects, expert judgement was used as prior belief (Verhagen 2006, Ducke et al. 2009), see Ducke et al. (2009) for an example of expert judgement. For conditional belief, some archaeologists have used expected versus actual site located ratio of the study area (Verhagen 2006, Clarke et al. 2009). Ultimately, this method compares old models/beliefs/interpretations with new data and then makes an adjustment layer to represent a new, more accurate or precise result through statistical procedure.

This approach has the advantage of allowing greater flexibility. When using previous site predictive models, regardless of type, it is possible to make improvements to the data without having to completely re-calculate every aspect of the model (Verhagen 2006). As new data is acquired these Bayesian models are automatically updated (Verhagen 2006, Wescott 2006, Finkea, Meylemans et al. 2008). In effect, modellers no longer have to completely disregard previous models because of flaws in the data but can build upon these imperfect datasets and adjust the models accordingly. It also means that biases in data collection can be taken into account and adjusted.

Dempster-Shafer

Dempster-Shafer Theory (DST) is built around the concept of *belief*, a generalised version of mathematical probability (Ducke et al. 2009). This method enables the use of uncertainties and beliefs as inputs for a predictive model alongside more traditional datasets (Canning 2005). The method works by taking one or more hypotheses and comparing them against variables that might be related to the belief of the hypothesis's outcome (Canning 2005, Stancic and Veljanovski 2006). 'Belief' being different then to 'probability', as the latter involves a more rigid mathematical framework. This is done through the following mathematical formula:

$$m(A) = m_1 \oplus m_2 = \frac{\sum_{B \cap C = A} m_1(B) m_2(C)}{\sum_{B \cap C \neq 0} m_1(B) m_2(C)}$$

Equation 10**:** Formula for Dempster-Shafer

This chapter will not go into the details of this method as the best description can be found in Canning (2005) and Ducke et al. (2009).

Even if Dempster-Shafer or Baysian methodologies are not used predictive modellers can simply list areas where the data may be weak or potentially captured in a bias way, as was the case with the MDOT model (Figure 13). That model labelled areas that had very little archaeological data but high potential as 'suspected high' while areas with strong survey data and high potential as 'high'. These adjusted users' expectations so that they realised the model could be wrong in certain areas where the data was weak. Signposting can be an effective tool when dealing with possible problems with the data.

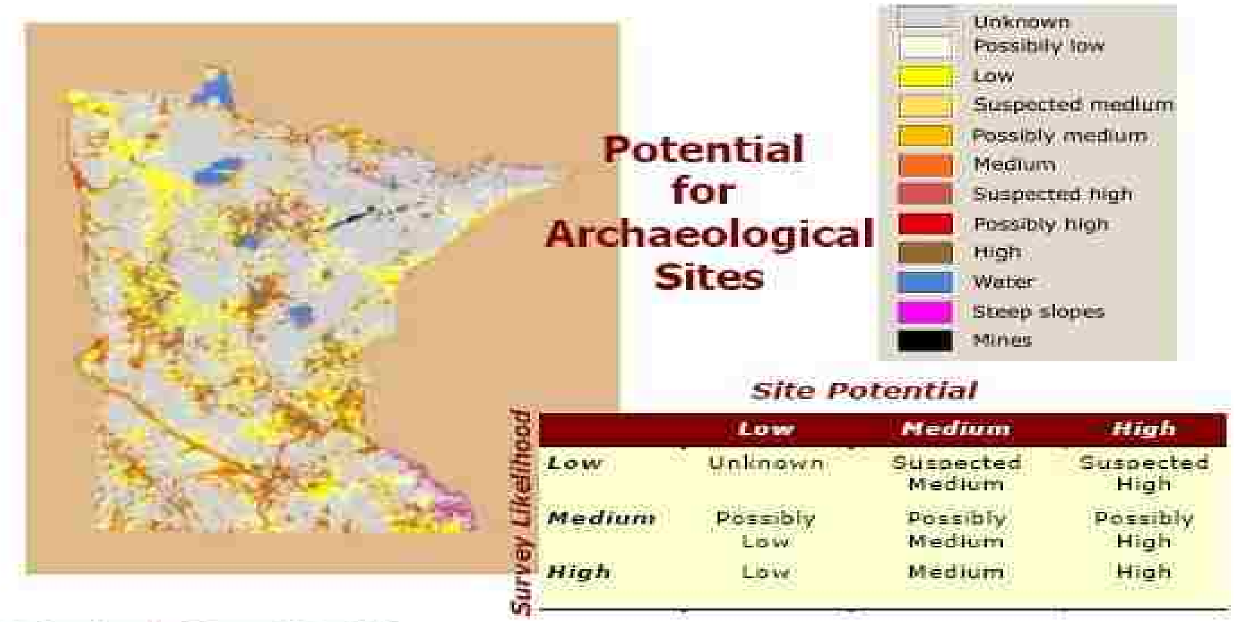


Figure 13: Minnesota Department of Transportation Assignment of Values for Confidence in Archaeological Site Probability. From MDOT website (MDOT 2009).

A final point is that the limited testing of haphazard gathering of data in database does not always affect models. The PUMP III project in New Mexico (Altschul et al. 2005) (discussed further in Chapter 5) was able to create a time depth analysis of databases to examine the biases in haphazard data collection resulting from CRM work and found that they had no effect on the predictive models (Figure 14). Further testing is needed but haphazard collection of data through CRM work might only be relevant in a few cases for predictive modelling.

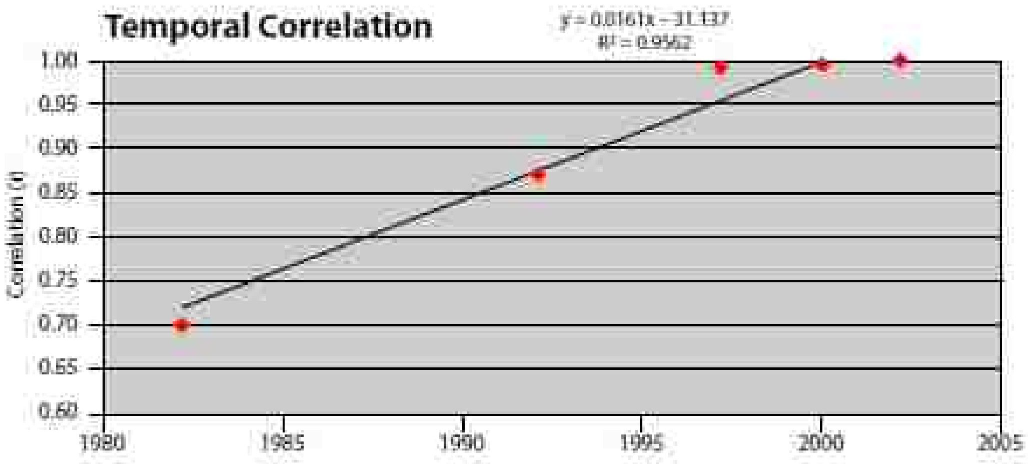


Figure 14: Correlation of logistic model by year of data available for PUMP III project. Very little deviation in models results due to data being added between years (Altschul et al. 2005 Figure 6.18 p. 101).

Given the work to develop methodologies to address the issue of biased data and that in some cases that haphazard data collection does not affect model outcomes there was little scope for further improvement.

Method-specific Problems

There are issues with specific methods as well as general predictive modelling problems.

> 16. 'Models based on statistics cannot meet random-sampling assumptions because most extant data were not obtained by random sampling.' (Kvamme 2006 p. 6)

Some statistical methods require random-sampled data (Beckman and Duncan 2000). This can be hard to achieve in archaeology. Site databases are the result of many different projects and surveys being compiled together (Kuiper and Wescott 2000, Kamermans 2007, Verhagen 2007c). Even when random-sampling surveys are undertaken they are not done correctly for predictive modelling (Berry 1984, Verhagen 2007c) or these projects can have different standards or goals that lead to errors in the data and a distinctly non-random sampling strategy (Marozas and Zack 1990, Garcia Sanjuan and Wheatley 1999, Beckman and Duncan 2000, Kuiper and Wescott 2000, Altschul et al. 2004, Kvamme 2006).

> 17. 'Models derived from random cluster sampling are mis-specified because they do not adjust for underestimated variances.' (Kvamme 2006 p. 6)

This is what has been referred to as the Teotihuacan Quandary (Altschul and Nagel 1988). It is so named because some sites, due to their political power and cultural significance, like Teotihuacan in Mexico Valley, would greatly influence the location of the surrounding archaeological record but could easily be missed by random sampling (Altschul and Nagel 1988). Of course, anyone conducting a survey of the Mexico Valley region would not miss Teotihuacan, due to its importance and immense size, but it would have to be excluded by the strict requirements of random sampling. Correlative methods are stuck in a problem where they must exclude important sites to meet random sampling but by excluding important sites they are potentially damaging their models.

There are additional problems with correlative models, and in particular with the regression-based models:

- Some modellers misinterpret results, as the results are usually relative and not absolute in mathematical terms. The answers provided by these methods do not specify that area X has a higher probability of having archaeological resources than area Y. In practice, the model suggests that X's probability is Z and Y's probability is V. Misrepresenting this outcome is quite common in many site predictive models (Woodman and Woodward 2002).
- Collecting data samples can be difficult as tests have shown that significant amounts of the landscape need to be surveyed: <10% in (Altschul et al. 2005), 7.5% in (Gazenbeek and Verhagen 2007) otherwise the model will fail (Stancic and Veljanovski 2006). This requirement cannot always be met from existing data and the acquisition of this additional

data can be costly and time-consuming, or impossible in some cases (Kuiper and Wescott 2000).

- Problems specific to linear regression algorithms mean that this equation assumes a linear relationship between sites and landscapes. This means that linear regression models fail with non-linear datasets (Warren 1990a). Furthermore, it is necessary for the data to be measured on an ordinal scale or higher. Unfortunately, spatial data is frequently nominal rather than ordinal (Gillings and Wheatley 2002).
- A technical problem with the logistical regression method is that it can be biased towards prediction of the larger class examined. In archaeology, the larger class is usually site absence (Gillings and Wheatley 2002).

Other traditional methods, like Boolean and Weighted, have problems of their own that make them just as undesirable. Boolean methods are very poor for explanatory modelling or modelling in general as they result in binary mutually exclusive outcomes. As mentioned in Chapter 1, even if an area fulfilled 99 out of 100 of the desirable characteristics for containing sites it would still be marked as poor by Boolean methods. Medium or mild archaeological intensity cannot be captured using a binary, mutually exclusive, result (Altschul et al. 2004).

Weighted methods have been found to rely on overly simplistic assumptions, e.g. people lived near water; people liked flat land. These are criticised because they ignore theoretical approaches to landscape and settlement, except in the most basic sense. Moreover, many weighted methods undergo little testing to determine to what extent each attribute influences the model (Brandon et al. 2000).

More Methods, More Problems

It is for these very reasons that there has been a recent push to promote and develop new methodologies (Brandon et al. 2000, Kamermans and van Leusen 2005, Mehrer and Wescott 2006, Verhagen 2007a, Kamermans et al. 2009a). However, the limited testing of these new methods against older methodologies has not proven to be very successful (Ducke et al. 2009) and there remain issues with these methodologies.

For instance, testing the Dempster-Shafer method has resulted in mixed outcomes. This method derives from research in the Netherlands, but when applied to different periods it returned varied levels of performance (van Leusen, Millard et al. 2009). In some instances, this method was found to be an improvement on previous methods (Palaeolithic and Mesolithic period sites), but in other instances it was found to perform worse (late medieval period sites). Since the model has not provided significant improvements over previous methods it remains to be seen if there is scope for significant research to be devoted to this method in the future.

A serious problem with Bayesian methods is that any data used to establish conditional belief needs to be independent of the previous data and non-biased (Millard 2005, Verhagen 2006, Clarke et al. 2009). This may sound simple, but obtaining new data can be complex and time-consuming.

> 'The bottleneck in this statement is the fact that the new data should be independent from the old data in order to be able to adapt the posterior belief; otherwise there is the very real danger of self-fulfilling prophecies. This is precisely the problem of using archaeological predictive maps for guiding surveys: there will be a natural tendency to select those areas where high site densities are predicted, and this may lead to an ever-increasing amount of biased-sample data.' (Verhagen 2006 p. 187)

Another problem with Bayesian statistics is that it is not capable of dealing with revisions of the original hypothesis. If a model is wrong and needs additional parameters, then it will have to be completely rebuilt. This has led those who have tested this method to conclude:

> 'The claim, sometimes made, that Bayesian statistics constitute a superior way of doing statistical analysis, is in my view exaggerated. Given the doubts and complexities surrounding the subject, I have not further pursued its development.' (Verhagen 2007f p. 91.)

Summation of Issues

After reviewing the issues with predictive modelling and whether they were being resolved or not, the following table of problem status was created to help visualise areas that still needed to be solved:

Category	Problems Raised with Predictive Modelling
General to all of archaeology or unknown how it affects predictive modelling	6. 'We must be able to model and understand the archaeological formation process, both natural and cultural, before we can model where sites might be found.' 20. 'Environmental variables shown to be important to site locations may only be proxies for variables that were actually important.' 22. 'One must understand and model complete behavioural systems before archaeological models can be built.'
Problems with methodologies created to solve the issues	1. 'Many archaeological sites are buried, and we cannot model them because we do not and cannot know their distribution.' 2. 'Known site distributions in existing government files and databases are biased because of (a) the haphazard way in which many were discovered and (b) variations in obtrusiveness, visibility, and preservation.' 4. 'One cannot model archaeological site distributions because "site" is a meaningless concept; human behaviour did not occur in discrete bounded areas but formed a continuum over the landscape.' 9. 'Models based on landscape variables are meaningless.' 10. 'Models based on site presence-absence criteria are mis-specified because one cannot assume site absence.' 11. 'Blue-line features on topographic maps are frequently arbitrary and unreliable indicators of water.' 13. 'GIS data have insufficient resolution and poorly represent the real world.' 14. 'GIS data are inaccurate.'

Problems with potential for further research, but some with unknown extent of issue	3. 'Many known sites are inaccurately located on maps and in databases.' 5. 'Functional, temporal, or cultural site types cannot be readily determined for most sites in an archaeological database, yet profound locational differences must exist between the types.' 7. 'We do not know the locations of resources important in the past times, such as water sources, springs, edible-species distributions, lithic raw material sources, and the like.' 8. 'Past environments were very different from present ones, so we cannot model the past based on the present.' 12. 'Modern soil types are meaningless because they are changed from the past and, in any case, are frequently irrelevant to past farming practices.' 18. 'Grouping sites of many types into a single, site-present class creates too much variability to model.' 23. 'The most interesting sites are the (idiosyncratic) ones that do not fit the pattern.'
Method-specific	16. 'Models based on statistics cannot meet random-sampling assumptions because most extant data were not obtained by random sampling.' 17. 'Models derived from random cluster sampling are mis-specified because they do not adjust for underestimated variances.' Some modellers misinterpret results. Collecting data samples can be difficult as tests have shown that significant amounts of the landscape need to be surveyed for regression methods. Linear regression algorithms assume a linear relationship between sites and landscapes. Logistical regression method can be biased towards prediction of the larger class examined. Boolean assumes mutually exclusive binary outcomes. Weighted methods can be simplistic. Bayesian methods: data used to establish conditional belief needs to be independent of the previous data, and non-biased, which can be difficult to obtain.

Table 3: Project categories of problems facing predictive modelling.

There were issues that had been addressed in the preceding years since they were first raised. Most methods also have various issues, some of which are inherent in the methodology and cannot be fixed, e.g. Boolean methods are binary and so mutually exclusive. There were issues with potential for further research as they had only been briefly been explored or were unresolved.

While considering which of these problem to pursue another possible avenue of research presented itself as one worthy of investigating. A project in Australia was able to improve the performance of previously created predictive models by adding cultural datasets to environmental ones (Ridges 2006). Ridges' first model was able to capture 71% of the known sites in the high-probability level. This model was then modified with an explanatory preference for site location, in this instance the cultural preference for rock art. That work was based on a solution put forward by van Leusen and influenced by Zubrow (1994):

> 'By applying an ED (environmentally deterministic) model to a dataset, one can eliminate environmental patterning in the data, leaving a clearer view of whatever cultural factors may influence the data.' (Gaffney and van Leusen 1995 p. 370)

The result was an improvement in the model results from 71% to 75% of sites captured (Ridges 2006). Trying to determine why people choose locations, as opposed to simply looking for correlations, improved the performance of the predictive model.

This was not a one-off experiment either; similar work by Maschner came to the same conclusion (Maschner 1996). He found that correlative models based on environmental factors accounted for the majority of site locations but that explanatory aspects were important for some of the sites in his study area. Complementing these findings are a steady stream of archaeologists advocating looking at more explanatory sources for some site locations (Gaffney and van Leusen 1995, Allen 2000, Brandon et al. 2000, Ebert 2000, Whitley 2003, Whitley 2006, Goel et al. 2009, Verhagen and Whitley 2011). Other archaeologists have specifically called for more explanatory based work in combination with CRM goals to improve models (Whitley 2004a, Verhagen and Whitley 2011).

This approach is intertwined with a larger issue, that predictive modelling use has been primarily limited to CRM and it has not seen widespread use by university-based archaeologists:

> '… the development of predictive modelling has veered away from mainstream archaeological thought and theory and has now become a largely self-contained activity – enjoying reasonable success as a tool for CRM, but not commanding much respect from academic scholars ...' (Verhagen and Whitley 2011 p. 50)

In general, academic archaeologists have been sceptical of, and sometimes even averse to, predictive modelling as practiced in CRM (Kamermans and van Leusen 2005). This is due to a failure to engage theory or an explanatory framework in models (Brandon et al. 2000, Ebert 2000, Harris and Lock 2006, Verhagen and Whitley 2011).

This difference between CRM and university-based archaeologists' interest has sometimes been portrayed in binary fashion and using different terms (Table 4) in predictive modelling literature. This has further tied into debates in the predictive modelling literature about inductive and deductive methods. However, these descriptors are not very accurate; for example, the first nation-wide predictive model for the Netherlands used inductive methods (Verhagen 2007a, Kamermans et al. 2009a).

Inductive	Deductive	Source
Correlative	Explanatory	(Judge and Sebastian 1988)
Data driven	Theory driven	(Gillings and Wheatley 2002 p. 149)
American	European	(Lock and Stançic 1995)
CRM (Cultural Resource Management)	Academic	(Podobnikar, Oštir et al. 2001, Whitley 2003, Stancic and Veljanovski 2006)
Bottom Up	Top Down	(Podobnikar and Šprajc 2007 p. 3–4)

Table 4: Historic divisions in the descriptions of site predictive modelling methods.

In the author's personal opinion, the divisions of 'CRM' and 'Academic' archaeology are oversimplifications and not very productive. That being said there is agreement that predictive modelling has not been as popular with academic archaeologists as it has been with archaeologists working in CRM. The lack of interest in the academic sector has had a detrimental impact on the use of predictive models throughout archaeology. For example, in the UK most academic archaeologists who have weighed in on the subject of predictive modelling have been mainly critical (Gaffney and van Leusen 1995, Gillings and Wheatley 2002, Wheatley 2004). Consequently, there have been very limited uses of predictive models in the UK, with only two published models being completed (see Wilcox (2009) or Graves (2010)), and only by postgraduates not under the supervision of those critical of predictive modelling. Of course, this is a complex problem and the aversion of UK academic archaeologists may not be the only factor that influences such an outcome. However, the vast majority of archaeologists receive their initial training at universities in Europe (Aitchison, Alphas et al. 2014) and North America (Zeder 1997). It is this formative experience that determines the views archaeologists carry throughout the rest of their careers (Zeder 1997). If predictive modelling were to move beyond a niche use and gain wider usage and acceptance in archaeology, it would require the interest of archaeologists based in universities who can pass on this interest and the required skills to future generations of archaeologists.

Initially, this project aimed to focus on performance issues and not those of theoretical disagreement, but further investigation presented the possibility that they were linked. The proposition of both increasing performance and gaining wider use/acceptance of predictive modelling at universities seemed like an excellent direction to take the research into. If moving away from environmental correlative and into more explanatory modelling could accomplish this then this seemed like the best route to take. The project's objective was thus set to **increase the explanatory abilities of predictive modelling**. The next chapter goes into the details of how this project creates a methodology to meet that objective.

Hindsight

This chapter has presented the views held at the beginning of the project, but as will be demonstrated later in this thesis they were not always correct. The impact of blue-line features, i.e. water locations on maps, was underestimated at this point in the project, being dismissed as not a high priority issue. This and several other assumptions turned out to be incorrect but this could not have been known when the project began.

Chapter 3: Routes to Improving Predictive Models

The review of the history of predictive modelling showed its development over the years and how at the beginning of this project it had not yet reached its fullest potential. Chapter 2 presented many of the problems holding back predictive modelling. It presented the aim of this project – to improve the performance of site predictive models for CRM archaeology uses. The review also found that in several instances more explanatory methods of predictive modelling could lead to improved performance and that many archaeologists have been calling for such models to be made for decades now (Gaffney and van Leusen 1995, Allen 2000, Brandon et al. 2000, Ebert 2000, Whitley 2003, Whitley 2006, Goel et al. 2009, Verhagen and Whitley 2011). This meant that the project created an objective to increase the explanatory abilities of predictive modelling. This chapter explores in more detail what was envisioned as an explanatory predictive model and the tools that would be tested to create one, agent-based modelling. This chapter covers part of the second planned activity for this project — to create a methodology to solve the problem(s) that lead to poor model performance.

Problems to Solve

Improve the explanatory abilities of predictive modelling – newer methods, while full of potential, have not yet proved themselves capable, as the review of methodological problems in the previous chapter demonstrates. Most of the methods to create predictive models involve correlative methods and their performance (see Table 1) and utility were doubted. As Verhagen has said of some of the newer correlative methods, ‘Given the doubts and complexities surrounding the subject, I have not further pursued its development.’ (Verhagen 2007f p. 91.) Besides the technical failures of these methods most lack any way to explain why sites are located where they are. Inductive models can show connections between different variables and site locations but cannot explain why those connections exist (Brandon et al. 2000, Ebert 2000, Harris and Lock 2006, Verhagen and Whitley 2011). While they can test archaeological theory, it is very hard to incorporate theory into the models. The so-called deductive-leaning/explanatory methods, like the weighted or Boolean methods, are simplistic and in many cases provide less information about possible past human behaviour than the correlative methods because they are reduced to simple assumptions such as distance to water, slope, etc. without any consideration of what those assumptions actually represent. Lock and Harris discussed such generalisations of site locations and the potential to not understand cause and effect in one of their papers:

> ‘… now consider the following predictive model. In similar fashion to the above approach, a new study area was designated that again emphasised distance to water (buffered from 100, 500, and >500 m), low slopes and valley bottomlands (<18°), fertile alluvial soils, favoured

> old-field sites, and at a preferred elevation from just above sea level to 300 m and up to 770 m (Figure 2.2). Without stretching the point too far, these parameters are not greatly dissimilar to those specified above for the prehistoric site model. However, these latter parameters correspond not to the location of archaeological sites, but to the habitat of the common sycamore tree *(Platanus occidentalis L.),* a fast-growing, long-lived tree and one of the most common trees in eastern U.S. deciduous forests (Wells and Schmidtling 2001) …' (Harris and Lock 2006 p. 43)

These issues with past methods led this project to explore new possible tools and methods for creating predictive models.

New Methodology

This idea of pursing more explanatory predictive modelling did not occur in isolation. A little over ten years before this project began several predictive modellers, Brandon, Burgett and Church (2000), laid out a plan to create such a model through an eight-point manifesto in their paper, 'GIS Applications in Archaeology: Method in Search of Theory'. This raised issues that were incorporated into the early thinking that drove the project's planning, while others were not taken into account. The first point of the manifesto was,

> 'We advocate that the body of theory and methods that have come to be termed "landscape ecology" has much to offer to the study of prehistoric populations. We are fully cognisant of the pitfalls in borrowing from other disciplines. In regards to this we agree with Keene, who stated, "The source of the problem is borrowing without modification and a tendency to adopt rather than adapt." (Keene 1983 p. 142) However, we find that arguments against any use of ecological method and theory in archaeology are provincial and arrogant.' (Brandon et al. 2000 p. 159–60)

This was a rebuke to the Post-Processual critique that began in the 1980s and carried on into the 1990s about the environmental determinism aspect of predictive modelling. Their second point continued that rebuke but also acknowledged that cultural factors have a role to play in site location:

> 'We accept the argument that, as societies have developed, the constraints imposed by the surrounding environment have been increasingly mitigated by cultural responses. However, we believe that during almost all of North American prehistory ecological forces have limited and shaped prehistoric population activities to a substantial degree. We therefore argue that an understanding of the ecological system and its interaction with the geomorphological and culture systems is essential to interpreting the archaeological record.' (Brandon, Burgett et al. 2000 p. 159–60)

These were interesting arguments but ones that were not applicable to all sites, regions or cultural histories around the world and I would argue not even true for 'almost all of Northern American prehistory' either. As such the use of 'landscape ecology' as the only body of theory was not incorporated into the theoretical drive of this project. Such a theory was viewed as being incorporated into the models but not as the sole theoretical provision. It would be their other points that would be embraced at the beginning of this project when creating a methodology. For example, there was much to agree with the statement that, 'Arguments to the effect that human behaviour is too complex to model and, therefore, any attempt to model cultural systems will be so generalised as to be useless are nihilistic and passive in viewpoint. No one is denying the complexity of the task, but that is the challenge, not an excuse' (Brandon, Burgett et al. 2000 p. 159–160).

There was further agreement with concerns about correlative models,

> 'We strongly believe that the current additive strategy in archaeology where information from a number of points or sites is used as the base to build a picture of regional prehistory is theoretically shallow, methodologically costly, and ultimately misleading. The archaeological record is more profitably used to validate hypotheses generated by models than as a basis for model-building itself.' (Brandon, Burgett et al. 2000 p. 159–60)

Though several of the other points were not relevant to the goals of this project, there was a final point that was influential in conceptualising what the explanatory predictive model should look like:

> 'The key factor that makes this proposed approach superior to traditional correlative models is the flexibility of data. Traditional predictive models are generalised models of static variables. With the predictive model structure proposed here, managers can generate data that incorporates temporally variable aspects, post-depositional processes that might obscure, alter, or destroy the archaeological record, as well as flag areas having a high probability of sites, including those more specialised sites that are often ignored in traditional predictive models. "Understanding patterns in terms of the processes that produce them is the essence of science and the key to the development of principles for management." (Levin 1995 p. 278)' (Brandon, Burgett et al. 2000 p. 159–60)

At the very earliest stages of the project, in the Fall of 2010, some of the ideas presented by Brandon, Burgett and Church influenced some of the early conceptions of how to meet the goals of this project. Adopting parts of Brandon, Burgett and Church's manifesto such as not giving up because the process might be complex, injecting more theory into model creation and a flexible enough model to incorporate dynamic data, created a template to examine possible new methodologies. However, parts of their vision were also discarded. Focusing on a single theoretical framework when archaeologists have diversified the theory they use, was seen as counterproductive. Archaeological theory and methods had moved on since they wrote their paper and to ignore these developments would have alienated many archaeologists. Considering that one of the underlying drivers of this project was to garner the interests of academic archaeologists, who have different views on theory, this counterproductive aspect was dropped.

Enter Agent Based Modelling

With these guiding principles of stronger use of theory to drive modelling and the need to handle dynamic data sets, agent-based modelling (ABM) was stumbled upon, a chance finding when searching for terms ‘modelling’ and ‘archaeology’ in Google Scholar Search, as a possible tool/methodology that could meet the needs of this project. ABMs are a class of computational models used for simulating the actions and/or interactions of autonomous agents. An agent can be anything that can be represented as a single unit: a person, a plant, even a rock. If you can represent something as a single object, then it can be modelled by an ABM program.

In ABM each agent individually assesses its circumstances and makes choices on the basis of a set of programmed rules. These individual actions then combine to create more complex systems. The process is one of emergence from the lower-level decisions of systems to a higher level (Bonabeau 2002). Agents may undertake any variety of behaviours depending on the model-maker’s needs (Axtell and Epstein 1996, Axelrod 1997).

> ‘Repetitive competitive interactions between agents are a feature of agent-based modeling, which relies on the power of computers to explore dynamics out of the reach of pure mathematical methods.’ (Bonabeau 2002 p. 7280)

At the simplest level, an agent-based model consists of an agent and its relationships to its surroundings. Sophisticated ABM sometimes incorporates neural networks, evolutionary algorithms or other learning techniques, to allow learning and adaptation by agents. The models can be as complex or as simple as needed.

History of ABM

ABM began in the 1940s, with von Neumann and Ulam’s investigations into cellular automata and modelling of ‘cells’ and their interactions with one another through mathematics (von Neumann 1951). One of the earliest agent-based models was Thomas Schelling’s segregation model (Schelling 1971). Although Schelling used coins and graph paper rather than computers, the model personified the basic concept of agent-based models as autonomous agents interacting in an environment. In the early 1980s, Robert Axelrod hosted a tournament of Prisoner’s Dilemma strategies with an agent-based concept (Axelrod 1997). The first use of the term ‘agent’ as the definition currently used today is attributed to John Holland and John H. Miller’s (1991) paper (Hussain and Niazi 2011).

ABM can use complex and significant calculations which require significant computer power. Due to its high demand on computers ABM did not become widely used until the 1990s, with the increase in affordable computing. A major thrust for ABM was the book *Growing Artificial Societies* (Axtell and Epstein 1996). This showed how simple and interpretable rules for agents could simulate behaviour for a wide variety of disciplines such as sociology, anthropology and economics. By the end of the decade the use of ABM was growing exponentially and continues to do so (Hussain and Niazi 2011).

Swarm, an agent-based modelling program which was designed by Chris Langton at the Santa Fe Institute to explore 'artificial life' (Askenazi, Burkhart et al. 1996) was created in 1996. A variety of agent-based modelling programs have been created over the years; the next chapter reviews over 70 of them. Each program has its own advantages and disadvantages as well as common and unique tools. Currently no program dominates ABM, which means most models are created with different programs and cannot be easily transferred between platforms.

ABM in Anthropology and Archaeology

ABM-based work has gained prominence in archaeological and anthropological circles in recent years (Gumerman and Kohler 2000, Baden and Beekman 2005, Kohler and van der Leeuw 2007, Costopoulos and Lake 2010). This is partly due to the fact that archaeologists have played key roles in the development of the software for ABM. Kohler's models of group dynamics in southern Colorado originally started out as a GIS-based investigation, but quickly switched over to ABM (Kohler and Van West 1996). This work played a key role in developing ABM at the Santa Fe Institute (Gumerman and Kohler 2000). Archaeology has been involved in ABM almost from the beginning.

ABM is Simulation

Agent-based modelling is a form of simulation. Simulations are the action of experimenting/working with a model in order to understand system behaviour and/or its underlying causes (Shannon 1977, Breitenecker and Popper 2011). In this case a model is when, 'one creates some kind of simplified representation of "social reality" that serves to express as clearly as possible the way in which one believes that reality operates' (Gilbert 2008 p. 2). Simulations cover a much wider range of tools in archaeology, from 3D re-creations of sites in gaming engines to using algorithms to estimate past behaviours. Simulation in archaeology has a much deeper history than agent-based modelling stretching back to the 1960s and 1970s. Since this project finished, two reviews of simulation in archaeology, specifically focused on ABM but including much of this earlier work, have been published and are worth reading to gain a better understanding of the history, see Lake 2014, Cegielski and Rogers 2016.

Why use ABM/Simulation?

Formalisation, understanding and prediction were the three of the most relevant reasons, of many (Epstein 2008, Gilbert and Troitzsch 2009), to use ABM for predictive modelling. For example, formalisation means removing those generic statements seen so often in Boolean or weighted predictive models, i.e. sites are within 400m of water. Mathematical and computational modelling forces one to be precise. Using natural language to express theories and models explains work but are rarely precise (Epstein 2008). With ABM, ambiguities cannot be tolerated if the model is to run, as every aspect of the model needs to be laid out. Basic assumptions have to be tested and building a model highlights logical gaps and data gaps in the verbal models and theories. ABM 'allows us to formalise our thinking about how the past worked' (Graham 2009 p. 2).

Simulating the past helps us understand the past by enabling us to explore our assumptions systematically and experiment with them (Graham 2009). A model can be set up and run under changing conditions, as any variable can be introduced and changed. Once changes are made simulations are observed to see how the conditions affect the systems. Essentially, a series of 'what-if scenarios' can be run effectively. Such an approach can identify underlying dynamics and help illuminate causal relationships. This reaches the heart of trying to create explanatory predictive models, understanding relationships between archaeological resources and their location in landscapes.

The last aspect and most relevant for this project is prediction, though not in the same terms that most ABM modellers imagine. Traditionally, predicting with ABM was seen in the following way:

> 'If we can develop a model that faithfully reproduces the dynamics of some behaviour, we can then simulate the passing of time and thus use the model to "look into the future". A relatively well-known example is the use of simulation in demographic research, where one wants to know how the size and age structure of a country's population will change over the next few years or decades. A model incorporating age-specific fertility and mortality rates can be used to predict population changes a decade into the future with fair accuracy.' (Gilbert and Troitzsch 2009 p. 4)

For this project the future is now, but the idea is the same— faithfully model the dynamics of a behaviour to create an understanding of how that behaviour will manifest itself at some point in time, being past, present or future. A tool commonly used for prediction is exactly what this project needed.

Complex systems

An additional attraction for ABM was its ability to deal with complex systems. One of the influences when searching for a new predictive model method was the idea that,

> 'Arguments to the effect that human behaviour is too complex to model and, therefore, any attempt to model cultural systems will be so generalised as to be useless are nihilistic and passive in viewpoint. No one is denying the complexity of the task, but that is the challenge, not an excuse.' (Brandon, Burgett et al. 2000 p. 159–60)

This is an issue of complex systems, which are systems that are composed of many different parts. Interactions of the systems' different parts lead to unpredictable behaviour at the system level. In other words, cause and effect relations are non-linear. Consequently, system behaviour cannot be explained by looking at the system parts (Bentley and Maschner 2003, Kohler and van der Leeuw 2007). ABM is well adapted to deal with such systems. ABM can model this system behaviour through actions and interactions of individuals (agents) from the 'bottom up'. 'The ability to study non-linear dynamics generated from the bottom up not only distinguishes agent-based models from top-down, deterministic models, but it also makes them especially attractive to social scientists' (Premo 2008 p. 44). This ABM an even more attractive tool for this project.

Basically, with ABM it is possible to build artificial societies and landscapes which can then be experimented with (Dean, Gumerman et al. 2000, Premo 2008). Jim Doran has stressed that ABM allows the exploration of human cognition (Doran 1999). Such a process enables archaeologists to explore our ideas about the past, experiment with those ideas and test hypotheses.

Issues with ABM

While there are many benefits to using ABM there are drawbacks to its use. Creating an ABM model can require a substantial effort in time and other resources, even with simple models (Gilbert 2008). ABM enables the exploration of ideas about the past but that means it is not modelling the past itself but rather ideas about the past (Premo 2008). A simulation does not tell archaeologists how things were even when a simulation's results perfectly match the archaeological record. A match with the archaeological record is not conclusive proof that that is how it happened in the past (Premo 2008, Premo 2010). This has led to a discussion among archaeology modellers about the direction agent-based modelling should be headed (Costopoulos and Lake 2010, Lake 2010, Premo 2010). This discussion centres around whether models should try to emulate the real world as closely as possible or be as simple as possible and be mainly used for exploration and experimentation.

For this project a simulation not exactly representing the past was not an issue in one sense. The goal of the project was to improve performance. If the results could do that, then whether the simulation exactly represented the past or not was not important. However, a subtext to the need for explanatory models was to increase the interest of other archaeologists in predictive modelling. If this is an issue for some archaeologists than this is a concern for the project. Even then, one has to wonder why this has been specifically targeted at modelling when this is applicable to all of archaeology. Barring the creation of a time machine to check, all archaeological knowledge is assumptions made with the best available information. There was no reason to think that ABM was any more problematic than archaeology in general.

ABM for Predictive Modelling

Serendipitously for this project, there were other predictive modellers, Verhagen and Whitley, concerned about the lack of explanatory aspects in predictive modelling and how this had caused a siloing of its use, who had a paper published early in this project's development (online first on February 1st 2011) to address this issue of explanatory modelling. Verhagen and Whitley used GIS to attempt to achieve their goal of a theory-based predictive model. Because their models were agent-focused but did not use ABM they referred to their models as agency/agent models. Agency in archaeology has unfortunately been turned into a catch-all term to describe a plethora of different thoughts. Dobres and Robb (2000) mention that out of 16 papers on the subject in archaeology there were 12 different definitions for the term 'agency'. The only common point that was agreed between these papers was that 'agency' should be included in archaeological research.

> 'Archaeologists appear to actively employ popular terms such as "agency" without specifying or perhaps even understanding their definitions.' (Verhagen and Whitley 2011 p. 67)

Originally developed out of a series of small projects, this methodology has progressed into a defined endeavour to instil theory in predictive models so as to stimulate academic interest into predictive modelling (Whitley 2000, Whitley 2003, Whitley 2004a, Whitley 2006, Goel et al. 2009, Burns and Whitley 2012). Verhagen and Whitley (2011) have attempted to define the general concept of agency within predictive modelling as follows:

1. Any investigation of behaviour that deals with spatial decisions and placement would incorporate a perspective held by the agent. The understanding of past human behaviour should be based on situations in which agents are presented with local conditions, that may or may not be representative of regional or global conditions, from which decisions are made on.
   1.1. Agents are most usually represented as an individual but group agency is possible. Some material culture manifests itself as a group decision instead of that of an individual.
2. Agency encompasses the Post-processual emphasis on pluralism. That is the idea that there are multiple different angles from which to look at a problem. This is seen as the way forward for archaeological sciences. BUT, this embrace of pluralism does not include the 'anything goes' attitude that is seen in some Post-processual writing. Instead, pluralism is incorporated into agency in the refutationist tradition, where multiple models and theories are tested and compared. The goal of this is to reveal weaknesses in those models or theories and thus contribute to the advancement of theory (Bell 1994).
3. By not embracing the 'anything goes' mantra, agency-based predictive modelling does not accept all interpretations as being equally valid. This is in direct conflict with some Post-processual thought that advocates the belief that all interpretations are valid and that only the interpreter needs convincing of the validity of their beliefs. Predictive modelling, which is created for the needs of modern people and societies, cannot abide by such an extreme and simplistic view of archaeology. Models are created to address the needs of laws dictated by the current social and political views of modern people. To fulfil their purpose, models will need to follow those beliefs and constraints to their operations. This means not all sites or landscapes

will be considered of equal importance and neither will all archaeological explanations and interpretations. This is a compromise that must be reached to satisfy both groups, CRM and academic, that are to benefit from this model.

(Verhagen and Whitley 2011 p. 60-63, 70-73, 86-87)

To use an idiom, Verhagen and Whitley attempted to thread the needle between Post-Processual and Processual schools of thought by embracing aspects of both. It takes on board the Post-Processual view of the individual (agent) as a driving force in the phenomenon that creates the archaeological record and that there can be multiple interpretations of the past, yet embraces more Processual-leaning concepts like refutationism.

This is not the first attempt at a bridge between Post-Processual and Processual schools of thought; cognitive archaeology in the mid-1990s attempted this very same undertaking. Renfrew and Zubrow (1994) explicitly include cognitive (social) aspects in a systemic view of human behaviour. But, as Whitley and Verhagen point out, 'cognitive archaeology arose at a time in which the two extremes of archaeological theory were at their most divisive, and no real conciliation between them was sought or expected by either side' (Verhagen and Whitley 2011 p. 62).

GIS was tied up in these early attempts at peace-making, and still is (Harris and Lock 2006). It was thought that GIS could be used to produce a quantitative example of the phenomenological results favoured by Post-Processualists. In this way, the quantitative methods of Processual archaeology could serve as an underlining bridge between the two groups. This led to the creation of models of land use from the perspective of agency (Robb and Van Hove 2003, Trifković 2005), view sheds (Llobera 1996, Witcher 1999, Llobera 2000, Llobera 2001, Llobera 2003) and Ingold's creation of taskscapes as a way to view the world (Ingold 1993, Ingold 2000). Some of these experiments have since made their way into traditional predictive models as datasets. Verhagen and Whitley's view of agency can be seen as an evolution from these earlier ideas and methodologies.

The idea of perspective from an individual/agents point of view is also not dissimilar to Lock and Harris's (2006) proposal to humanise the landscapes of predictive modelling. They proposed the use of viewsheds and cost–path analysis to show 'location based on the subtleties of vision and movement as well as historical context' (Harris and Lock 2006 p. 51). While not explicitly calling this 'agency' the results are very similar and based on an individual/agent's perspective of local landscapes.

Meeting Expectations

Their paper demonstrated the reasons why ABM was chosen in the first place by this project— formalisation, understanding and prediction. The use of theory is inseparable from the process. Whitley's paleoeconomic model of the Georgia coast, covering the time period of 2000BC–

AD1650 (Goel et al. 2009, Verhagen and Whitley 2011) worked on human energy budgets, the caloric resources available spatially and cognitive features of human behaviour. These were combined through the application of Optimal Foraging Theory (Emlen 1966, MacArthur and Pianka 1966), Central Place Foraging (Orians and Pearson 1979, Krebs and Stephens 1986) (for examples in archaeology see also Simms (1987), Jones and Madsen (1989), Barlow and Metcalfe (1992), Kelly (1995), Kennett and Winterhalder (2006)). Additionally, theoretical studies such as the Diet-Breadth Modelling (Hames and Vickers 1982, Hawkes and O'Connell 1984, Winterhalder 1987, Smith 1991, Delpech and Grayson 1998) and Prospect Theory (Kahneman and Tversky 1979, Kahneman and Tversky 1992, Machielse, Timmermans et al. 2003) played a role in the creation of this cognitive-paleoeconomical model.

Whitley's model applied accumulated caloric distances onto pseudo-topographic surfaces which could be translated into multiple cognitive presentations, such as the representation of territories, social dominance, perception and pathways. That model integrated theories such as Optimal Foraging Theory, Central Place Foraging, Diet-Breadth, Prospect Theory, etc. all while supporting an understanding of the locations of cultural resources. It moved predictive modelling out of the pure concentration on CRM resources to one that informs both CRM and academic interests.

It will not go without notice that these examples of theories focus on a specific side of the archaeological theory spectrum. It was imagined that simulation would actually be of interest to archaeologists who are interested in more Post-Processual theory like phenomenology. The author once heard the following phase at a conference, 'GIS is a method in search of theory while phenomenology is a theory in search of method'. It was thought at the beginning of this project that computer simulations, not just ABM, could greatly benefit phenomenology by allowing for the exploration of many different experiences with more scientific rigour than has occurred in the past.

In terms of prediction Verhagen and Whitley translated the accumulated caloric distances onto pseudo-topographic surfaces which formed pathways. These pathways would form corridors 'along which resource gatherers would routinely spend a great deal of their time' and then they 'would expect that sites resulting from the loss or discard of artefacts associated with daily activities would occur along these pathways, and a predictive model could be generated to capture them' (Verhagen and Whitley 2011 p. 81).

> 'The (deductive) cognitive modeling framework is extremely flexible, easier to operate and understand, better suited for testing purposes, and as far as we can tell produces better predictions than the currently prevailing alternatives…' (Verhagen and Whitley 2011 p. 90)

Different Routes for the Project

Even though Verhagen and Whitley created an 'agent'-based predictive model that successfully highlighted the benefits of such work, this project was originally unaware of that work and so there were key deviations that put this project on a different path of exploration, and thus it did not simply run their methodology in a different location. The first of these differences was the goals of the work. Verhagen and Whitley were specifically aiming to reconcile academic and CRM divisions with predictive modelling:

> '… many researchers working in archaeological applications of GIS and predictive modeling have struggled to come to terms with the ensuing theoretical debate in archaeology between the Processual and Post-processual schools of thought (see, e.g., Wheatley 1993, 2004; Witcher 1999) [...] As an undesired consequence, the development of predictive modeling has veered away from mainstream archaeological thought and theory and has now become a largely self-contained activity—enjoying reasonable success as a tool for CRM, but not commanding much respect from academic scholars […] We hope that this will point the way out of a debate which we feel has been unduly polarized along the lines of CRM versus academic research, as we are convinced that predictive modeling can be a useful instrument for both fields of application.' (Verhagen and Whitley 2011 p. 50)

This was also a goal of this project, but not the only goal. Originally this project's aims were to improve predictive modelling performance and the idea of a more explanatory modelling helping to achieve that was pursued. The fact that such a framework could also address the CRM and academic divide in predictive modelling was an added bonus. This led to decisions in this project that focused more on trying to improve and measure performance, i.e. choice of project area.

Verhagen and Whitley saw great potential for the use of ABM in predictive modelling to implement their agency theory. They point to Lake's (2000) dynamic Mesolithic foraging model on the island of Islay (Scotland). That model looked at the correlation between lithic scatters and the gathering of hazelnuts. Lake discovered that there was a poor correlation between known lithic scatters and areas where the agents would most likely be harvesting hazelnuts. This threw out the theoretical assumption that hazelnut gathering was the determinant of site location on the island of Islay in the Mesolithic era. These findings were negative, but were a positive example of the potential that agent-based predictive modelling has to add to our understanding of the past. Even the negative results provide insights into the past and help create a better predictive model through elimination of potential factors in site locations. For this reason, Verhagen and Whitley see 'real potential in these techniques for predictive modelling, especially since the cognitive predictive modelling approach does not need to be significantly adapted in order to be used in a dynamical modeling context' (Verhagen and Whitley 2011 p. 87-88).

Yet, they did not use ABM software for their agent/agency-based predictive modelling. A review of the available resources by Verhagen and Whitley found a lack of 'GIS capacity' with ABM programs:

- MAGICAL software, developed by Lake for the above-mentioned Islay Mesolithic project, was for the GIS software GRASS, version 4, and has not been upgraded in the decade since, which effectively eliminates this as a possible program to use.
- SWARM, which is maintained by the SWARM initiative (www.swarm.org), is updated pretty regularly but does not have GIS capacity. In addition, to operate the software it is necessary to possess advanced programming skills, which is usually beyond the capacity of most archaeologists.
- An examination of the freeware agent-based modelling package NetLogo found that it recently acquired the capability to import GIS data, though the examiners found that it has not been used to create a single archaeological model (http://ccl.northwestern.edu/netlogo/models).

In this case 'GIS capacity' means the ability to import data to represent the real world, which can create a major problem if one is trying to create a model of the real world. As a result they concluded that there remains 'lukewarm interest in these modelling techniques from the side of archaeologists despite the optimism expressed in Bentley and Maschner (2003) and Beekman and Baden (2005)' (Verhagen and Whitley 2011 p. 88).

This project had discovered ABM before Verhagen and Whitley published their paper. As discussed in the next chapter this project found that several ABM programs had developed GIS capabilities and thus were suitable for use. The difference in results from this project and their work is not unsurprising given how quickly technology advances, and a publication in 2011 probably meant that the actual work occurred several years before, when ABM was less developed. Because of this Verhagen and Whitley used GIS as the tool to pursue agent/agency-based predictive models while this project used ABM. In a sense this project could pursue this type of modelling with 'better' tools.

GIS vs ABM

This project did not wish to fetishise tools such as ABM or GIS. Using a term like 'better' to describe ABM against GIS needs to be quantified because ABM and GIS are tools and nothing more. GIS is good for some uses but struggles in others. Archaeologists have found that representing complex theoretical factors using GIS is time-consuming and hard to do (Judge and Sebastian 1988, Kvamme 1988b, Gaffney and van Leusen 1995, Joolen 2003, Kvamme 2006, Deeben et al. 2007). Many times in GIS complex cultural phenomena are simply reduced to buffer zones around existing resources to indicate an attraction. A problem of model creation being driven by the tool not the needs (Harris and Lock 2006).

ABM is also strong in some aspects but not in others. Most of the ABM programs have almost no GIS data editing abilities (discussed more in Chapter 4). They can import layers but cannot alter the data, like clipping large datasets . Being able to modify datasets to represent 'real life' conditions as possible through GIS is critical for ABM modelling:

> 'In most cases, dynamical models will be able to give a good idea of the kinds of pattern that will emerge from certain behaviors but not of the exact location or the chronological order in which they will appear. This is due to the fact that these so-called non-linear models are extremely sensitive to initial conditions and the accumulation of small variations (Allen 1997) […] Any dynamical model that would aim for (reliable) prediction would therefore have to be based on "real" initial conditions and have benchmark data available of intermediate and end conditions as well in order to limit the outcome of the simulations to more or less realistic scenarios.' (Verhagen and Whitley 2011 p. 87)

Currently, most ABM programs have very limited or non-existent abilities to edit datasets that create the initial conditions. At some point in the future this distinction may blur as GIS gains more ABM features or vice versa. Until then they are very different tools with very different abilities at the same time complement each other. This project was the first test of ABM to create agency-based predictive models, but it also used GIS where appropriate. Indeed, most of the work undertaken in Chapter 8 was done with GIS because it was the tool that was needed. 'Better' is subject to the needs of the model and what the tool can deliver.

Projects Agent/Simulation Based Predictive Modelling

These differences resulted in a methodology that was somewhat different from Verhagen and Whitley's but still similar. It has as its overarching concepts:

- The final product needed to work in a CRM legal context. While archaeological resources can take many different forms from isolated artefacts to whole landscapes not all of them will be acknowledged as such by the laws that govern CRM archaeology and so only those archaeological resources relevant to CRM in the project area would be focused on. It was understood that such a stance would not be appealing to all archaeologists. However, the aim of this project is to improve predictive modelling performance for CRM uses first and foremost.
- This project would not be a slave to ABM. The best tool, be it GIS, ABM or a mathematical model, was to be used. This method is about simulating behaviours

and this is really a simulation-based predictive model. ABM happened to be a good tool for running the simulations that this project wants to undertake.

- The person's perspective was to be the driver of the models. This was about locations that past people inhabited. Understanding why they undertook the behaviours they did that led to the current archaeological record was key to being able to predict site locations. This is not to say that agents could not be other things and all agents had to be people.
- Multiple theories and scenarios should be explored. The power of simulation is the ability to test out many different assumptions. A model should not just stop because the results somewhat match the expected outcomes. Different theories should be tested to find one that best fits the available data.
- Level of detail will be dependent on what gives the best results. If an abstract model can help pinpoint the location of certain sites then that will be used. As discussed above, I believed that the level of detail in the model should be dependent on the desired outcomes and there was no need to get bogged down in arguments about abstraction, at least to meet the aims of this project.

In more practical terms, it was envisioned that this project would undertake the following steps:

1. A project area would be found that had previous predictive models so that the results of the explanatory modelling could be compared against traditional methods of predictive modelling;
2. The prevailing theories about why sites were patterned in that landscape would then be examined;
3. Each theory would then undergo simulation, using the best tools GIS, ABM, etc., to attempt to understand which theory/theories best represented the available data;
    a. It was understood that this might result in several bespoke and possibly not interchangeable models.
    b. The level of abstraction would be dependent on if it helped pinpoint site locations. Although this project would be looking for sites in a specific project, it was imagined that data to represent local conditions would be needed but that the level of detail in that data might vary.
4. The resulting simulations would then be used as a predictive model to represent areas of archaeological potential in the project area.

Chapters five through eight describe in detail how this process was implemented. This was just the initial plan of investigating the use of ABM/simulations for predictive model creation.

Chapter 4: Finding an Agent Based Modelling Program to Use

As reviewed in Chapter 3, there were issues with past predictive modelling methods that led this project to explore the use of agent based modelling as a possible tool and methodology to create more explanatory predictive models. Verhagen and Whitley (2011) reviewed ABM programs and come to the conclusion that they were not yet viable as a tool for archaeologists. However, as explained, this project had conducted an independent review and came to a very different conclusion about the abilities of ABM programs. This chapter covers that review of the ABM programs as part of the second planned activity for this project— creating a methodology to solve the problem(s) that lead to poor model performance. The purpose of the work undertaken in this chapter was to determine if ABM was a viable tool for predictive modelling use.

This chapter begins with a discussion of the development of ABM programs at the time beginning of this project, circa 2010. It will then review the criteria created for selecting a program and the use of that criteria to narrow down the selection of programs. Those programs that were short listed are discussed in more detail. Finally, the chapter ends with a summation of the program chosen, NetLogo, and the reasons for choosing that program.

State of ABM Programs

At the beginning of this project in 2010 there were close to 80 agent based modelling (ABM) programs that one could use; possibly more (Madey and Nikolai 2009). In terms of knowing the quality of the software there have been several general reviews of ABM software (Madey and Nikolai 2009, Allen 2010) and a review of GIS capabilities of ABM programs (Castle and Crooks 2006). These studies had found that while all programs follow the same concept, modelling through agents, the quality of features and the ease of use vary greatly between programs.

For archaeology modelling, Verhagen and Whitley (2011) reviewed a half dozen programs and concluded that none of them could undertake the work needed to create ‘agency’ based models, at that point in time. However, at the beginning of this project Verhagen and Whitley’s findings had yet to be published and an independent review of the ABM software was undertaken as part of this project. This review found several programs that had the needed features; a striking difference in only a period of a year or two from Verhagen and Whitley’s findings, but not surprising given how fast technology advances.

The Review

To gain a clear understanding of the capabilities of ABM in relation to creating a predictive model this project conducted an investigation into over 70 different ABM programs. The first step in this process was the creation of criteria for making an objective as possible assessment of the ABM programs that were going to be reviewed. What follows are the criteria considered and why:

GIS Capabilities

Verhagen and Whitley (2011) listed GIS capabilities as one of the criteria for an ABM program. They believed that for archaeological use an ABM program must be able to import GIS data or use the ABM software in conjunction with a GIS software program. Because all of the desirable datasets that would be used in this project were in the form of GIS data it was concluded that GIS capabilities would be a non-negotiable criterion. Entering in by hand all of the different data points would not be economically feasible, e.g.

- thousands of site locations in the form of ArchGIS vector data
- thousands of previous archaeological investigations' data, e.g. archaeological surveys, is in the form of ArchGIS vector data
- millions of points in the raster DEM, which is only available in raster form

This project came to that same conclusion about GIS capabilities and so it was considered the primary characteristic for judging programs.

File Formats

Many different data formats can be read and altered by GIS programs. For instance, a TIFF (Tagged Image File Format) file commonly used for photos can be read and used by GIS programs. If an ABM program can import data through TIFF files then it can still use applications available to GIS programs. As such, the definition of GIS compatible is more flexible than being able to import ArcGIS files.

Documentation and Support

The next criterion that was set was a need for strong learning options. Considering I would have to personally learn how to operate the software I wanted good documentation on how to use the software. A program could meet all the other needs but if it is impossible for a user, in this case myself, to understand how to operate the system then it is useless.

The same goes for a support system. If there is no support system to help work through any problems that may arise and cannot be answered in the documentation, then there is a real potential for projects to hit a dead end and fail. Such support could be:

- personal communication avenues with the software developers, e.g. email
- discussion forum
- email list

Training materials and a way to obtain outside help were considered sufficient enough support for this project.

Cross Platform Use

Compatibility with different modern computer operating systems, like Windows version (X), Linux (at least one version), or Mac OS, is another aspect that was considered important for this project. This could be in the form of compatibility with each specific operating system though platform-specific versions of the software. Or an alternative was that the agent based modelling program could run on something like Java which has cross-platform support. Because one of the goals of this project is to widen the use of predictive modelling, cross-platform use was considered a requirement of the system

Open Licence or Free Use

A second component of access is cost. Being able to access software on most computers is not very helpful if one cannot afford to use the software. Thus, to ensure the widest possible use, a requirement that the software be Open Source or free was implemented. Open Source was preferable because if a creator or vendor stopped supporting/updating a software system, future users or researchers could potentially pick up supporting it and continue to update it. However, free use for proprietary software was considered acceptable as well, as long as the licenses ensure protection for free use in the future.

Continual Development

The last requirement for an ABM program was for ongoing support for the software. While Open Source licensing makes it possible to use programs that are no longer supported or updated it was preferred that an ABM program had a strong development backing. There are several reasons for this. One was the development of new features. GIS capabilities are new features for ABM but one which this project could not be run without. While no one can predict what features will be needed in the future a strong development team will ensure they are added and that the ABM program stays relevant.

Moreover, ongoing support means that should an issue have arisen that needed outside support there would have been someone to approach for guidance. This goes back to the help requirement but takes it further. Sometimes problems arise that might require part of the code of the program to be rewritten. This ability may be outside of the abilities of the user and can only be undertaken by those that developed the program. In such a case continual development of the program would be needed.

3D capabilities

3D capabilities were not a requirement for this project but they were noted. A benefit of ABM is that the simulation run on them allows one to observe possible real world interactions. In most cases a 2D bird's-eye view is sufficient but someone may want to view simulations in 3D.

Measuring Criteria

Based on these criteria the data was collected in the following forms:

- GIS Capabilities - this came in the form of Yes/No. So few programs support GIS that asking for specific GIS software was not considered.
- 3D modelling - this also came in the form of Yes/No. Because this was considered a bonus there was no need to collect extensive details.
- User support - any and all support listed was collected.
- License - all licenses and terms of use were examined and noted. Though only those that were Open source or free were recorded as usable.
- Operation systems - all platforms the program could be used on were listed.
- Continual development. The date of last upgrade/release of the software was collected. To determine support a maximum date of last release was set at three years (2008) from when this review was conducted. This was decided upon after looking at a large sample of programs and seeing that if a new version of model is not released within two years most programs cease to be supported.

Programs

A total of 76 agent based modelling programs were examined to see which of the above criteria were met. The list of programs was gathered from varied sources such as Nikolai and Madey's paper (2009), agent-based-modeling.com, Swarmforums, agent based models and GIS blog, etc. This list, while extensive, probably has missed a few ABM software programs created. Still, there is high confidence that the majority of programs have indeed been reviewed.

Each program's website was investigated in full to gather the relevant data. In addition, second party information such as other websites, research papers, and documents were also examined. This was because in several instances outside researchers have created their own add-ons or workarounds for the software which expanded or increased its capabilities. Not all of these additions were always listed on the primary websites. The full table of results can be seen in Table 32.

Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
A-Global	http://agents.felk.cvut.cz/aglobe	2008	Free	Yes	Yes	Yes	Yes	Tutorials, Manual	Yes	No
AnyLogic	http://www.xjtek.com/	2011	Proprietary- very expensive	Yes	Yes	Yes	Yes	Demos, training, consulting, knowledge base, online forum, ask a question, documentation, selected references, book	Yes	Yes
GAMA	http://gama.ifi.refer.org/mediawiki/index.php/GAMA	2010	LGPL	Yes	Yes	Yes	No	Contact authors, report bug, tutorials, guide	Yes	No
MASON	http://cs.gmu.edu/~eclab/projects/mason/	2011	Academic Free License (Open Source)	Yes	Yes	Yes	Yes	Mailing list, documentation, tutorials, third party extensions, reference papers, API	Yes	Yes
NetLogo	http://ccl.northwestern.edu/netlogo/	2011	Free, not Open Source,	Yes	Yes	Yes	Yes	Documentation, FAQs, selected references, tutorials, third party extensions, defect list, mailing lists	Yes	Yes
Repast	http://repast.sourceforge.net/	2011	BSD	Yes	Yes	Yes	Yes	Documentation, mailing list, defect list, reference papers, external tools, tutorials, FAQ, examples	Yes	Yes
SeSAm (Shell for Simulated Agent Systems)	http://www.simsesam.de/	2010	LGPL	Yes	Yes	Yes	Yes	Tutorials, mailing list, FAQs, wiki, author contact	Yes	Plugin available
TerraME	http://www.terrame.org/doku.php	2011	Open Source	Yes	No	No	No	Tutorials, examples, courses, references	Yes	No
AMP (Agent Modelling Platform)	http://www.metascapeabm.com/content/view/57/120/	2011	Open Source	Yes	Yes	Yes	No	Documents, forum, guide, wiki, bug report	coming soon	coming soon
Cormas (Common-pool Resources and Multi-Agent Systems)	http://cormas.cirad.fr/indexeng.htm	2011	Free to modify but not to distribute	Yes	Yes	Yes	No	Training, selected references, examples, online forum, email developers, documentation	Yes	No

Table 5: ABM programs with GIS capabilities

Results

Just looking at the GIS capacities of the different ABM programs cut down the usable programs to ten (Table 5). AMP (Agent Modelling Platform) was in development and while GIS capacities were on the list of coming features, when it was first looked at, this development did not materialise in time for this project, eliminating this program as one to investigate. In addition, AnyLogic was a proprietary system that was quite expensive which eliminated it as an option. A-Global had not been updated in several years which eliminated this as an option. Finally, TerraME was not cross-platform and eliminated. This left seven possible software programs to investigate as an option for the dynamic agency based site predictive model:

NetLogo

NetLogo was first created in 1999 by Dr. Uri Wilensky at the Center for Connected Learning and Computer-Based Modeling, then at Tufts University. The CCL moved to Northwestern University in 2000 and all work has been directed from there since then. NetLogo grew out of StarLogoT, which was authored by Wilensky in 1997. StarLogoT can trace its beginning to the StarLogo developed at the MIT Media Lab in 1989-1990 and which ran on a supercomputer called the Connection Machine. In 1994 a version was developed for the Macintosh computer, MacStarLogo. StarLogoT was essentially an extended version of MacStarLogo with many extra features and capabilities. NetLogo could be considered a fork from StarLogo and both have been developed separately since then.

NetLogo is written mostly in Java with a few elements written in Scala, such as BehaviorSpace and the compiler. The Scala code compiles to Java byte code and works with Java and other JVM languages. The major releases of NetLogo were: 1.0 in 2002; 2.0 in 2003; 3.0 in 2005; 4.0 in 2007 and 4.1 in 2009. Version 4.1.3 was available at the beginning of this project and version 5.0 was released in 2012. It is primarily aimed to be an educational tool for all ages but it has been used in many research settings as well.

At the beginning of this project Netlogo was distributed under its own licensing terms:

> ‘NetLogo software, models and documentation are distributed free of charge for use by the public to explore and construct models.’

All users were granted permission to copy or modify the NetLogo software for educational and research purposes only. These licensing was changed during this project and the software is now licensed under an Open Source license.

The 4.0 version of NetLogo runs on Windows 7, Vista, 2000, and XP for Microsoft operating systems. For Apple computers Mac OS X 10.4 or newer is required. (NetLogo 4.0 was the last version to support 10.3 and 10.2.) It should also work on any platform in which Java 5 or later is installed.

Netlogo does have 3D capacities to represent agents but some older, less powerful systems may not be able to use the 3D view or NetLogo 3D.

Cormas

Cormas is maintained by the Green research unit from CIRAD, a French research centre working with developing countries to tackle international agricultural and development issues. Cormas is based on the VisualWorks programming environment which allows the development of applications in the Smalltalk object oriented language. It has mostly been applied to management of natural resources, namely studying the interaction of human societies with the Earth's eco-system.

Cormas is protected by authors' rights but it can be downloaded for free after filling out a brief form to inform the authors of the reason for use. The purpose behind this is to place the software in the public domain so that others can use it as a 'tool for exchanges and dialogs regarding the problem of natural-resources management'. There is no mention of restrictions on commercial use and it is free to distribute.

Cormas was developed with the non-commercial version of VisualWorks from *Cincom Systems,* which means to run the program one has to first install a compatible version of VisualWorks, which is version 7.6. Cormas is not compatible with version 7.7 or above. VisualWorks works with multiple operating systems, including Windows, Mac OS X, Linux, and several versions of Unix.

At the time of this review in 2011 the last version of Cormas was released in March 2008. A new version has since been released in 2014 but that was too late for this review.

GAMA

GAMA was developed by the research team MSI (located in Hanoi, Vietnam). It is a simulation platform, which aims at providing a complete modelling and simulation development environment for building spatially explicit multi-agent simulations. It was originally created in 2007 and the version available during the reviews was 1.2, released on January 19, 2010. According to the website:

The most important requirements of spatially explicit multi-agent simulations that GAMA fulfils are:

1. The ability to use complex GIS data as an environment for the agents;
2. The ability to handle a vast number of (possibly heterogeneous) agents;
3. The ability to offer a platform for automated controlled experiments (by automatically varying parameters, recording statistics, etc.);
4. The possibility to let non-computer scientists design models and interact with the agents during simulations.

Beyond these features, GAMA also offers:

- a complete XML-based modelling language, GAML, for modelling agents and environments
- a large and extensible library of primitives (agent's movement, communication, mathematical functions, graphical, ...)
- a cross-platform reproducibility of simulations
- a powerful and flexible plotting system
- a user interface based on the Eclipse platform
- a complete set of batch tools, allowing for a systematic or 'intelligent' exploration of models' parameters spaces

The software is under a LGPL license which means anyone has free access to the code and edit it or/and redistribute it under the same terms. GAMA runs on most operating systems but is constrained by the amount of memory available; some models may need up to 2GB of free memory. It works on: Windows 7; Vista; XP; 2000; NT; ME; 98; Mac OS X 10.4 or newer plus 10.3 and 10.2; Linux machines but Ubuntu is recommended; any platform on which a Sun Java Virtual Machine, version 1.5 or later, is installed.

During the reviewed period it appeared that GAMA was undergoing a change in support. No new code updates have been made since 2010 and the main website had a message saying the page will no longer be updated and directs visitors to a Google code project version of the website (http://code.google.com/p/gama-platform/) but that website did not appear to be kept up. Changes have since been made but the review came to the conclusion it had stopped development.

SeSAm

SeSAm (Shell for Simulated Agent Systems) provides a generic environment for modelling and experimenting with agent-based simulation. The key aspects of that SeSAm is claimed to provide are:

- easy visual agent modelling
- flexible environment and situation definition
- the whole power of a programming language
- integrated graphical simulation analysis
- distribution of simulation runs in your LAN
- and many further features....

The first version of SeSAm was available in the winter of 1998 as result of the PhD thesis of Franziska Klügl. In 2000 it was fully re-designed and re-implemented using JAVA, allowing it to be used on most computers.

MASON

As the website humorously says, MASON stands for '**M**ulti-**A**gent **S**imulator **O**f **N**eighborhoods... or **N**etworks... or something...'. MASON is an ongoing project between George

Mason University's Evolutionary Computation Laboratory and the GMU Center for Social Complexity. It was designed by Sean Luke, Gabriel Catalin Balan, Keith Sullivan, and Liviu Panait. They received help from Claudio Cioffi-Revilla, Sean Paus, Keith Sullivan, Daniel Kuebrich, Joey Harrison, and Ankur Desai in the process of creating the program. The website describes it as, 'a fast discrete-event multi-agent simulation library core in Java, designed to be the foundation for large custom-purpose Java simulations, and also to provide more than enough functionality for many lightweight simulation needs.' The highlighted features of MASON are:

- 100% Java
- fast, portable, and fairly small
- models are completely independent from visualisation, which can be added, removed, or changed at any time
- models may be checkpointed and recovered, and dynamically migrated across platforms
- can produce results that are identical across platforms
- models are self-contained and can run inside other Java frameworks and applications
- 2D and 3D visualisation
- can generate PNG snapshots, Quicktime movies, charts and graphs, and output data streams

As Mason runs on Java 1.3, or higher, it can run on almost any computer and with any operating system as long as Java is installed.

MASON has GIS capabilities through the extension GeoMason, an optional extension that adds support for vector and raster geospatial data. GeoMason natively supports reading and writing ESRI shape files. Conversely, there is optional support via third party libraries, such as GeoTools, GDAL, and OGR for other data formats, like PostGIS, Web Feature Format, SDTS, DTED, GeoTIFF, USGS Digital Orthophotoquad, SDTS, GML, TIGER, S57, KML and NTF formats.

MASON is licensed under the Academic Free License ('AFL') v. 3.0. This grants users a worldwide, royalty-free, non-exclusive, sublicensable license, to reproduce the program and modify it in any way the user sees fit. User can also distribute copies of the original program or modified version, with the modified version under any license they see fit.

RePast

RePast is a cross-platform Java-based modeling system that runs on Microsoft Windows, Apple Mac OS X, and Linux. RePast Simphony (RepastS) models can be developed in several different forms including the ReLogo dialect of Logo, point-and-click flowcharts, Groovy, or Java, all of which can be fluidly interleaved. NetLogo models can also be imported. RepastS has been used in many different applications from future hydrogen infrastructures models to ancient pedestrian traffic.

It was developed by a team, led by Michael North of The Center for Complex Adaptive Agent Systems Simulation in the Division of Information Sciences at Argonne National Laboratory. RePast was originally created at the University of Chicago. Since then it has been maintained by organisations such as Argonne National Laboratory. RePast is now managed by the non-profit volunteer RePast Organization for Architecture and Development (ROAD). Developers of the software include Mark Altaweel, Dariusz Blachowicz, Mark Bragen, Carl Burke, Nick Collier, Robbie Davidson, Tom Howe, Charles Macal, Bob Najlis, Michael North, Jonathan Ozik, Miles Parker, Eric Tatara, Jerry R. Vos.

RePast has GIS capabilities both within the system and through the extension Agent Analyst, developed by the Redlands Institute and Argonne National Laboratory. Agent Analyst, unlike most other integrations of GIS and ABM, brings the ABM capabilities into a GIS instead of the other way around. It is specifically designed for the ESRI's ArcGIS suite of products. The extension has its own website, http://www.spatial.redlands.edu/agentanalyst/Default.aspx, which boasts that 'Agent Analyst allows users to create, edit, and run RePast models from within the ArcGIS 9 geoprocessing framework, including access through ArcToolbox, ModelBuilder, and ArcMap. The graphical Agent Analyst tools allow the user to create agents, schedule simulations, establish mappings to ArcGIS layers, and specify the behavior and interactions of the agents.'

RePast is licensed under a 'New BSD' style license which states:

> 'Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:
>
> Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
>
> Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
>
> Neither the name of the Argonne National Laboratory nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.'

Testing of Programs' GIS Capabilities

The closer examination of the programs eliminated Cromas and GAMA, because of the uncertainty surrounding their future support. Thus the remaining programs, SeSAm, MASON, RePast, and Netlogo, underwent testing to see if they could meet the projects needs in terms of GIS use.

The first operation test was the importing of a DEM of the project area (see Chapter 5 for details on the area) into each program. In the case of RePast, which brings the ABM software into a GIS program, this step was reversed with the ABM program being brought into the GIS program. For SeSAm the GIS capabilities of this model were found to be insufficient. With the use of a plugin it was theoretically possible to import vector data. There was a limit to what could be imported as a second plugin was required to produce polygons using SpatialInfo. The lack of documentation for how to operate these plugins led to several unsuccessful attempts to integrate GIS with this ABM program. While the initial investigation found documentation for the ABM program a further investigation found none for the GIS plugins. Taking into account these unsuccessful attempts and the limited range of the GIS capabilities, this program was eliminated as a possible program for this project.

Choosing a Program

This process had narrowed down 70+ choices to three, MASON, RePast and NetLogo. Out of the three programs NetLogo had the best documentation, tutorials, and was, and still is, considered the easiest program to use (Figure 15). Moreover, it is not as underpowered as shown in Figure 15, that image was created by the RePast team and it is not without its biases. Importantly, RePast has ReLogo which converts NetLogo models into RePast models. That meant that models could be created in NetLogo, an easier to use program, and if required more advance computing power, like that of a High Performance Computer (aka super computer), it could be converted to a RePast code and used on RePast S. Using NetLogo meant that the advantages of both NetLogo and RePast could have been utilised for this project if they were so needed. MASON did not have this ability and as such was removed from contention. This project undertook the agent based model creation in NetLogo.

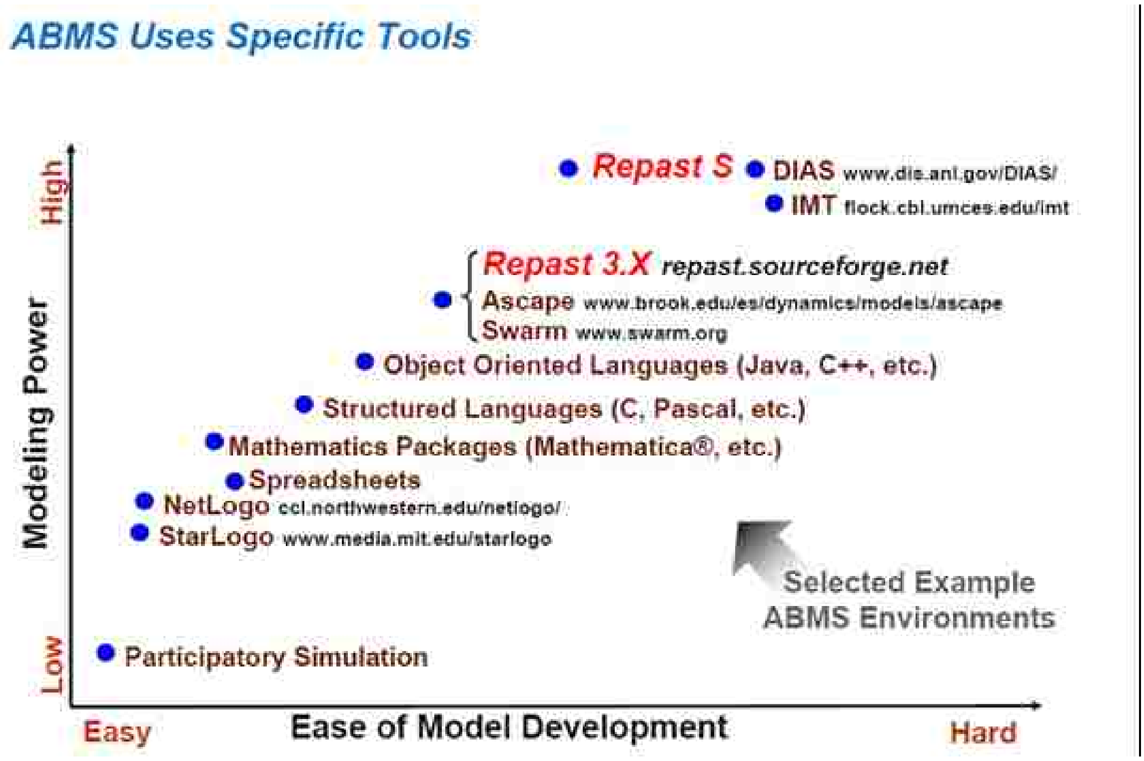


Figure 15: Modelling Power vs Ease of Model Development. Figure from (Macal and North 2006).

Chapter 5: Subject Area and Data

In the beginning this project planned to compare the proposed solution results against independent models to determine its effectiveness, activity four. Basically, compare the old methods against the new to see if there was in fact any improvement in performance. To achieve such an aim meant finding a project area that previously had a predictive model made for it by others. That way my personal biases and skills deficits would not influence the creation of models that used older methods. This chapter presents the project area that was found and some of the other reasons it was chosen, as there are many dozens of areas of with predictive models that could have served as the project area. After that discussion the rest of the chapter gives background information about the project area like cultural history, geology, ecology, etc. This background information helped shape some of the model creation discussed in the proceeding chapters.

Subject Area

The primary subject area chosen to trial the agent based predictive models was located in south eastern New Mexico, USA (Figure 16). It consists of eight USGS 7.5-minute quadrangles (quads) maps, aligned in a matrix of four quads across, east to west, and two deep, north to south. A quad is a 1:24000 scale map that measures 7.5 minutes on all sides. The name of the subject area, Azotea Mesa (named after one of the quadrangle maps), was derived from a previous site predictive model project that used the same area: the PUMP III project (Altschul et al. 2005). The dimensions of the project were also derived from that project which used digitized 1:24000 scale maps as the bases for its data; this project used a different dataset but the same dimensions.

Why this area and why use the dimensions of a previous project? There were several reasons for the selection of this area:

1. The Azotea Mesa project area had been the subject of previous site predictive models for the PUMP III project using traditional methodologies (Boolean, Weighted, and Regression). This allowed for a comparison between several older methodologies and the newer agent based modelling approach. As discussed this was the primary reason for choosing this area.
2. Previous investigations also checked for survey biases in the known archaeological record. By only using data available in each year of 1982, 1992, 1997, 2000, and 2002 they created models for those years. The results showed that after about 10% of the subject area had been surveyed the output of the models created rarely change, with a correlation of .98 between years (Altschul et al. 2005). This eliminates a possible outlier in the results, i.e. data bias.
3. This area has been the subject of this author's Master's thesis on site location accuracy and its effects on predictive modelling. The methods developed during the Master's work will

make it possible to ascertain locational biases in site location data. This eliminates site location errors as a possible problem in the testing of the performance of the diverse methodologies.

4. A further advantage is that this section of the state has some of the most intense survey coverage due to the natural gas and oil developments, as discussed later in this chapter. This means that an excellent database already exists of recorded sites to use as the basis for testing the different models' precision and accuracy.
5. Finally, data for this subject area is free and easy to access. Previous archaeological data is held both in a central state database as well as the regional offices of the Bureau of Land Management which is free to access remotely for researchers through databases. Digital elevation and biodiversity data has been collected by the United States Geological Survey (USGS) which it provides to the public. Conditions that significantly reduce the cost of gathering data for a predictive model and making it financially feasible to undertake this project.
6. Open Data – because most data in the project is free and in public domain, it can be redistributed to other researchers or the general public so that they may check the results or expand upon the research. A caveat to this is that the site location data cannot be shared with the general public without permission of ARMS (Archaeological Records Management Section) of the New Mexico State Historical Preservation Division. However, ARMS does give access to this information to any serious researcher, after filling out a short form.

All of these benefits combine to make this study area the right location for this project. The ability to compare this model against previous independent ones is why this project has opted for the exact PUMP III study area size and location.

Background of Previous Site Predictive Models

The PUMP III project was a component of larger project known as *The Adaptive Management and Planning Models for Cultural Resources in Oil and Gas Fields in New Mexico and Wyoming* (DE-FC26-02NT15445), a partnership between Gnomon, Inc., the U.S. Department of Energy (DOE) and the National Energy Technology Laboratory (NETL). It was funded through DOE's Preferred Upstream Management Practices grant programme and examined cultural resource management practices in two major oil and gas-producing areas: south eastern New Mexico and the Powder River Basin in Wyoming. The project was started in 2002 and completed in 2005 (Altschul et al. 2005).

As part of this investigation into cultural resource management in the specific gas fields of New Mexico and Wyoming, site predictive models were created for three separate zones of the south eastern New Mexico portion of the project. These areas were Otero Mesa, Azotea Mesa, and Loco Hills, each named after one of the quadrangles (quad) in their respected study areas. The quads are used as a unit of measurement for the project, because at the time there was limited digital data available. Some of the datasets had to be digitised from existing quad maps, thus the boundaries were defined by the quad maps. The Azotea Mesa area had three predictive models created using each one of the different methods – Boolean, Weighted, and Logistic Regression.

The other PUMP III areas, Otero Mesa and Loco Hills, were not used for this project. Loco Hills did not have three different predictive models created, only two – Weighted and Logistic Regression, for it so it was less valuable in terms of testing past methods. Another reason was the poor performance of the previous predictive models in those different areas. Results of the different PUMP III models, when compared against the previously recorded archaeological record of the areas, can be seen in Table 6. The criteria, used by the PUMP III project, for evaluating success in these models, was Kvamme's gain value.

A perfect gain score would be 1 but as was noted in a previous study (Rocks-Macqueen 2010) and by other modellers the Gain Statistic does not take into account the area a site must occupy (Kamermans 2006, Kvamme 2006, Verhagen 2007c). For example, in the area for this project known sites occupy 3% of the currently surveyed area. To accurately capture all sites at least 3% of the area has to be labelled as containing sites. In that case the maximum gain value could only be .97 instead of 1. Thus expectations needed to be adjusted on what the maximum possible gain value will be, which was calculated for the project area.

The relative poor performance of the traditional site predictive models in the Azotea Mesa area leaves the usefulness of site predictive models in this area in question and the opportunity for an agent based model approach to prove itself. For the Otero Mesa area this was also an opportunity but given its high performance any improvement would have been smaller and thus the possibility of other factors causing slight changes in outcomes that could be misinterpreted. Moreover, only Azotea Mesa had its site locations fully investigated to confirm data accuracy through the Master's work conducted in this area (Rocks-Macqueen 2010). There were not the resources available to conduct spot checks in Otero Mesa and Loco Hills which could have potentially biased the outcomes. Because of these reasons only the Azotea Mesa area was used in this project.

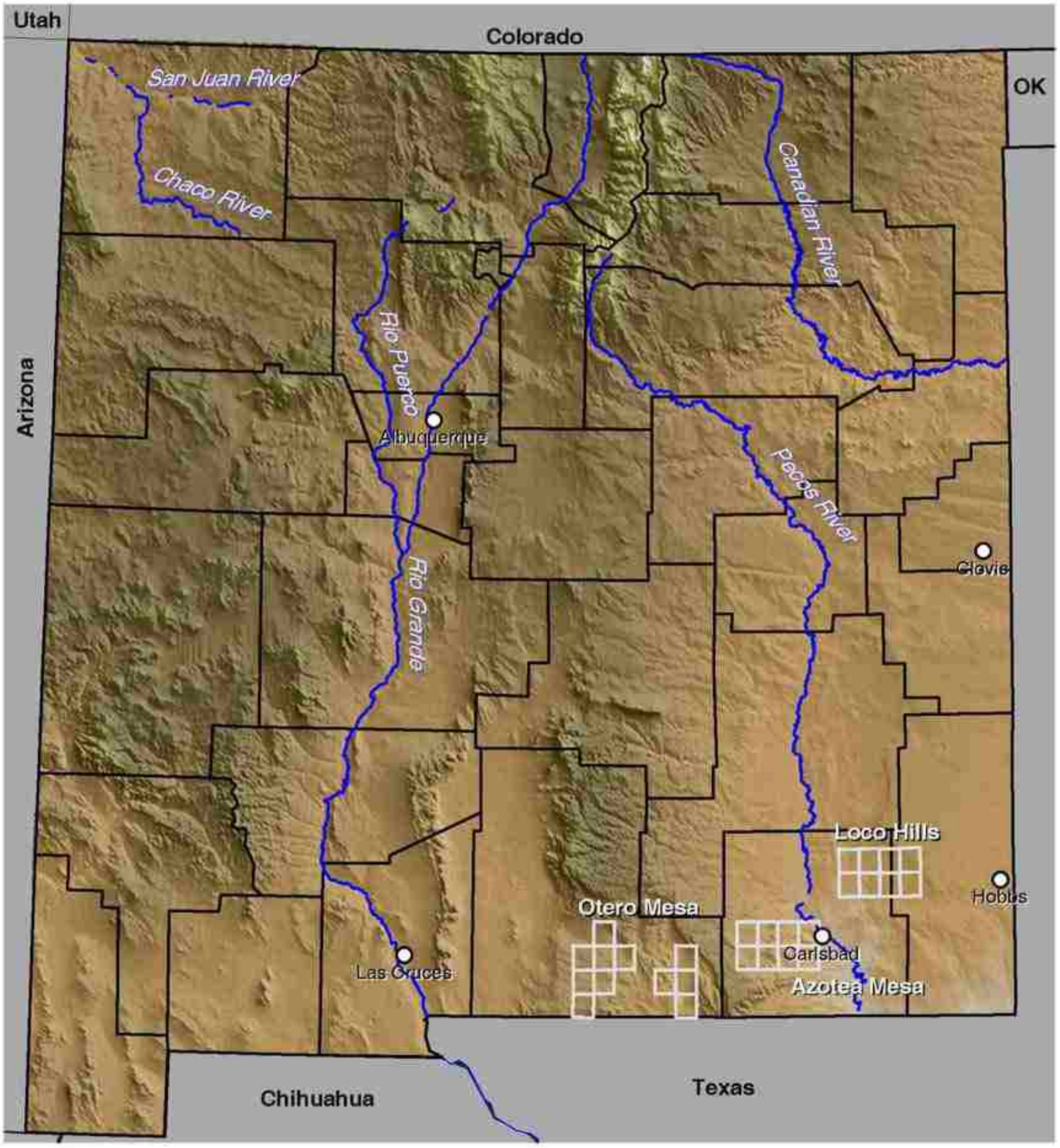


Figure 16: Study areas of the PUMP III project (Altschul et al. 2005 p.2 Figure 1.2).

Study Area	Max Gain Values	Boolean	Weighted	Logistic Regression
Loco Hills	0.960	-------------	0.46 41% 76%	0.25 71% 95%
Azoto Mesa	0.970	0.21 46% 58%	0.39 43% 70%	0.26 63% 85%
Otero Mesa	0.995	0.62 (East) 31% 81%	0.71 (West) 25% 87%	0.62 (West) 34% 90%

Table 6**:** Max gain values and data they are based on from the project area. (**Black**) Gain Values (**Blue**) Land cells covered in medium to high probability zones (**Red**) Site cells captured. Percentage of land that constitutes sites on surveyed land are used to estimate the max land coverage of sites: Loco Hills 4%; Azoto Mesa 3%; Otero Hills .5%.

Potential Issues

After their creation, the use of the PUMP III predictive models was sporadic at best, with only anecdotal evidence that they were ever used for cultural resource management. During visits in 2009 and 2012 the local and state BLM office were unable to locate any copies of the models. By contacting the authors of the report it was possible to obtain some datasets however their formats were unusable in the GIS software available, ArcGIS or QGIS. There were concerns that this would be a detriment to the project as it meant relying on the results presented in the report and not being able to check them in GIS.

However, it was decided to go forward with this project area, despite this lack of access to the direct data, because:

1. Getting access predictive model data is very difficult. As discussed in Chapter 2 most models are created and then placed in drawers. If a model is not constantly used, then the datasets are not preserved. Attempts were made to find other suitable locations but most modelling results are only available through publications and not datasets. This would have been an issue for most locations.
2. The other factors for choosing this project area were still valid. The site locations have been previously tested for accuracy, the area had been tested for survey biases, multiple methods had been independently created, etc. There were many reasons to choose this project area, not just the previous predictive models created there.

Measuring Performance

Because the PUMP III models used Kvamme's Gain Statistic and because that is most used standard of measuring performance (see discussion in Chapters 1 & 2) it was decided that this would be the measurement of performance used for the agent based predictive models. As will be discussed in the conclusion of this project it is probably not the best measurement of model performance. However, that conclusion was only reached after undertaking the project and so this project started with the expectation that a Gain Statistic would provide easily measurable outcomes to compare with the models.

Regional Geology and Geography

The Azotea Mesa study area is located west and south west of the city of Carlsbad. Indeed, some of the edges of the study area are part of the Carlsbad city boundaries which also includes the Pecos River. The river runs through the north east corner of the study area and the edge of its river valley runs north-south through the eastern half of the study area. The western edge of the boundary captures the lower slopes of the Guadalupe Mountains. This means that the elevation of the project area ranges from roughly 1,700 metres (5,600 feet) in the south west to 950 metres (3,100 feet) in the north east corner (rounded to the nearest 50 metres) (METI and NASA 2012).

These larger geographic entities, the Guadalupe Mountains and the Pecos River, are significant components of the regional landscape. The Guadalupe Mountains continue south from the study area into Texas. The highest elevation in this range is Guadalupe Peak in Texas, at 2,667 metres (8,751 feet) (National Geodetic Survey 2012). Runoff on the east side of the mountains drains into the Pecos River watershed and includes the drainages running through Azotea Mesa. The runoff from the west side of the mountains drains into the Salt Basin (Altschul et al. 2005). This mountain range is both important in terms of ecological resources provided but also in terms of raw materials. There are several lithic procurement sites in the Guadalupe Mountains that would have provided raw materials for many people in the region (Hogan 2006). Furthermore, a review of the archaeological record shows that several lithic materials quarries are also located within the project area.

The Pecos River Valley flows for roughly 926 miles, from its headwaters in the Sangre de Cristo Mountains in north-central New Mexico, all the way into Texas where it joins the Rio Grande river (Kammerer 2005, Office of the State Engineer 2005). This river has been the primary perennial water source for most of human occupation in the area (Altschul et al. 2005). It also butts up against the Mescalero Plain which is defined as the pediment surface sloping westward from the base of the Mescalero Ridge to the Pecos River. It is estimated that about 80% of the Mescalero Plain is covered by what are called the 'Mescalero' Dunes and the rest of this flat region is spattered with dry and seasonal streams, sinks, and resistant rock outcrops (Reeves 1972).

Further east of that is the Llano Estacado plateau, which covers some 32,000 square miles in east New Mexico and west Texas (Reeves 1972). The plateau consists of large sections of caliche, a layer of soil that has been hardened by minerals, and has low rainfall which makes this plateau semi-arid. In some areas the caliche is buried below sandy and clay deposits. The broad uplands between are dotted with thousands of shallow depressions, many containing playas (old dry lake beds) with lunate dunes on their leeward margins (Hawley 1986), all inhospitable geology for many plants and animals to live in, let alone humans. These conditions in the Llano Estacado and Mescalero Plain make the Pecos River Valley one of the few suitable locations for agriculture in the surrounding region.

Current Climate

Azotea Mesa's climate is semi-arid, with very hot summers and mild winters. During the summer the average high temperature is 35-38 Celsius (mid-90s to over 100 degrees Fahrenheit). It is not uncommon for temperatures to reach up and over 46 C (115 F). The winters sees daytime temperatures of around range -3 to 13 C (high 20s-50s F) though outliers of -10 to 24 C (teens to 70s

F) are not uncommon, even in the same day (WRCC 2012). The high altitude and lack of cloud cover account for these wild swings in temperatures. For the most part, cold temperatures are not too common, with frost free days averaging more than 200 days a year (Altschul et al. 2005). The annual rain fall varies from 10 to 16.5 inches per year (25.4-41.9 cm) (WRCC 2012) but this varies substantially with elevation (NMAES 1971). More information on rainfall and weather is discussed in Chapter 8.

Paleoenvironment

There are paleoenvironmental studies from the Southern Plains (Reeves 1972, Stafford 1981, Hall 1982, Johnson and Holliday 1989, Johnson and Holliday 1995) and the northern Chihuahuan Desert, but the only reconstructions in the study area are those of the Guadalupe Mountains (Van Devender, Spaulding et al. 1979, Van Devender 1980, Roney 1985). Preference would have been to have a paleoenvironmental reconstruction encompassing the study area but given no alternatives these studies can serve as acceptable proxies.

Humans are assumed to have arrived in south eastern New Mexico during the last Late Pleistocene/Early Holocene pluvial, roughly 11,000 to 4,000 BC. At the beginning of this period there was increased precipitation and lower summer temperatures. This created an environment that had small and large playas (ponds/lakes) surrounded by mixed grassland/open woodland vegetation. This all changed around 9000 BC when the climate became much drier, with warmer summers. Shifts also included possibly cooler winters and a concentration of precipitation during the winter months. The result of which was a shift to grasslands. This period also saw the extinction of megafauna, like mammoths and *Bison antiques* (Altschul et al. 2005).

From around 8500 BC to 7500 BC this area saw some fluctuations between wetter and drier periods. Yet, the general pattern was towards increased dryness in the area. By 7000 BC all of the woodland vegetation had disappeared from the Southern Plains and the majority of playas had dried up. The landscape turned into a desert grassland very similar to what it is today. A brief period of higher moisture during the San Jon pluvial 6500-5000 BC intersects this period of dryness but for the most part the climate has stayed the same. This is assuming that the subject area followed a similar pattern to the surrounding landscapes which have paleoenvironmental studies.

What this means in terms of human behaviour is that for the majority of possible human occupation there is no significant change in resource availability. Changes in technology can affect how well the local populations could exploit these resources but overall there was a stable environment. However, brief periods of extreme weather such as droughts do need to be taken into account. In terms of modelling this will mean fewer variables to account for. Though more environment variability is not necessarily a negative aspect for predictive modelling, it just requires more work.

Vegetation

The vegetation data gathered for the PUMP III project was from the Gap Analysis Program (GAP) of the USGS. This data is divided into 17 subcategories (Figure 17) and the majority of the vegetation in the study area was listed as Chihuahuan desert grassland (Table 7).

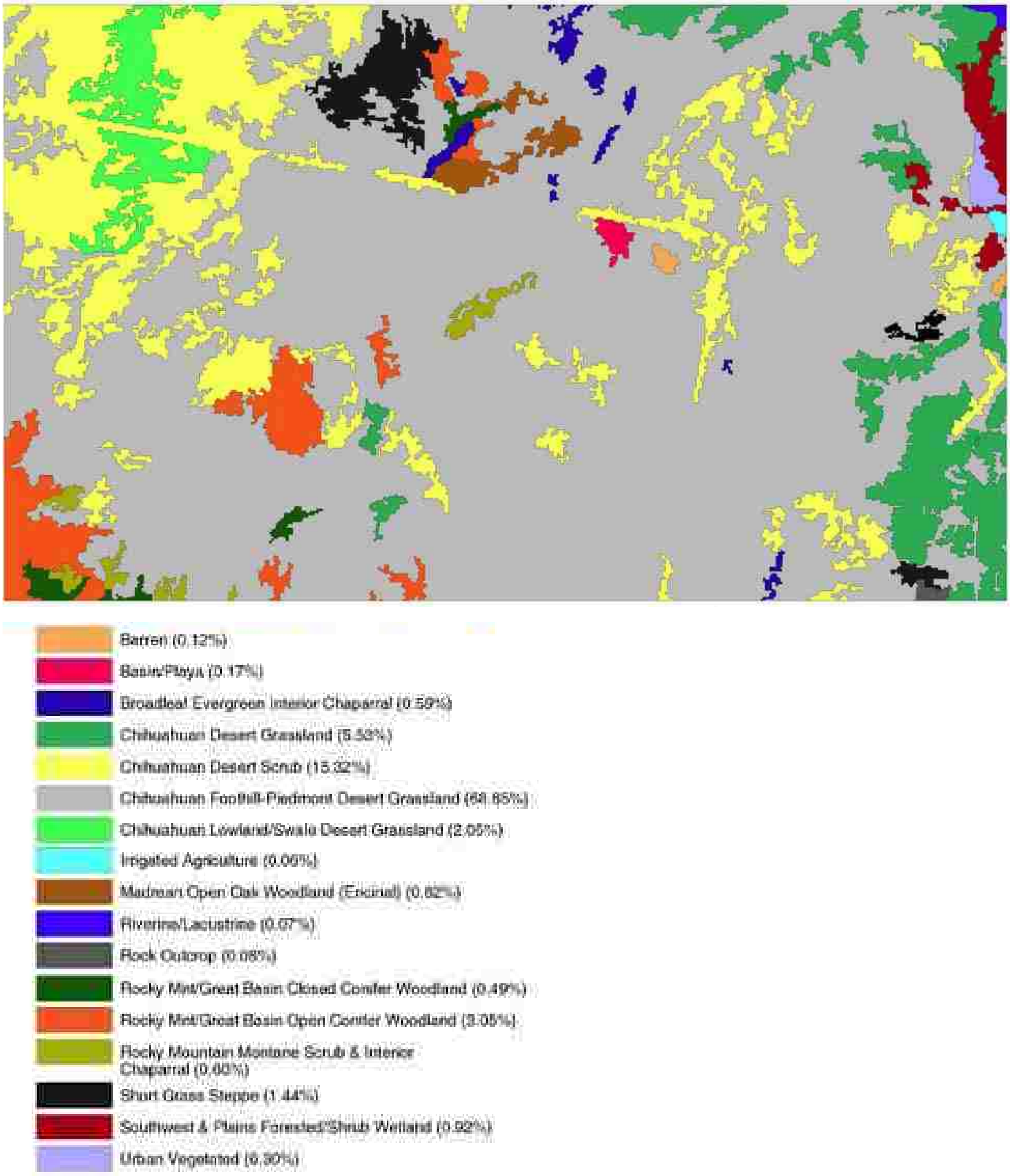


Figure 17**:** GAP vegetation data from the PUMP III predictive model project (Altschul et al. 2005p. 82 Figure 6.5).

PUMP III	Percent of Land Coverage	GAP Data for this project	Percent of Land Coverage
Chihuahuan Foothill-Piedmont Desert Grassland	68.65%	Chihuahuan Creosotebush, Mixed Desert and Thorn Scrub	46.70%
Chihuahuan Desert Scrub	15.32%	Apacherian-Chihuahuan Semi-Desert Grassland and Steppe	32.98%
Chihuahuan Desert Grassland	5.53%	Western Great Plains Shortgrass Prairie	10.48%

Table 7**:** The largest vegetation zones in the study area from PUMP III and GAP (Altschul et al. 2005, USGS 2012b).

As discussed at the beginning of this chapter, the original dataset was no longer available for use in this project. This meant that new data from the GAP program was gathered (USGS 2012b).

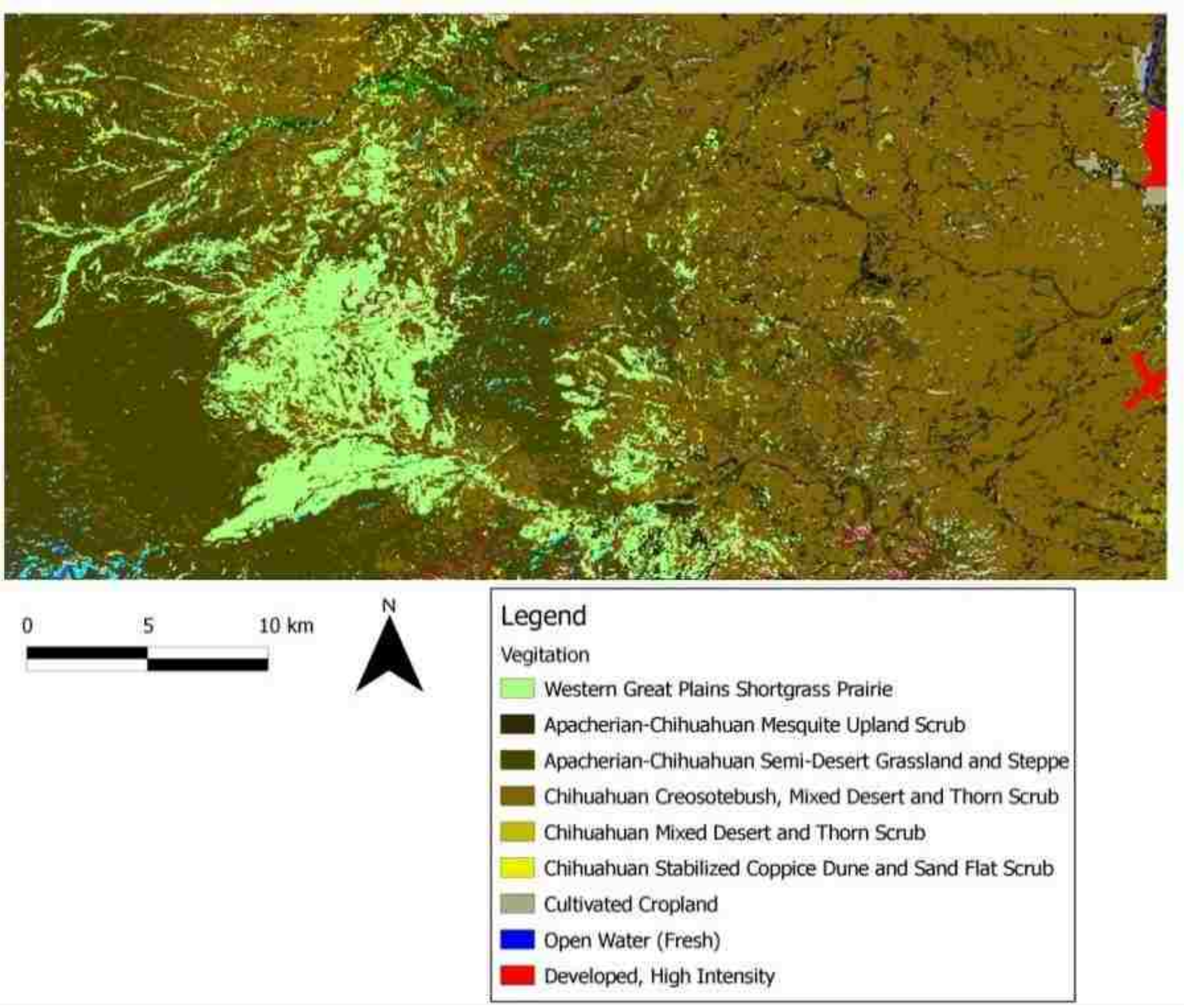


Figure 18: Gap data distribution of vegetation. Only those classes making up more than 1% of the project area are listed.

In the newer GAP data the largest ecological zone is the Chihuahuan Creosotebush, Mixed Desert and Thorn Scrub covering roughly 47% of the study area. This land cover type is composed of two ecological systems, the Chihuahuan Creosotebush Xeric Basin Desert Scrub (CES302.731) and the Chihuahuan Mixed Desert and Thorn Scrub (CES302.734). This cover type includes xeric creosotebush basins and plains and the mixed desert scrub in the foothill transition zone above, sometimes extending up to the lower montane woodlands.

The next largest ecology system is the Apacherian-Chihuahuan Semi-Desert Grassland and Steppe system, which is broadly defined desert grassland, mixed shrub-succulent or xeromorphic tree savannah. It covers roughly 33% of the subject area. It is found on gently sloping bajadas, mesas and steeper piedmont/foothill slopes in the Chihuahuan Desert. It is characterised by typically diverse perennial grasses

The last major ecological zone in the subject area is the Western Great Plains Shortgrass Prairie, which covers about 10.5% of the subject area. This zone occurs primarily on flat to rolling uplands. Higher grasses may be found in this ecological zone but they are of secondary importance to the sod-forming short grasses. In areas with sandy soils there can be a higher cover of Hesperostipa comata, Sporobolus cryptandrus, and Yucca elata. Moreover, scattered shrub and dwarf species can be found in this zone. A detailed breakdown of the vegetation systems is found in Appendix D.

The new GAP data was significantly different than that listed in the PUMP III project (Table 7) with more detailed categories. This is due to higher resolution data collection by the USGS which changed the accuracy of the GAP data. The PUMP III project noted this low resolution data:

> 'The gross scale at which the vegetation is mapped and the general nature of the vegetation categories do not allow us to observe or model the effects of relatively small patches of highly valued resources such as succulents or seed grasses on the location of past human activities. At best, we can only evaluate general land-use patterns related to vegetation categories.' (Altschul et al. 2005 p. 82)

As the newer data had higher accuracy it was tested in some of the model creation but as will be demonstrated it contributed very little to the models. Because, even the newer more detailed GAP data did not play a significant part in the model creation no attempts were made to try reconcile the two different GAP datasets. This information is mainly presented to give an understanding of the ecology of the project area for background purposes.

Project Area Geology

For this project, soil data was obtained from the Natural Resources Conservation Service (NRCS) Soil Survey, a division of the United States Department of Agriculture. These data were acquired from the web soil survey website (http://websoilsurvey.nrcs.usda.gov/app/HomePage.htm) and Chapter 7 discusses how this data was collected. The results of which present a similar, but not exactly the same, picture of soil types to the PUMP III data (Table 8).

NRCS Code	Type	Brief Description	Percentage of project area covered
EC	Ector stony loam, 0 to 9 percent slopes	Very shallow to shallow, well-drained, calcareous, stony and extremely rocky soils that are underlain by limestone.	24.60%
EE	Ector extremely rocky loam, 9 to 25 percent slopes		23.55%
ER	Ector-Reagan association, 0 to 9 percent slopes		8.45%
LT	Limestone rock land	Steep to very steep canyon walls and escarpments.	8.97%
RE	Reagan-Upton association, 0 to 9 percent slopes	Deep, well-drained, moderately dark coloured, calcareous loams that developed in old alluvium derived from calcareous, sedimentary rocks of the uplands.	9.34%
UG	Upton gravelly loam, 0 to 9 percent slopes	Calcareous, gravelly soils that developed in old alluvium derived from calcareous sedimentary rocks.	9.87%

Table 8: Major (>5% project area coverage) soil types from NRCS dataset (NMAES 1971, USDA 1981).

A more detailed breakdown of the soil attributes is found in Appendix D and a discussion on the water absorption abilities of the different soils in Chapter 7. The NRCS data shows that the majority of the project area is characterised as eroding bedrock with thin soils (Table 9). There are several percentage points differences between soil types in the PUMP III and NRCS datasets and their distributions are different as well (Figure 19 & Figure 20). Yet, the results are similar enough to say that the general geology of the subject area is one of eroding bedrock with thin soils interspersed with alluvium deposits. Several other soils types making up minor deposits throughout the project area to round out the geomorphology.

Soil Type	Distance to restrictive layer	% of coverage	Soil Type	Distance to restrictive layer	% of coverage
EC, EE, ER, LT, RPG, RTE, Up	<=15	66.43%	Aa, AH, Ah, Ao DP, GA, GC, Ha, Hk, Ku, LN, MXC, PD, Pe, PM, RA, Rc, Rdm, RE, RG, RM	>200	19.67%
At, DRG, DYE, SG, SM, TN, TPE, UG, Uo, UR, Ut	>15 and <84	13.83%	GA, W	N/A	0.07%

Table 9: Distance (cm) to a restrictive layer, e.g. bedrock, cemented layers, dense layers, and frozen layers (USDA 1971, USDA 1981).

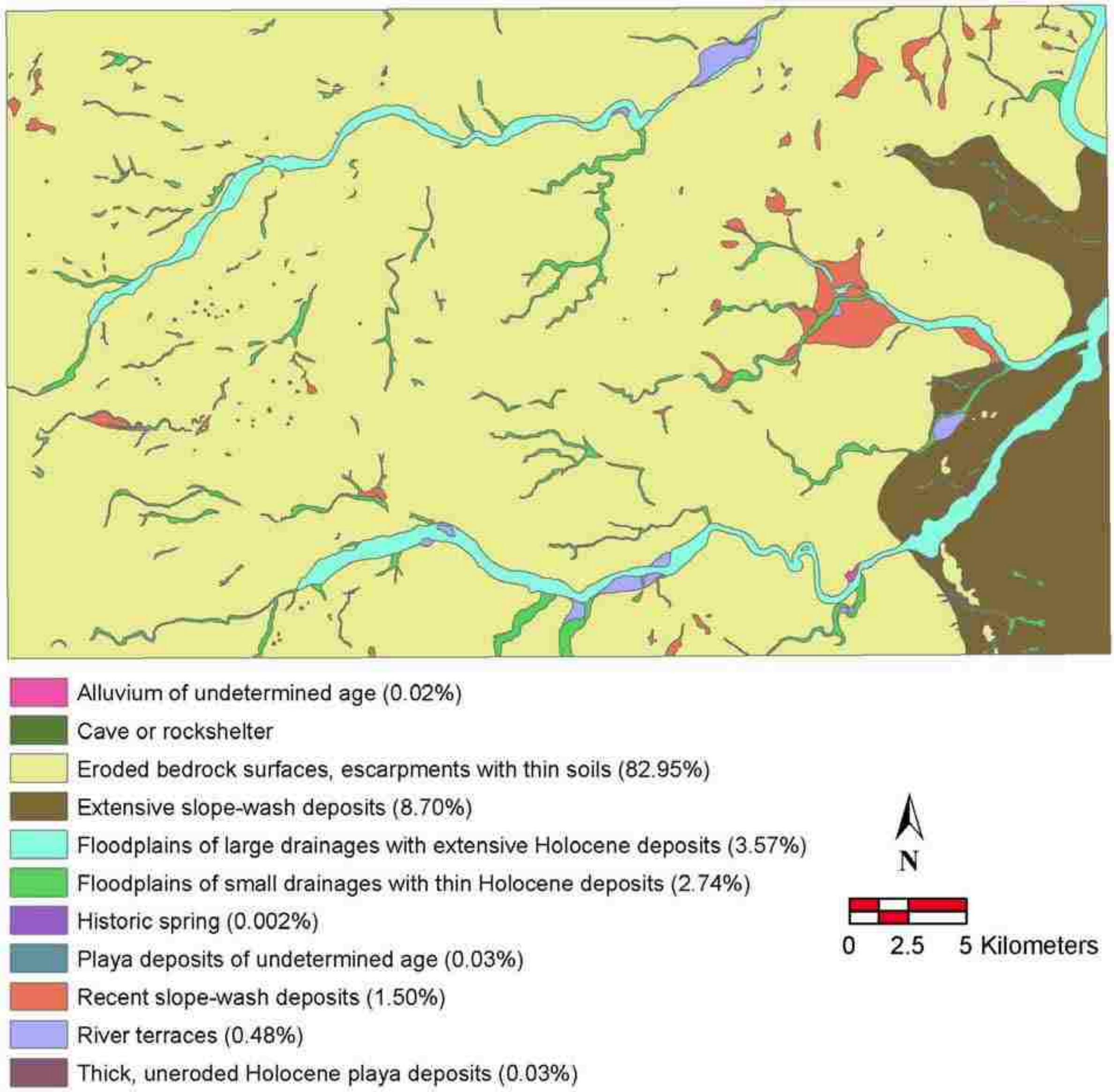


Figure 19: Geomorphology of Aztea Mesa study area (Altschul et al. 2005 p. 80 Figure 6.4).

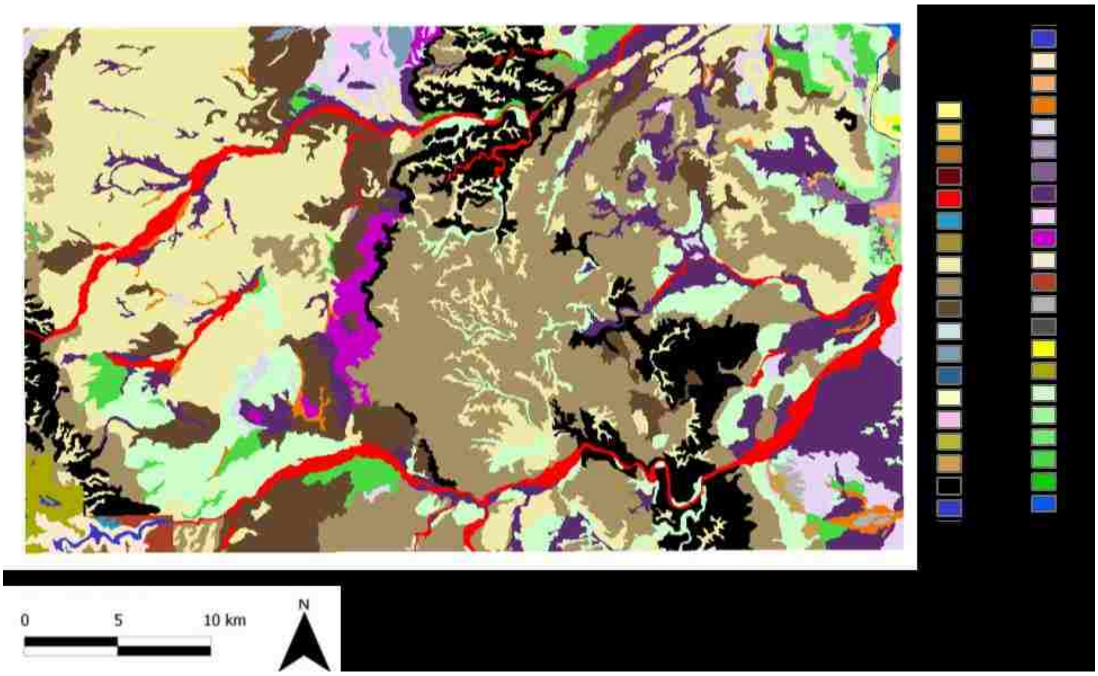


Figure 20: Study area geomorphology from NRCS dataset.

The differences between data sets are a result of how the data was collected. The PUMP III project procured geomorphologic data for the project area through the mapping of black and white stereo aerial photographs (scale about 1:52,000) and colour infrared stereo aerial photographs (scale about 1:86,000). This process involved:

> 'Landforms were identified from the stereo aerial photographs using a Topcon mirror binocular stereoscope at 3× magnification, and the location and spatial distribution of the landforms were then plotted on 7.5-minute topographic maps (scale 1:24,000), the base-map standard for this project. Landforms smaller than about 200 feet in greatest dimension (ca. one-tenth of an inch on topographic maps and smaller yet on the aerial photos) were not mapped.' (Altschul et al. 2005 p. 80-81)

This is similar to the Natural Resources Conservation Service (NRCS) Soil Survey data collection methods which also used aerial photograph to map soil distribution. However, the NRCS also set teams to investigate the soil profiles, take samples and test those samples in labs giving it a more detailed understanding of the soil profiles. It also meant additional characteristics such as soil depth or soil absorption rates were also collected. Given that these characteristics were needed for some of the model creation this data was used instead of attempting to recreate the PUMP III data from the report's images.

Elevation Data

The PUMP III project used a DEM dataset provided by the United States Geological Survey (USGS) to represent the topography of the project area. However, the specifics of this data were not discussed in the report with only the mention that the data was laid out on 30 x 30 metre squares.

The current project relied on the newest version of data provided by the USGS for the project area, Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) Version 2 (ASTER GDEM2), as a base for modelling the topography of the area. The data is distributed in 1 x 1 degree tiles and uses the WGS84 1984 World Geodetic System. This covers 3601 x 3601 pixels that are each 1 arc-second. 1 arc-second is equal to 31 meters and so the pixels are 31 x 31 meters. This is close to the 30 x 30 m squares of the PUMP III data. However, it is unknown if the PUMP III data was originally in a 30 x 30 format or if this data has undergone any sort of interpolation. Thus it is assumed that these datasets are similar but not the same.

The ASTER GDEM2 data was collected by using nadir- and aft- looking near infrared cameras by the National Aeronautic and Space Administration (NASA) and the Ministry of Economy, Trade, and Industry (METI) of Japan. In October 2011, NASA and METI released GDEM 2. It has an overall accuracy of around 17-m at the 95% confidence level, and a horizontal resolution of approximately 75-m. It is given out for free by METI and NASA with condition the following message is included in a publication: 'ASTER GDEM is a product of METI and NASA' (METI and NASA 2012).

Culture History

There have been several detailed major syntheses for this general region of the state; see- (Larralde and Sebastian 1989, Rothman 1998, Duran, Eidenbach et al. 2001, Katz and Katz 2001). The following is a brief general overview of the cultural history of the subject area. The chronology of the area can be grouped into four broad temporal periods: Paleoindian (first humans – 5,000 BC), Archaic (5,000 BC – AD600/900, Ceramic (AD600/900 – AD1500s) and Protohistoric (Spanish Arrival in AD 1500s to 1860s/1880s with the arrival of US settlers) (Katz and Katz 2001).

Paleoindian

The first human inhabitants in this area were the Paleoindians. There are very few Paleoindian sites in the region and most are identified and classified by the projectile points they have left behind, e.g. Clovis (ca.9000 BC) and Folsom (ca. 8500 years BC). Both the Clovis and Folsom period peoples used the fluted spear points, for which they are named, to hunt the, now extinct, megafauna such as mammoths and *Bison antiquus*. The butcher and kill sites from these animals are the most commonly found archaeological sites from this period (Rothman 1998). However, a lack of archaeological data makes it hard to determine if this was their primary subsistence or not. Sebastian and Larralde (1989) contend that Clovis, in particular, was probably a broadly-based hunter-gatherer adaptation instead of specialised big game hunters. Recent evidence from the Gault Site in central Texas tends to support this assertion (Collins 2002). So far research does not indicate any sites older than Clovis (ca. 9000 years BC) in south eastern New Mexico (Hogan 2006).

Concern has been raised about dating and clarification of sites through projectile points (Cordell 1979, Larralde and Sebastian 1989). However, since then the number of radiocarbon dates associated with paleopoints has tripled (Eighmy and LaBelle 1996) and now allows for a more accurate chronology and identification of Paleoindian site through projectile points (Hogan 2006). However, because these sites are mainly identified through their distinctive projectile points, selective collection of projectile points, both prehistorically by native groups for reuse and in recent times, by amateurs for collection purpose creates a problem of biased sample data. This has probably limited our ability to identify these sites as Paleoindian instead of the generic label of lithic scatter. As a consequence, in the project area there is only one site that has had an identified Paleoindian component to it.

As mentioned in the Paleo-climate section above, there were changes in the climate 8500-6500 BC years ago and this is reflected in the cultural record of the Later Paleoindian sites (8500–6500 years BC). The mass extinction of megafauna caused the local populations to concentrate on smaller prey, such as the modern bison (*Bison bison*) (Larralde and Sebastian 1989, Rothman 1998, Judge 2007). It seems that the later Paleoindian toolkits (those associated with Scottsbluff and Eden projectile point forms) were employed, which indicates a more generalised subsistence strategy and a settlement pattern focused on playas (temporary lakes) and springs (Larralde and Sebastian 1989, Rothman 1998, Altschul et al. 2005).

Archaic

This generalised subsistence strategy would continue into what is defined as the Archaic Period (7,000 years ago). It also coincides with the last major shift in the climate to create the environment seen today. The majority of what is known about early archaic adaptation to this region comes from excavations in dry caves of the Guadalupe Mountains and of open-air sites in the Pecos Valley and southern Tularosa Basin (Altschul et al. 2005). These excavations have yielded a wide range of artefacts including organic and trade goods like woven articles of high quality such as yucca mats, and bracelets made from *glycimeris* shell, originating in the Gulf of California (Rothman 1998). As will be discussed in the archaeological data section, there have been very few excavations in the actual study area. Much of the information about the cultural history is reliant on information gathered from the surrounding areas.

It is not known how the Paleo people relate to those in the Archaic. It had been argued that Paleoindian hunters withdrew from the south west as climatic conditions deteriorated and were replaced by Archaic populations moving into the region from the west (Irwin-Williams 1979). The other view is that the existing Paleoindians became Archaic hunter-gatherers as they adapted to the environmental changes during the early Holocene.

Ceramic

Domesticated corn arrived in the area, from Mexico, in the mid to late Archaic period, before 1000 BC, the earliest known dates being 1500 BC (Carmichael 1982). It played a minor role in diets until the first millennium (Larralde and Sebastian 1989), with the earliest known evidence of cultivation from around AD 1-50 (Wiseman 1996). The beginning of cultivation was due to the introduction of more productive type of corn and other domesticated plants such as beans and possibly amaranth. The domestication of plants plus the introduction of ceramics and the bow and arrow is viewed as the end of the Archaic and beginning of the Ceramic period, roughly between AD 600 and 900. This led to the use of corn agriculture and more sedentary village or farmstead settlements on alluvial fans at the edges of the desert basins and along the tributaries of the Pecos River, all firmly established by AD 900 (Katz and Katz 1985).

What is not known is whether the Archaic form of subsistence and living continued in other areas of the region, outside of the Pecos river valley. No evidence of agriculture has been found in the Guadalupe Mountains. It has been suggested that a mobile hunting and gathering lifestyle persisted during the Ceramic period (Applegarth 1976), with hunting focused on small game and the collection of plants. This hypothesis is supported by the radiocarbon dates on ring middens, which are predominantly dated between AD 500 and 1450 (Katz and Katz 1985). It is possible that this sort of living continued with the incorporation of ceramics and bows and arrows. It is also possible that different groups of farmers and hunter-gatherers both inhabited the region (Larralde and Sebastian 1989).

More permanent settlement that comes with agriculture gives way to substantial pithouse sites by the AD 1000s and 1100s. This even leads to modest pueblo sites by the AD 1200s and what is called the Formative period. Though unlike other areas of New Mexico, the Formative cultures in this region seem to concentrate on agave and shin oak for subsistence instead of corn, beans, and squash (Leslie 1979); a continuation of the trends seen in the preceding Archaic periods. This period of farming is quickly followed by a shift from sedentary farming to nomadic bison hunting between about AD 1250 and 1350, (Jelinek 1967).

Protohistoric

By more or less AD 1400 south eastern New Mexico was largely abandoned by agricultural-based peoples. This is attributed to localised droughts throughout the 1300s and the beginning of the Little Ice Age (Rothman 1998). When European contact brings a written record to the area it is inhabited by the Kiowa and Mescalero Apache, nomadic hunting groups who lived in short-terms camps and kill/butchering sites on the Plains. These groups pushed out or absorbed the preceding local populations such as the Jumano, who were assumed to be descendants of the local populations stretching back to the Archaic (Rothman 1998). Data suggests that these groups were sedentary or at least foraged within a much more restricted range during a part of the year, and were only seasonally mobile until the introduction of horses which greatly increased their range (Hogan 2006).

Early Spanish expeditions into the south west did not visit the project area (Weber 1994, Rothman 1998) until the Espejos expedition, which travelled along the Pecos River in 1583 (Kelly 1937, Weber 1994). An expedition in 1590 led by Gaspar Castaño de Sosa also passed through the Pecos River area and found a found a cache of shelled corn in an *olla*, near the location of present-day Carlsbad (Rothman 1998). However, they found very little signs of life and after that Spanish explorers ignored this area of the state. This meant that in 1595, when King Philip II of Spain commissioned Juan de Oñate to conquer and settle 'New Spain' this conquest took place mainly in the north of New Mexico near Santa Fe and along the Rio Grande (Faunce 2000). Consequently the establishment of the Camino Real, or Royal Highway, to connect the settlements of the northern frontier, was well to the west of the Guadalupe mountains (Schneider-Hector 1993).

The Spanish largely ignored this area in the following centuries. The only real interaction was the launch of punitive expeditions against the Apaches to keep them at bay and prevent them from raiding in the El Paseo Area. In 1786, Bernardo de Gálvez, the viceroy of New Spain, expanded upon this policy of punitive raids with the promise of supplies if the hostile bands would agree to live in peace. This stick and carrot approach worked relatively well and from 1793 to 1821 the region experienced peace (Hawthorne 1994, Faunce 2000).

By the 1700s the Apache were displaced from east of the Pecos with the arrival of the Comanche who are the last American Indian group to move through this region (Hogan 2006). The Pecos River was usually considered the western border of the Comanche lands, leaving the areas west, and the project area, inhabited by Apaches (Hogan 2006).

Historic

In 1821 Mexico won independence from Spain which at that time included all of New Mexico but still showed no interest in the area. It was not until the USA gained control of the region after the Mexican-American war that non-American Indian people began to show interest in the area (Faunce 2000). However, raiding by Apache bands continued, all the way through 1860s, which discouraged settlement and forced the US military to launch several military campaigns in the area (Faunce 2000). Eventually this led to the creation of the Mescalero Apache Indian Reservation (Mehren 1969, Hawthorne 1994) and the subsequent removal of the last remaining American Indian groups to this reservation. This resulted in the pacification of the area and moving in of Anglo (White) ranchers, sheep and cattle, and miners during the 1880s.

The 1880s and 1890s brought homesteaders who were drawn to the area by the water provided by the Pecos River and railroad developments. The town of Carlsbad was founded in 1888 as the town of Eddy but would change its name to Carlsbad after the famous European spa of

Carlsbad, Bohemia (City of Carlsbad 2012). Drilling for water wells led to indications of oil and gas resources in the area and further exploration. In April of 1924 the Illinois #3 became the first viable producing well in south east New Mexico, which is the second largest oil and gas pool in the United States (Altschul et al. 2005). Oil and gas would prove to be the driving factors in development of the area since that point and a key driver of the archaeological research in the area.

Archaeological Data

The site location data for this project was gathered from the New Mexico Cultural Resources Information Systems (NMCRIS), the state-wide archaeological record database, of the Archaeological Records Management Section (ARMS), a part of the New Mexico Historic Preservation Division (NMHPD). As noted, state laws prohibit the unauthorised distribution of site location information. Thus that data is not included in this publication and the images are purposely designed to be vague about site locations (NMHPD 2012).

Archaeology of the Area

Gas and oil development over the last 90 years has worked with more recent heritage legislation to create an intensely surveyed area. The majority of the land in the area is government owned, thus requiring survey work. This fact, combined with aspects of the oil and gas industry, such as seismic survey and the relative density that wells can be placed through a landscape, means that a large proportion of the region has received archaeological survey, relative to other areas of the state (Altschul et al. 2005, Hogan 2006). The study area alone has had roughly 35,000 acres surveyed.

See Confidential Appendix- Figure 88, Figure 89, Figure 92

NMCRIS data

In the project area, over 858 archaeological sites have been identified in NMCRIS. Although previous investigations (Rocks-Macqueen 2010) have found that some sites have been recorded more than once and given different identification numbers. To ensure a detailed archaeological record for this project the NMCRIS site reports were examined and the resulting relevant information was manually entered into a database (Table 44). A significant number of sites have identifiable components, with some sites having multiple components. But another problem with dating the components is the plethora of terms and dates used between reports. For example, some sites were labelled as Apache but the date range from1539 AD to 1846AD or 1870AD or 1890AD (Table 44) with six different ways of describing the time/cultural period of those sites. This data was standardised, as seen in Table 43 and the results are as follows:

Period	Number of Sites
Clovis to Late Paleoindian	1
Late Paleoindian 8000BC to 6600BC	2
Terminal Paleoindian 6600BC to 5500 BC	1
Early Archaic	2
Early Archaic to Middle Archaic	1
Middle Archaic	3
Middle Archaic to Late Archaic	9
Early Archaic to Late Archaic	2
Late Archaic	55
Late Archaic to Unspecific Archaic 800BC to 200AD	1
Archaic	2
Unspecific Archaic	10
Plains Woodland to Panhandle Aspect 250AD to 1400 AD	1
Early Pithouse	1
Early Pithouse to Early Pueblo	1
Early Pithouse to Late Pithouse	2
Early Pithouse to Late Pueblo	15
Late Pithouse	16
Late Pithouse to Early Pueblo	7
Late Pithouse to Late Pueblo	69
Unspecific Jornada	28
Early Pueblo	7
Early Pueblo to Late Pueblo	28
Late Pueblo	8
Apache	11
Historic	92
Unknown/ Unspecific	541

Table 10: Site period component breakdown for project area. Sites can have more than one component.)

This surface data links up well with the general cultural history of the area. Limited activity in the Paleoindian to early Archaic periods with a steady increase in activity after this all the way up until the major cultural shift right before the Protohistoric period. The number of sites does not equate quality of sites (Table 11), as many sites are small lithic scatters of under 100 lithics.

Number of Lithics	**Number of sites**
0	76
1 to 9	106
10 to 99	303
100 to 999	156
1000 to 9999	39
10000 to 99999	10
Unknown	109
Number of Ceramics	**Number of sites**
0	585
1 to 9	82
10 to 99	29
100 to 999	4
1000 to 9999	1
10000 to 99999	0
Unknown	96

Table 11: Artefact distribution in known sites in the project area. Full data in Table 44.

See Confidential Appendix- Figure 93, Figure 94, Figure 95

Time Period Issues

There were significant issues with identifying the temporal period for those sites. Most sites had no C14 dates and almost all dating conventions were based on diagnostic surface artefacts, like projectile points or ceramics, which has problems. Projectile points could be reused 100s or 1000s of years after they were first created. Ceramics in the project area are not easily distinguishable. A common ceramic type is Mogolleon brown ware, which looks very similar to the later Protohistoric Apache ceramics that are also brown. The identifications are made in the field and it is unknown how many identifications were by a ceramics expert. In the Author's personal experience of CRM in New Mexico it would be rare to have a ceramics expert there to make an identification.

Of those sites with identifiable artefacts the NMCRIS database lists 328 sites as having prehistoric components. However, an investigation into the individual site reports for the project area found very different numbers. For example, 92 sites have a historical component according to their NMCRIS site reports, which is 20 more sites than listed by the NMCRIS database.

Lack of Excavations

A review of the ARMS records shows that as of the 1st of March 2012, out of 2040 archaeological activities listed, 2035 were surveys and only five were not survey actions. This was some limited testing on several sites, though very little was found (Clifton 1996; Gibbs 2003; Griffiths and Sciscenti 1996). The most promising result was the excavation of the Punta de los Muertos site which yielded C14 dates (Wiseman 2003). However, this site was heavily disturbed by looters and the information it provided is suspect. This lack of excavation data is indicative of the whole region. Only 51 site excavations were undertaken between 1990 and 2005 in south eastern New Mexico (Hogan 2006), an area comprising of Chaves, Curry, De Baca, Eddy, Guadalupe, Lea, Lincoln, Quay, and Roosevelt counties. It encompasses 31,590 sq. mi, an area about the size of the state of South Carolina.

This lack of excavation is because the study area is mainly owned by government agencies, there is little pressure to place oil and gas wells and supporting roads in any specific location due to partial ownership, i.e. they only have permission to build on certain owner's land. This situation means that when archaeological resources are encountered for oil and gas development these can simply be moved over to the next available area without archaeological resources. There is no pressure to only build in a limited area (Altschul et al. 2005). An example of this can be seen at site LA129788; the site report indicates that on encountering potential cultural remains, a hearth feature and only three lithics, the support road to the well was moved (Figure 21).

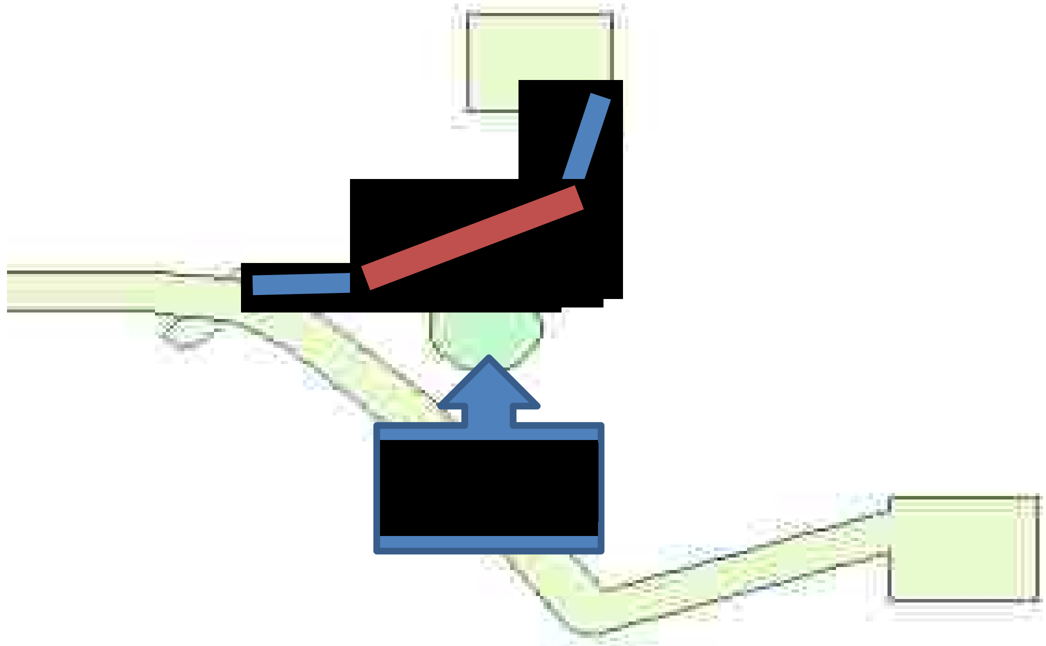


Figure 21: Image of construction near LA 129788, GIS shape files of surveyed areas (light green) and site (green). Blue line is original path of road and red line is the diverted road created to avoid LA129788.

Further investigations found that several excavations had taken place in the project area, mainly on cave sites. Some investigations had been done in the 1930s (Mera 1938) and in the 1960s (University of Wisconsin 1968); however, no legacy data exists for these projects in NMCRIS. In the case of the later excavations, the NMCRIS site record is the result of a letter sent in 1968 with very little further information. In essence, while in fact excavations have taken place in the project area, contrary to the NMCRIS activities query results, the data available from these projects is of very limited value for determining sites period(s) of use.

The almost complete lack of C14 dates, none from undisturbed contexts, and dating based on surface finds, and the associated issues with that method, meant the estimates of sites age in the database were almost complete guesses. They may have been educated guesses but still guesses and ones that covered a minority of sites in the project area. The majority of sites had no estimation of age. For this project that did not influence the outcomes as will be demonstrated in the discussion the area appeared to have the same use regardless of time period, except the recent historic. However, this does mean that significant more field work does need to be conducted to determine if this hypothesis is correct and that there are no differences in use between time periods.

Chapter 6: Movement Across the Landscape of Azotea Mesa with Least Cost Path Analysis

With the project area determined, agent based modelling software chosen and background information reviewed the project moved into model creation, planned activity 3. This chapter reviews the first model created for this project. It presents some of the prevailing theories about the project area, which were then used to guide the building of the models. As discussed in Chapter four the plane for the testing of explanatory models was to use the site location theories for the project area for modelling. This chapter specifically focuses on the movement of people in the project area as an explanation for the pattern of site placement.

Site Pattern Theories

When deciding on which theories to model, this project turned to the prevailing theories in the project area about why people would have been in the project area and thus the site distribution pattern. This was summarised by the PUMP III project:

> ‘We know that people came into the area, possibly in small, mobile groups that exploited locally available resources and then left, or as travellers following a favoured route from the river to the uplands, or possibly even as part-time agriculturalists establishing opportunistic fields at favourable locations to capture runoff. What we don’t know is which one or ones of these strategies they were pursuing, where they came from and where they went, how use of this area changed through time, and whether the structure of use changed as a result of organisational changes at a larger scale.’ (Altschul et al. 2005 p. 106)

There were several different hypotheses about why people were in the project area but it was unknown which of these hypotheses was correct. The strength of agent based modelling is that it allows for the exploration of different hypothesis and situations. This project did not need to pick a single route to test but could utilise the strengths of modelling to test all the different possible explanations. The next few chapters examine these different situations: foraging groups, traveling groups, or part-time agriculturalists.

Movement of People Plausible?

The first hypothesis tested was the idea of ‘travellers following a favoured route from the river to the uplands’. This was a plausible situation, given the large amounts of evidence for continual habitation in both the mountains and along the river compared to a lack of evidence in the project area (Previous Chapter). As the previous predictive model was aimed at subsistence patterns, e.g. vegetation, distance to water, it was possible that the poor results were caused by a mismatch between the behaviours that formed the archaeological record and the data used to model these behaviours. The archaeological record could have been the result of movement across the landscape from the river to the mountains or vice versa.

At the beginning of preparing for model creation concerns were raised that that the archaeological record did not indicate simple layover campsites on journeys between the two ecological zones. Examination of the NMCRIS records found that at least 39 sites are listed as having between 1000-9999 lithics (Table 11), viewable pieces in a surface survey, and at least ten sites are listed as having more than 10,000 lithics. These are just estimates but when combined with the listings of a wide range of features from mortars to roasting pits, it is hard to view these sites as simple one-night camping locations on a journey. But one could not rule out travel as the main reason for site placement without testing the hypothesis. This project proceeded to test this hypothesis to attempt to confirm one way or the other if it accounted for site placement.

Least Cost Path Analysis

A common way to analysis potential routes of movement in archaeology is to use a least cost path analysis (Gillings and Wheatley 2002, Conolly and Lake 2006). Least cost path analysis calculates best path from a source to a destination based on the costs. Cost is defined by the modeller, from distance to calories used. In nearly all of these investigations, the aim has been to reconstruct ancient routes and route networks (see Fábrega Álvarez and Parcero Oubiña (2007), Schriever, Shackley et al. (2010)) or to identify the principal factors governing the construction of known roads or road segments (Bell and Lock 2000, Kantner and Hobgood 2003, Jeneson and Verhagen 2012) . Until recently, it was rare to see this in predictive models and only a few projects are currently undertaking such work (see Bertoncello, Brughmans et al. 2012).

GIS Methodology and Problems with Least Cost Path

At the beginning of this project, least cost path analysis appeared to be a possible tool to explore site locations, as determined by 'travellers following a favoured route from the river to the uplands pathways across the project area.' There had been critiques of least cost path analysis, like the issues of implementing anisotropic costs in GIS (for full review see Husdal (2000), Gillings and Wheatley (2002), Conolly and Lake (2006)). Anisotropic costs are direction dependent; an example of this would be slope. A traveller going directly up a hill with a 30% gradient will experience significant anisotropic costs. However, if they would travel perpendicular to the direction of the slope it would be the equivalent of walking on a flat surface, with a slight slant to one's left or right. Unfortunately, GIS programs have problems implementing these directional forces, as most layers can only account for force in one direction. Unless a person travels a straight line their path will alter direction multiple times, thus changing their aspect and how they interact with anisotropic costs.

Some GIS models will create multiple layers accounting for force from different directions but they are awkward and do not work well. A lack of integration of anisotropic costs was the primary reason for deciding to move beyond GIS into agent based modelling, which can account for these changes in direction. However, there are multiple problems that have been raised with regards to the use of least cost path analysis in GIS (for a good review see Herzog (2012)) and this was not the only reason to use agent based modelling:

- Incompatible results across different GIS platforms e.g. results gathered in GRASS GIS cannot be replicated in ArcGIS. (Doneus, Gietl et al. 2008).
- Different neighbour sampling methods result in sub-optimal routes (Bevan 2012).
- Some GIS packages do not implement Dijkstra's algorithm in all parts (Lee, Munro-Stasiuk et al. 2003).
- GIS software is incapable of creating zigzag or hairpin curves seen on many roads and trails (Llobera and Sluckin 2007).

Which Costs?

The first step in creating a least cost path analysis was to determine the costs involved in travel. One was vegetation, but after investigating this as a travel cost it was discarded. The majority of the project area contains grasses or small shrubs that would not impede travel. Even larger barriers, such as pinion trees, were not significant enough at the scale used. The topographic data is 31 x 31 m and none of the vegetation was that large or in densities that big.

Soils and waterways were also considered as costs. As discussed in the following chapter, waterways in the project area are not perennial or large. Except for the occasional flash flood, streams would not impede travel. Soil at first seemed more promising, as the soil reports listed many soil types as hard to transverse (see Chapter 8). Under closer examination it was determined that while some soils are not conducive to comfortable travelling, like walking over hard pointed rocks in moccasins, the greatest cost is actually associated with slope. Walking up or down a slope greatly affects the speeds with which one can cross a landscape and the energy expended in such an endeavour. Thus the primary variable used as cost in this project is slope.

Costs relating to social issues were also considered. For example, if part of the subject area was owned/controlled by a rival group, then one may not cross this area to avoid conflict. While surely a factor in many past cultures, unfortunately no research found any evidence to support the existence of such 'no-go' zones in the project area. A common occurrence with least path models as noted by Van Leusen (2002 p. 6), 'it is unlikely that 'social' costs can be established with any degree of objectivity'. This assertion is also supported by other archaeologists as well: Bell, et al. (2002) point out that political boundaries, religious taboos or attractions leave hardly any mark in the archaeological record. Without evidence, it is impossible to model such an event other than as a wild guess or as a what-if exercise.

Slope as Cost

The cost incurred from slope can be measured in several different ways; one is time. Archaeologists often represent slope effects on walking as a time cost using Tobler walking function (Equation 11). This formula was first introduced to archaeology two decades ago (Tobler 1993). The equation calculates the walking speed dependent on the slope of the transverse area and thus time to cross that distance.

Kmp = 6 (exp -3.5 * abs (dh/dx + 0.05))

Equation 11: Tobler's Hiking Function. Six multiplied by Napier's number (2.718) raised to -3.5 multiplied by the absolute value of the slope (rise over run) plus 0.05.

Another measurement of cost that has been used is the calorific cost: how many calories one uses to travel across a certain landscape. Herzog and Posluschny (2011) supported the formula presented by Llobera and Sluckin (2007), which was based on a large sample of metabolic cost measurements, as the best function for modelling pedestrian movement. However, (Herzog 2012) now argues that Minetti's (Equation 15) is the best model of travelcosts. It should be noted that Llobera and Sluckin were not trying to create an optimal path solution in the traditional sense. For example, they examined an optimal path for summiting a peak which, in most archaeological studies, is not done.

M(s) = 2.635 + 17.37s + 42.37 s^2 – 21.43 s^3 + 14.93 S ^ 4

Equation 12: Llobera and Sluckin's quartic polynomial. Units are in KJ m ^-1.

The GRASS GIS program has its own algorithm for calculating cost distance: r.walk. This linear cost function is based on the (Langmuir 1995) rule of thumb calculations. It is:

A * Δ_D + b * Δ_H_up + c * Δ_H_gd + d * Δ_H_sd

Equation 13: Δ_D is the distance covered in metres; Δ_H_up is altitude difference in metres; Δ_H_gd gentle and Δ_H_sd steep downhill differences. The default values for the varibles are: a = 0.72; b = 6.0; c = 1.9998; d = 1.9998.

Wheatley (2002) used a 'backpackers equation' (Equation 14) for his cost path analysis. In that case it was Erison et al. (1980):

$$\mathbf{\Delta\_D + 3.168 * \Delta\_H\_up + * abs\ \Delta\_H\_down}$$

Equation 14: 'Backpackers equation' Δ_D is the distance covered in metres; Δ_H_up is positive altitude difference in metres; Δ_H_down is negative altitude difference in metres.

The minimum cost function of the 'backpacker's equation' is for walking on level ground. However, studies have found that the actual optimal is on a 10% down slope (Ferretti, Minetti et al. 2002). A method to model this has been put forth by Minetti et al. (2002):

$$\mathbf{1337.9\ s\text{^}^{6} + 278.19\ S\text{^}^{5} - 517.39\ s\text{^}^{4} - 78.199\ s\text{^}^{3} + 93.419\ s\text{^}^{2} + 1.64}$$

Equation 15: Minetti's cost path equation. s is mathematical slope.

Any one of these equations would have worked for this project. While the formulas vary, their outputs are similar (Figure 22 & Figure 23); it is only when one gets to extremes slopes that there is a great deviation in the returns. Initially all the formulas were going to be tested to determine if even the smallest difference between the formulas would affect the outcomes. However, only Tobler's equation was used. It was the first equation tested and by the time other equations were to be tested it had already been determined the least cost path analysis was not a viable methodology, probably for any project.

Figure 22: LCPs between Hennef-Geistingen and Olpe (Herzog 2010 Figure 3 p. 434)

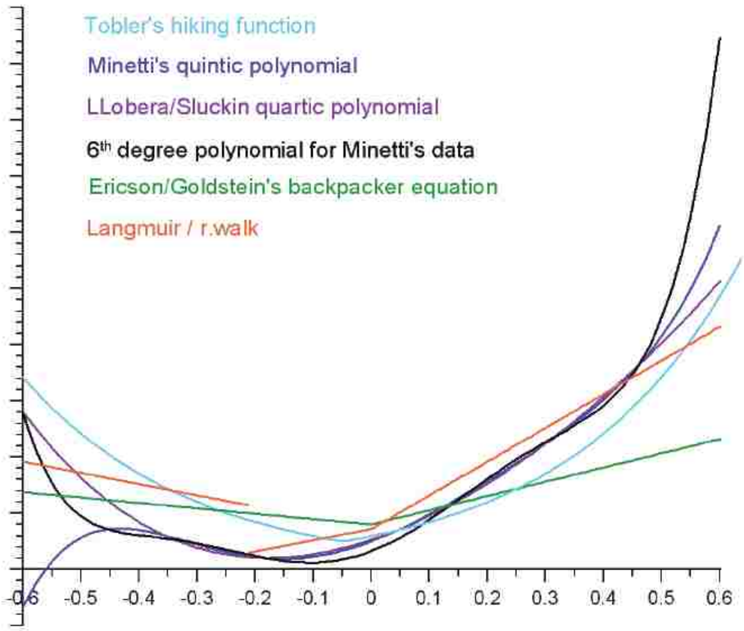


Figure 23: Some slope-dependent cost curves (Herzog 2010 Figure 1 p. 431).

Optimal Path

Determining costs does not produce a path; for that to occur; the costs need a decision algorithm. Traditionally, this has been done through a Dijkstra algorithm (Dijkstra 1959). Conceived by Dutch computer scientist Edsger Dijkstra in 1956, it is a graph search algorithm that solves the single-source shortest path problem. It produces a shortest path tree for nonnegative path costs. For a given source node in the graph, the algorithm finds the path with lowest cost between that node and every other node.

$$O(|E| + |V| \log |V|)$$

Equation 16: Dijkstra algorithm.

The Dijkstra algorithm works by:

1. Assigning to every node a tentative value: set to zero for the initial node and to infinity for all other nodes.
2. Marking all nodes unvisited and setting the initial node as current.
3. From the current node, considering all of its unvisited neighbours and calculating their tentative cost. For example, if the current node A is marked with a cost of x and the edge connecting it with a neighbour B has a cost of y, then the distance to B (through A) will be x + y. If this distance is less than the previously recorded tentative cost of B, then overwrite that distance. Neighbours, even examined, remain as an unvisited set.

4. When all neighbours have been considered the current node is marked as visited and removed from the unvisited set. A visited node is never checked again.
5. Setting the unvisited node marked with the smallest tentative costs as the next 'current node' and repeat.
6. The algorithm ends if the destination node has been marked visited. (Dijkstra 1959)

Implementation of Cost Path Analysis in Agent Based Modelling

The computer code examined in this chapter is divided into two types: least path calculating and optimising the running of the model. Calculating path is a self-explanatory utility and the code is laid out to help readers find any errors in implementation of the theory. The optimising code is code that has been inserted into the model so that it can run faster and complete the calculations in a timely fashion: hours instead of days. However, optimising a model can cause unforeseen glitches or errors in a model, which can affect the outcomes. Major sections of the code containing key assumptions are examined in the following section with the purpose of finding potential faults with the model. The full code is in Appendix C.

Agent Based Model Design

Before examining the code used it is important to clarify some NetLogo terms. The agents in NetLogo are referred to as turtles. These agents can move around anywhere in the NetLogo world in which the simulations run. There is a second type of agents known as patches. These agents are static and cannot move around the world. However, they can interact with each other and with turtles. Patches create the environment that the agents use to interact. In the models below, GIS raster squares are converted to patches.

GIS capacity is not a normal function of the NetLogo program and it must use an extension, the GIS extension, to work with GIS data. The full list of the custom commands for the GIS extension can be viewed in the NetLogo Directory and on the web version of the directory – http://ccl.northwestern.edu/netlogo/docs/. The primary developer of the GIS extension was Eric Russell, who makes use of several open-source software libraries including:

- Java Topology Suite
- JScience
- Java Advanced Imaging
- Apache Commons Codec
- Apache Jakarta HttpClient
- Apache Commons Logging

It also contains elements borrowed from My World GIS.

NetLogo currently only supports raster and vector files from ArcGIS and any dataset needs to be converted to those file types. There are no known issues with using ArcGIS files and the NetLogo extension but a problem with the extension could result in poor or inaccurate results.

ASTER GDEM2

For this project, the elevation data was the ASTER GDEM2 data discussed in Chapter 5. The data was first imported into ArcGIS where it was trimmed to fit the project area. The NetLogo GIS extension can handle GIS data as a layer of patches but a GIS program like ArcGIS or QGIS is needed to handle most functions such as reshaping project areas, combining datasets and buffering.

Convolution, Potential Problems

It was easier to import a single elevation layer into NetLogo and then use code to create separate aspect and slope datasets than trying to create the different layers in ArcGIS and importing them. This was done with code borrowed from the NetLogo GIS example in the Example Libraries. This process involves several methodological choices that, as will be demonstrated, can change the model outcomes. The following code was used to create these datasets in the NetLogo program.

```
let horizontal-gradient gis:convolve elevation 3 3 [ 1 1 1 0 0 0 -1 -1 -1 ] 1 1
let vertical-gradient gis:convolve elevation 3 3 [ 1 0 -1 1 0 -1 1 0 -1 ] 1 1
set slope gis:create-raster gis:width-of elevation gis:height-of elevation
gis:envelope-of elevation
set aspect gis:create-raster gis:width-of elevation gis:height-of elevation
```

Figure 24: Code example of NetLogo code used to convert DEM to NetLogo patches.

These commands create two variables, horizontal-gradient and vertical-gradient, using convolve. A convolution is a mathematical operation that computes each output cell by multiplying elements of a kernel with the surrounding cell values. In this case, a kernel is a matrix of values. The centre cell being the ‘key element’ that is computed from the surrounding cells.

In NetLogo the values of the kernel matrix are given as a list, which acts as the elements of the matrix following the pattern of left to right, top to bottom. For example, a 3-by-3 matrix would be listed in the following order:

1	2	3
4	5	6
7	8	9

The ‘33’ lets the program know that it is a 3 x 3 matrix and the numbers that follow are the values. In this model, the horizontal-gradient appears as:

1	1	1
0	0	0
-1	-1	-1

The vertical-gradient is:

1	0	-1
1	0	-1
1	0	-1

Convolution is a critical step in this methodology. It determines how the GIS data translated to the NetLogo world. Any change in this matrix will alter the values in NetLogo and thus potentially the outcomes. For example, one could use a Sobel Operator (-1 0 1 -2 0 2 - 1 0 1 & -1 -2 -1 0 0 0 1 2 1) or a Scharr Operator (3 10 3 0 0 0 -3 -10 -3 & 3 0 -3 10 0 -10 3 0 -3) to convert the GIS data to NetLogo values, with very different results. To test the potential difference in results the final cost path model was tested using three different operators: the one provided in the GIS example code, the Sobel Operator and the Scharr Operator. Figure 25 highlights the different outcomes of these operators with a limited sample of one location to 20 destinations. Any change in the operator used for convolution will greatly influence the outcomes: a potential problem that will be discussed later in this chapter.

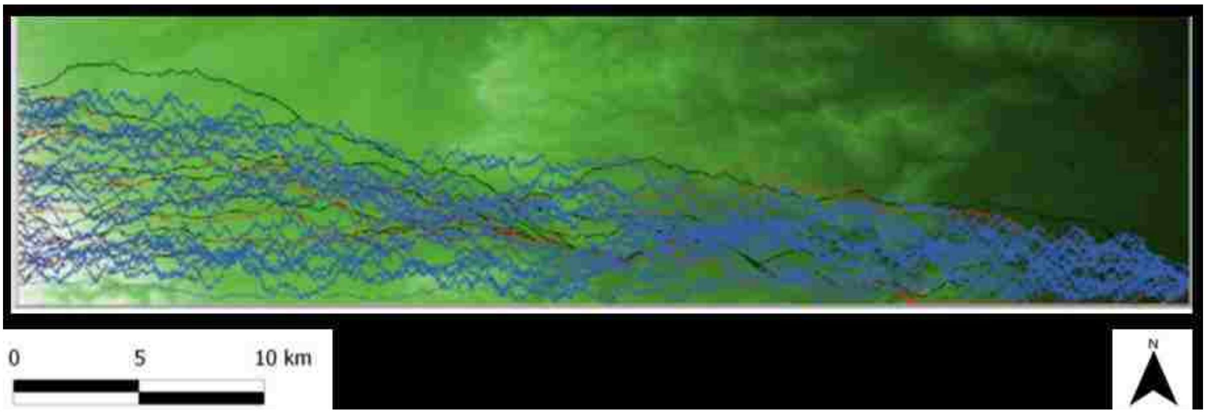


Figure 25: Different cost paths based on the different convolution operators. Black – normal, Red – Sobel, Blue – Scharr.

If convolution changes the outcomes of the model, then why create aspect and slope datasets; why not use elevation only for the model? It is possible to use the underlying elevation dataset to measure the slope between any two given points. The problem with this is that it changes the directional forces of the slope. It makes all calculations of slope relative to the position of the agent instead of based on the aspect of the landscape. This results in a different result for the least cost path (Figure 26). The limitations of the model forces the use of aspect.

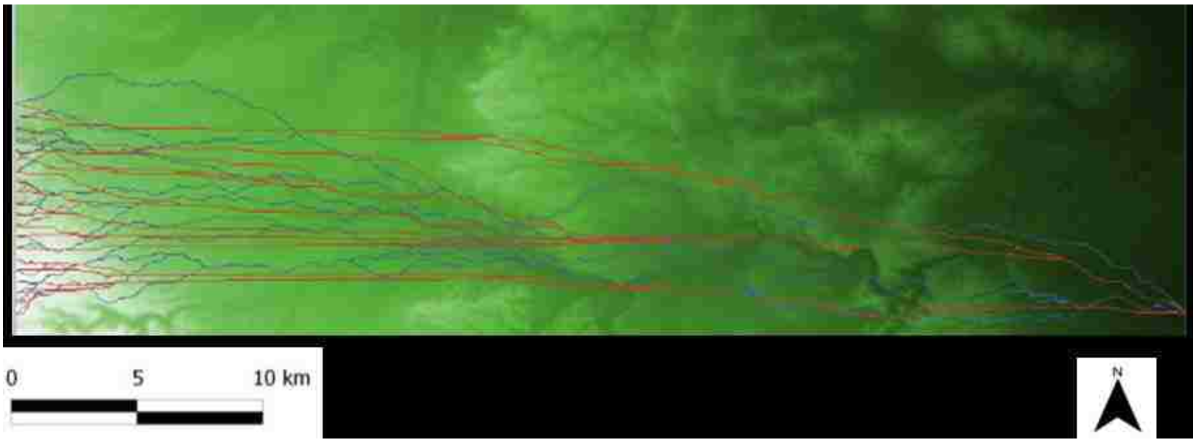


Figure 26: Comparison of least paths based on elevation (red) versus aspect and slope (blue).

Sampling Method Issues

If the NetLogo world is not set to the exact same size as the GIS data i.e. 100 patches for 100 raster cells, then the values have to be translated from the raster dataset to the patches. How one determines this translation could change the outcomes of the model. The sampling method could be any one of the following options:

- 'NEAREST_NEIGHBOUR': the value of the cell nearest the sampling location is used.
- 'BILINEAR': the value of the four nearest cells are sampled by linear weighting, according to their proximity to the sampling site.
- 'BICUBIC': the value of the sixteen nearest cells are sampled, and their values are combined by weight according to a piecewise cubic polynomial recommended by Rifman.
- 'BICUBIC_2': the value is sampled using the same procedure and the same polynomial as with BICUBIC above, but using a different coefficient. This method may produce somewhat sharper results than BICUBIC, but that result is data-dependent.

  These different sampling strategies were tested to see if they changed the model outcomes.

Table 12 runs through the different combinations tested, in order. Over a test of 20 paths, no differences in results were noticed for this model.

Elevation	Slope	Aspect
NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR
BILINEAR	NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR
BICUBIC	NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR
BICUBIC_2	NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR
NEAREST_NEIGHBOUR	BILINEAR	NEAREST_NEIGHBOUR
NEAREST_NEIGHBOUR	BICUBIC	NEAREST_NEIGHBOUR
NEAREST_NEIGHBOUR	BICUBIC_2	NEAREST_NEIGHBOUR
NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR	BILINEAR
NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR	BICUBIC
NEAREST_NEIGHBOUR	NEAREST_NEIGHBOUR	BICUBIC_2
BICUBIC_2	BICUBIC_2	BICUBIC_2
BICUBIC	BICUBIC	BICUBIC
BILINEAR	BILINEAR	BILINEAR

Table 12: Different sampling methods of least cost paths tested.

Potential Problems from Optimisation

Checking every single patch in the project, while thorough, is inefficient. To help increase the efficiency, reducing the time it takes to run the models, this project implemented a code to reduce the number of patches that would be searched. The assumption was that even in ideal conditions a route could only be so far away from a straight route to the designation. Essentially, the code ignored any areas too far away from the agent. A concern was that by limiting the range of calculations the model could be creating an edge effect. Where certain paths were excluded, not because they were unlikely to be the quickest route but because the optimisation code ruled out such a possibility.

As before, the completed model was tested to determine the influence, if any, this would have on the results of the model. As seen in

Figure 27, there were changes in possible routes when the range was shrunk down to a very thin range, i.e. a radius of 30 patches, but after about a radius of 100 patches there is very little change. Thus all models were run with the distance of deviation set at 100.

Figure 27: Difference between scanned zone. Radius 30 (Black) 50 (Red) 70 (Yellow) 100 (Blue) 150 (Pink).

Optimisation of Calculations

The cost path analysis behaviour model was based on a Dijkstra algorithm. Using this method, each squared is checked, and then the path to the next set of squares are calculated in order, until the goal is reached. Each square points to the previous square that had the lowest cost path to reach it. However, instead of having patches undertake this operation, it is quicker to have patches create turtles on a new patch to undertake the calculations.

To speed up the process, only certain patches, i.e. those within the radius discussed above, undertake the calculations. However, NetLogo will still have to search through every patch to find those that need to undertake the calculation; a very inefficient process. It is quicker to have the information transferred by turtle agents between the patches than to run through every patch.

The code creates eight agents (turtles with the breed mark) with each facing a separate 45 degree direction: 0 degrees is the top of the screen in the NetLogo world (Figure 28). The agents then calculate the cost between the centre point of the patch to the centre point of the patch it is facing. It then adds that cost to the inherited cost from the patch it started from and then passes that information onto the patch it is facing. A test of a version that ran with the slower patch only method was found to have no difference in cost path results, except that it took significantly longer, roughly 2.5x, than the model that employed the mark breed turtles. This method gave the same results at a much faster speed but could result in some unforeseen outcomes in the model.

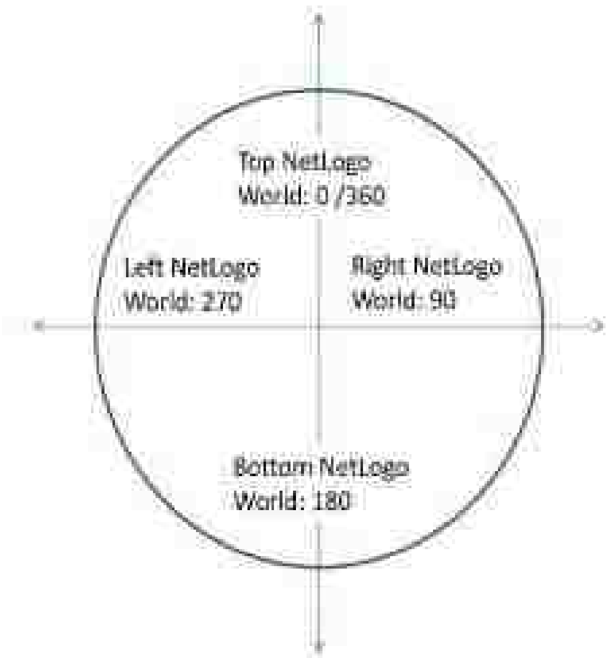


Figure 28: Direction in degrees in the NetLogo world.

Directional Costs

As discussed earlier, GIS least past cost path analyses rarely use multiple directional costs; instead, they deal with single directional costs. This agent based model was specifically designed to take into account direction from the agent's perspective. Yet, it was not known if this would actually affect the outcomes of a least cost path. A scenario was run to test this in which two agent based models were designed, one to take into account all directional resistances from the agents' perspective i.e. forwards, side to side, reverse, etc. and one to only take into account one direction cost: i.e. travelling west in the model or 270 degrees in NetLogo (Figure 28). Essentially, one model only calculated the costs of moving forward while the second calculated costs if agents moved laterally i.e. possibly creating a zig zag effect of movement up hills.

Figure 29 shows the results, which indicate that different directional forces do change the path taken.

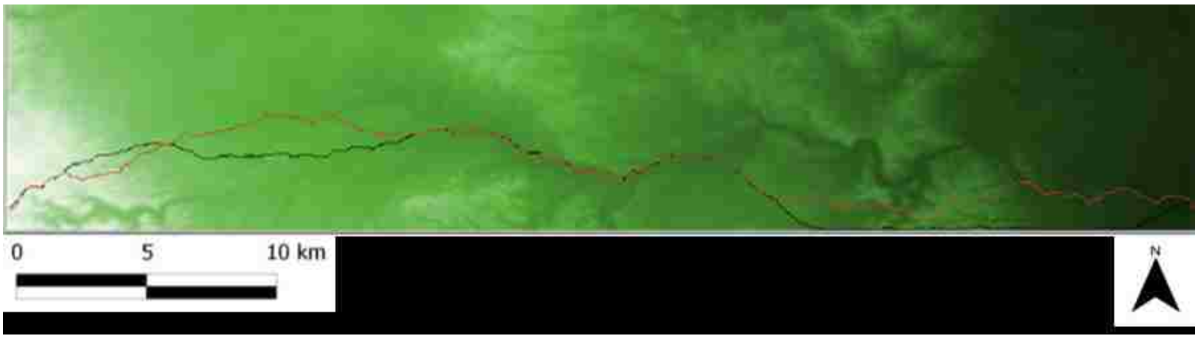


Figure 29: Multi-direction force cost path (Red) versus single direction cost path (Black).

When this is expanded to more interactions (Figure 30) it is possible to see that this is not a one-off occurrence. There are significant differences in the routes taken by the different methods across the whole project area. Some are minor fluctuations in the path and can be ignored as mirror deviations; however, others cannot. For example, several paths are over several kilometres from each other. That is a significantly different route of travel to the same location.

Figure 30: Multiple destination multi-direction path (Red) versus single direction cost path (Black).

Likewise, when examining the difference between a multi-directional and slope only, without directional applications, a difference is noticed in the paths (Figure 31). This was done by removing the code that dictates directional influences on costs. It is of note that when directional forces are removed the paths smooth out.

This project opted to use the multi-directional method for calculating costs. A decision made because a multi-directional model is a better representation of the actual forces one would encounter in the real world.

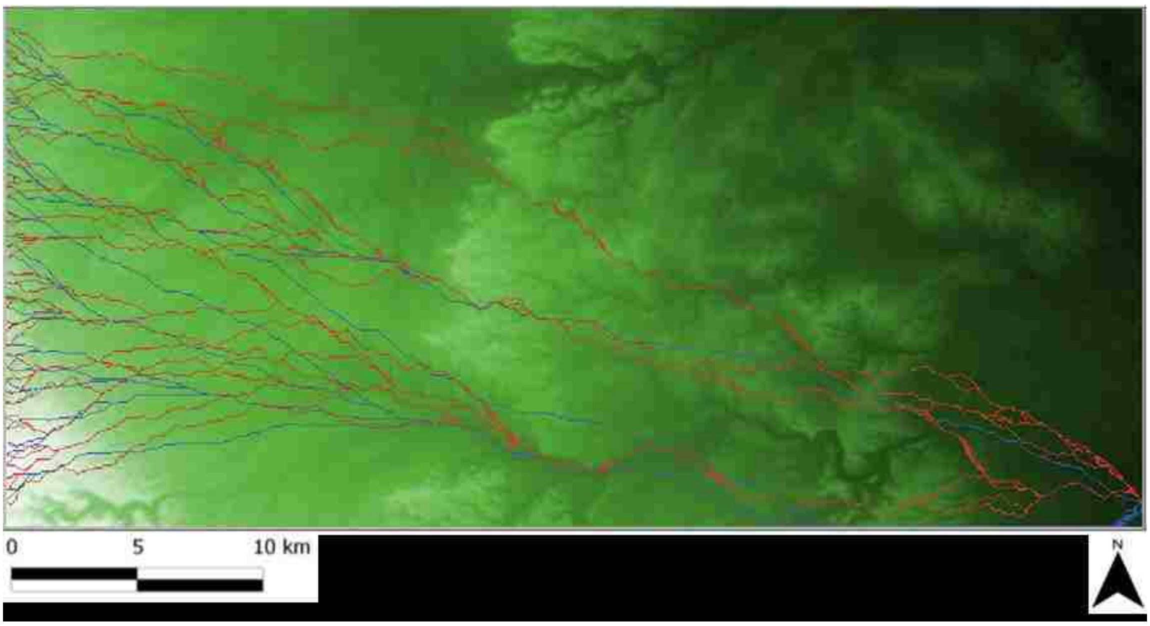


Figure 31: Multiple destination multi-direction path (Red) versus no directional cost path (Blue).

Simple Model Test and Distortion

It has been noted that some GIS software programs, like ArcGIS, fail the 'simple model test' (Bevan 2012). This was a test applied by Bevan to least cost paths; it states that if one were to have a perfectly flat surface then the least cost path should be a single line between two points. ArcGIS cannot produce a straight line; instead,, it produces a jagged one (Figure 32). Bevan attributes this to the fact ArcGIS only samples the immediate 8 surrounding pixels to calculate costs. He uses GRASS GIS as an example of a GIS system that samples 16 locations and passes his 'simple model test'.

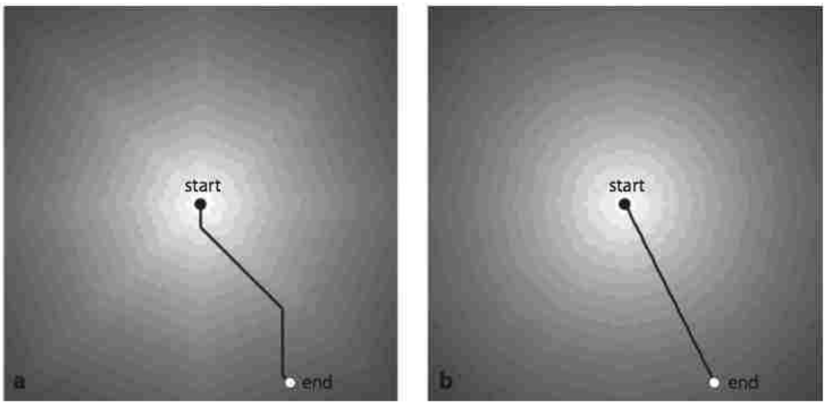


Figure 32: Simple models and cost surface problems. Cost surfaces and a least cost path calculated on a flat 100 x 100 surface, for two kinds of spreading algorithm: (a) D8 in ArcGIS, and (b) D16 in GRASS GIS (Bevan 2012 p. 6 Figure 1).

However, the solution proposed by Bevan, sampling more pixels, may not actually solve the problem. The jagged edge is the result of the use of a Dijkstra algorithm on a raster that causes a distortion. Due to this distortion, there are errors; the length of the actual path exceeds the length of the optimal path by about 20% in a D8 configuration (Gillings and Wheatley 2002, Adriaensen, Chardon et al. 2003). Increasing the number of sampled squares can lower this but only so much. As demonstrated by Herzog (2012) (Figure 33), the error rate is reduced to 11% for 15 neighbours but only 4.6% with 120 neighbours. It is likely the results seen by Bevan for the D16 sample conducted by GRASS GIS is the result of other factors; GRASS GIS uses a modified version of a Dijkstra algorithm and samples its data directly from the elevation layers. As seen in Figure 26, the results of working directly from elevation datasets can greatly alter, and notably straighten, the results.

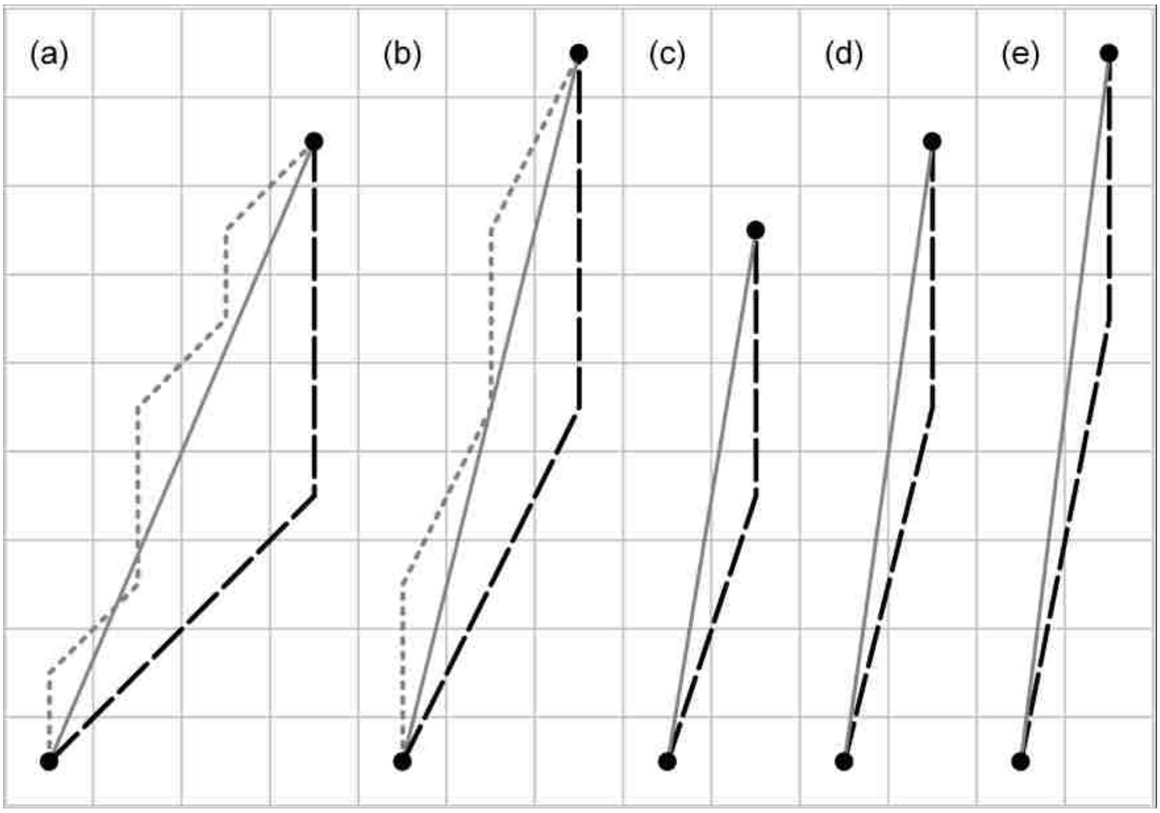


Figure 33**:** Worst case scenarios for different neighbourhood sizes, when a straight line is optimal but LCPs deviate from the straight line due to the raster to graph conversion. Neighbourhood sizes: (a) 3x3, (b) 5x5, (c) 7x7, (d) 9x9, (e) 11x11. The maximum distance of the worst case LCP (dashed black) to the true optimal path (grey straight line) decreases with increasing neighbourhood sizes. The dotted grey polylines are LCPs incurring the same cost as the worst case LCPs (Herzog 2012 p. 9 Figure 2).

Distortion and Agent Based Models

All of the models used throughout this project were based on an eight neighbour sampling system (Figure 34). However, the decision was made to investigate a wider variety of sampling ranges to see how they affect the agent based model. This involved more trigonometry to determine the correct angles and length of a patch that agent would cross (Figure 34).

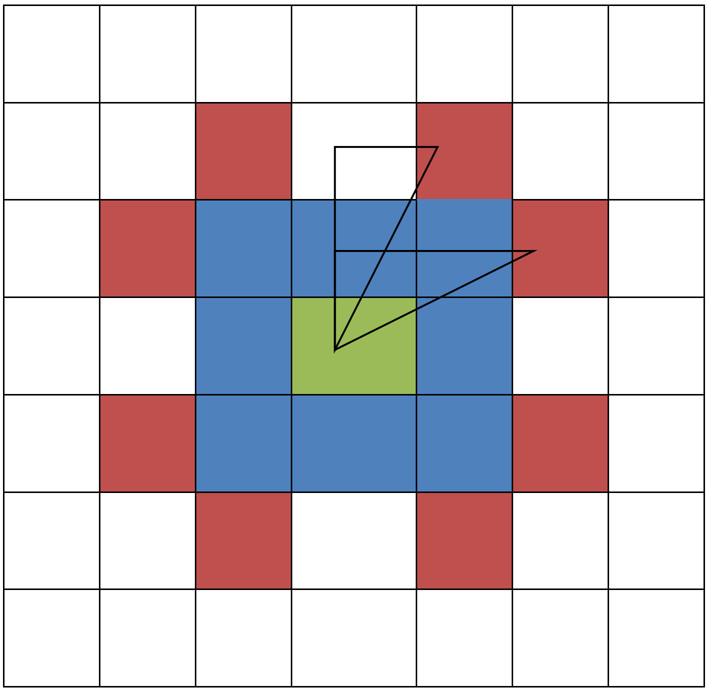


Figure 34: Eight and 16 square coverage of calculations. Angles to reach 16 square coverage are: 26.6 and 63.3 degrees.

Testing of the 16 neighbour model did not result in any deviation from the paths taken by the eight neighbour model (Figure 35). Moreover, it did not result in any improved ability to find the optimal path in the dump model (Figure 36). Essentially, the agent based models still suffer from distortion when making calculations about optimal routes to travel. The jagged paths found when all costs are removed are still seen.

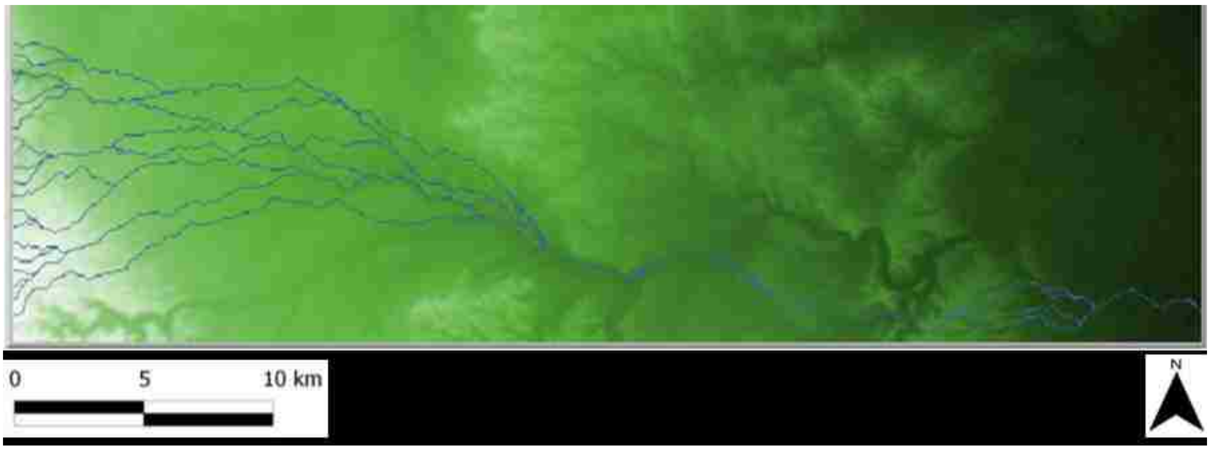


Figure 35: 16 neighbour (Blue) and eight neighbour (Red has been covered over by blue) models.

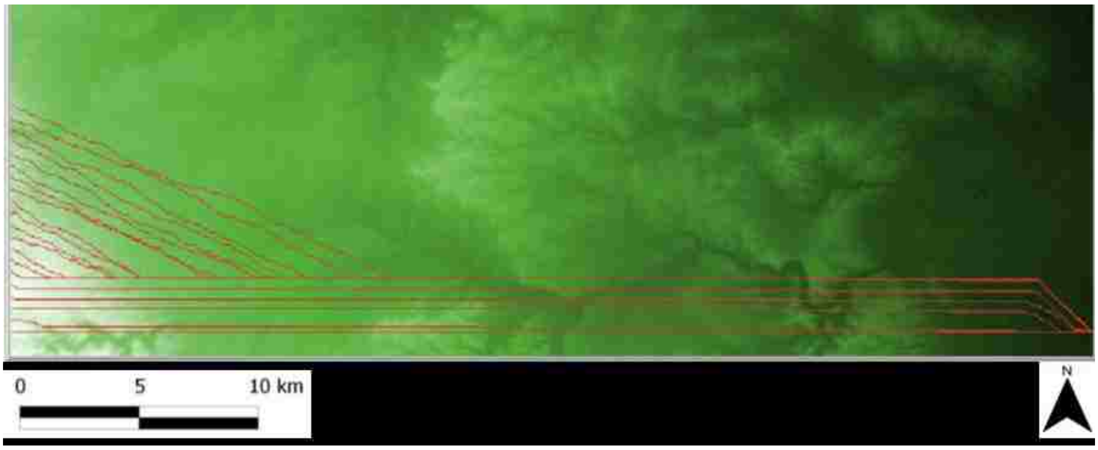


Figure 36: 16 neighbour (Blue-covered over) and eight neighbour (Red) solving the dump model.

Full Model Test

The fact that optimisation and sampling methods did not change outcomes was reassuring that the model would represent the real world. Moreover, a multiple directional cost calculation meant that the model was more representational of the real world. However, the fact that it still suffers from distortion and that convolution operators can dramatically change the digital representation of the area meant that errors could still affect the modelling outcomes. Even with these reservations, it was decided to test the model to see if any observable patterns could be discerned in the project area.

The first model was designed to examine the natural paths through the project area. This meant calculating multiple paths from different points along the natural Pecos River to the Guadalupe Mountains corridor, the main travel avenue hypothesised by the PUMP III report. A total of 40 points destination and beginning points on both edges were used. One point would plot paths to 40 other points on the other side of the project area. Figure 37 shows such a pattern, with all the routes from a single point on the eastern edge of the project area to the arbitrary points on the western edge. Also included, are the straight line paths to and from these locations.

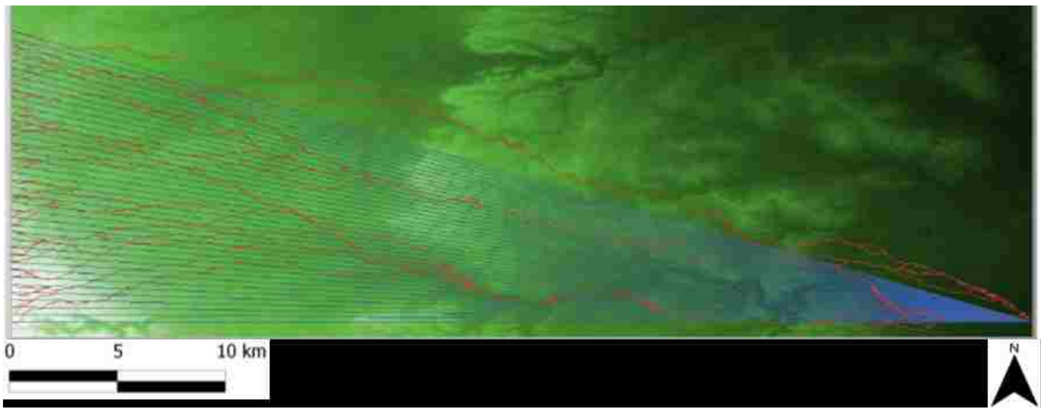


Figure 37: One location to 40 possible end points with straight line path (blue) and least cost path (Red).

Why Arbitrary Beginning and End Points?

At first, the known archaeological sites along the eastern and western edges of the project area were going to be used as starting and ending points. However, upon further consideration this option was discarded because we do not know the exact location people were coming from. They could be coming from sites hundreds of miles away, well outside the project area. Arbitrary even spaced points could capture the general travel routes throughout the whole area without having to know specifics.

East and West

The models were run from heading west to east (Figure 38) and vice versa (Figure 39). They were then combined to show all travel possibilities along the E-W/W-E travel corridor (Figure 40). There are a few locations on the edges that are missed but this is due to an edge effect of the model, which did not run routes near the edge. The results showed that the pathways missed some notable clusters of sites i.e. Black Canyon was missed (**Confidential Appendix Figure 96**). However, it did correlate to some sites and might explain the positions of those sites.

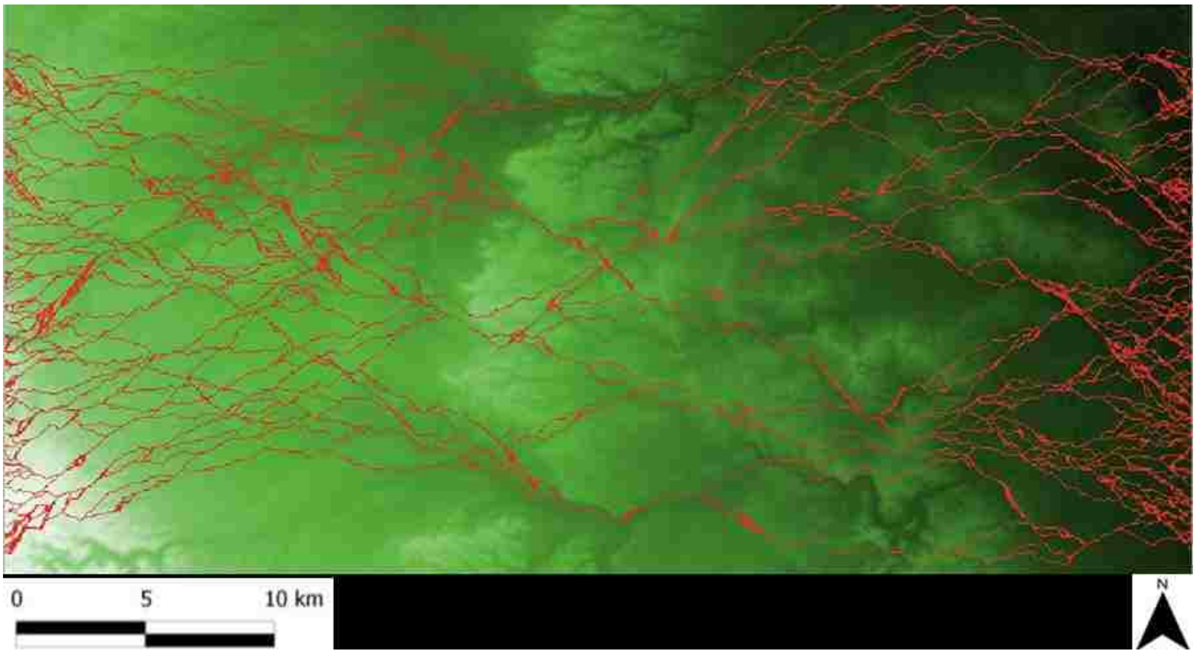


Figure 38: Western to eastern least cost paths across the project area.

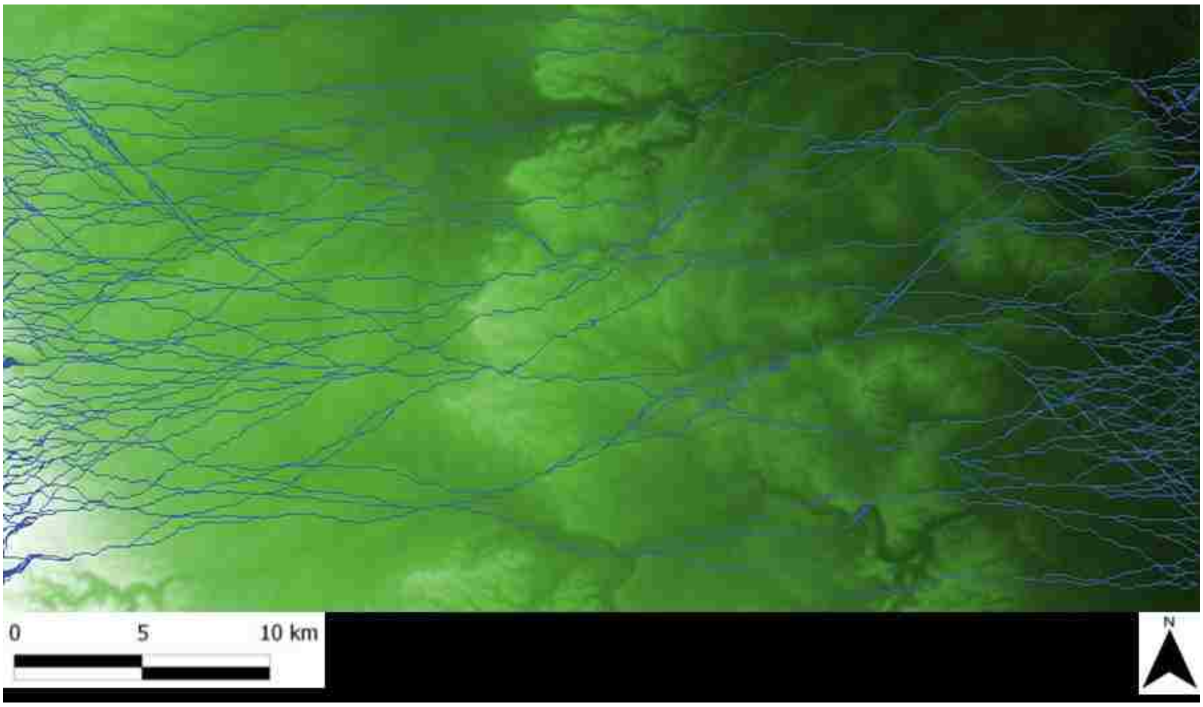


Figure 39: Eastern to western least cost paths across the project area.

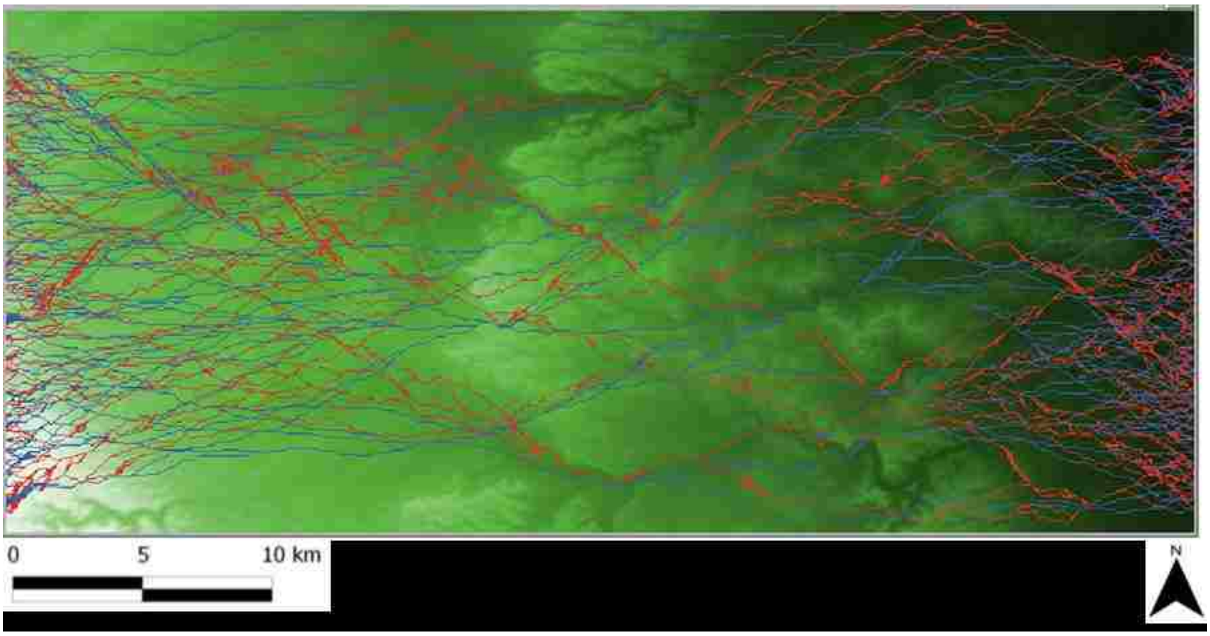


Figure 40: Combined eastern and western travel routes.

Vision and Cost Path

A possible explanation for the poor predictive results is that the least cost path analysis does exactly what it is meant to do: find the optimal route. An optimal route assumes perfect knowledge of all past environments and costs; an assumption that is not known. Constant travel in an area would cause a person to gain local knowledge of the landscape, to a level that might be close to perfect and, accordingly, the costs associated with travelling across it. Yet, this is an assumption and we do not know the level of knowledge individuals might have had.

It was decided to test a model that took into account both costs and the knowledge of those costs. This involved adding two new components to the model, a vie shed procedure and a decision procedure. The viewshed model is used to determine which aspects of a landscape a traveller would have been aware of as they travelled. This information is then acted upon by the decision procedure to calculate the best path according to the knowledge available to the agent. For this, the goal was set as the farthest point that the agent is aware of along the same direction as the destination. The agent knows it wants to end at a certain direction but it is navigating by heading to the farthest point in that same general direction that it is aware of; in other words, a waypoint model.

The results of this model show a significant difference in paths taken by the limited knowledge based model and the traditional perfect knowledge based model (Figure 41). The decision procedure for the model tested was based on distance travelled. After a set distance, the agent re-evaluates its surroundings and chooses a new path. Tests show that, as one would imagine, changing the decision procedure changes the path outcomes.

Figure 41: Local knowledge cost path based off of viewshed (Red) compared to the normal cost path (Blue). 70 Patch decision setting.

Errors – People are Not Perfect

Even when using a limited knowledge model the assumption is still made that people are able to calculate travel costs perfectly. That assumption seemed overly optimistic about a person's ability to estimate travel times. It is unlikely that a person would be able to calculate correctly the optimal path 100% of the time or even the majority of the time. Revaluating this assumption requires that the models be examined in the context of imperfect knowledge.

To create 'errors' in knowledge the cost of each patch was changed using different formulas. The first one tested was a random float rate function in the NetLogo program. Several different percentages were tested to see what would happen if agents incorrectly estimated costs of travel, e.g. over or underestimate costs by 5%, 10% etc. The results can be seen

Figure 42:

Figure 42**:** Cost path with random-float error rates. Error rates: black- 0%; grey- 1%; red- 2%; orange- 5%; brown- 10%; yellow- 15%; lime- 20%; turquoise- 25%; cyan- 30%; blue- 50%.

Using a random float rate there was very little change in the paths undertaken by the agents, even when set to a 50% error rate. This is because the random float method used could not guarantee 50% difference between the reality and the agent's view but any number between 0 and 50%. One patch could be off by 1% and another by 3% and a third by 45%. While there could be significant difference, the range blunted the impact.

To explore further the influence of errors in judgment on cost path analysis a different approach was undertaken. That was to eliminate the random number but keep the random + or – error in estimation. This involved setting a constant error rate with this code: set pc pc + (pc * (((random 2) * 2) - 1) * er). Figure 43 shows the results. Even with a difference of 50%, with the random + or – results in a potential difference of 100% between to patches, the results show only minor changes in paths.

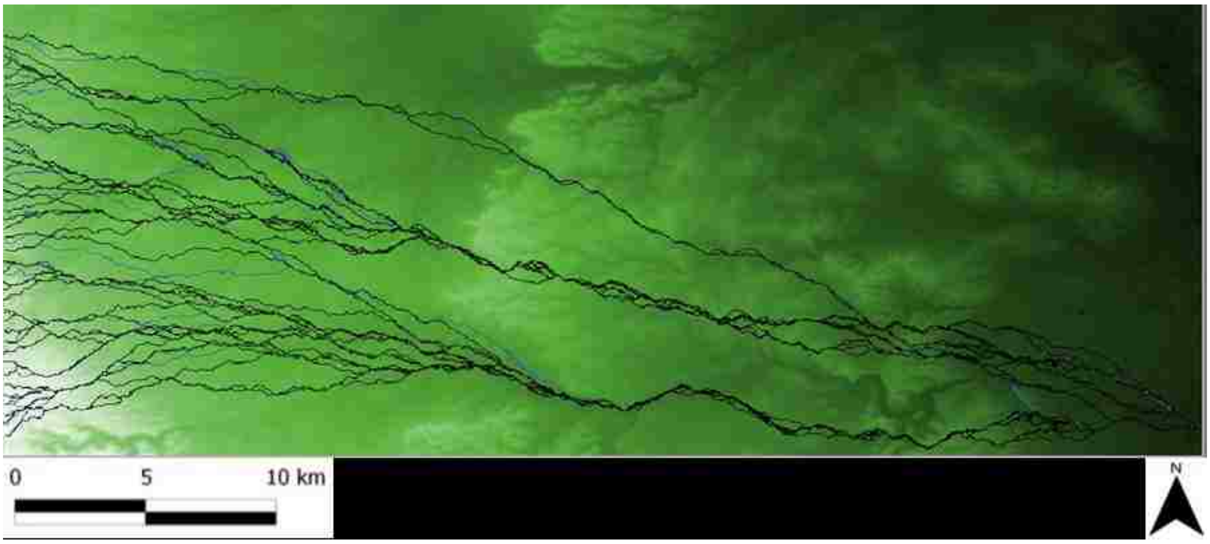


Figure 43: Paths with a set error rate in calculating costs. (Blue) normal cost path. (Black) plus or minus 50% error rate.

Other Error Rates

To undertake due diligence and explore as many representations of possible error rates in the cost path several other commands were used. One was a normal distribution of error rates from zero with different standard deviations. The results (Figure 44) are more significant for the normal distribution than the previous error rates. Even with a standard deviation of only 10%, there are significant swings in least cost paths routes compared to a perfect least cost path.

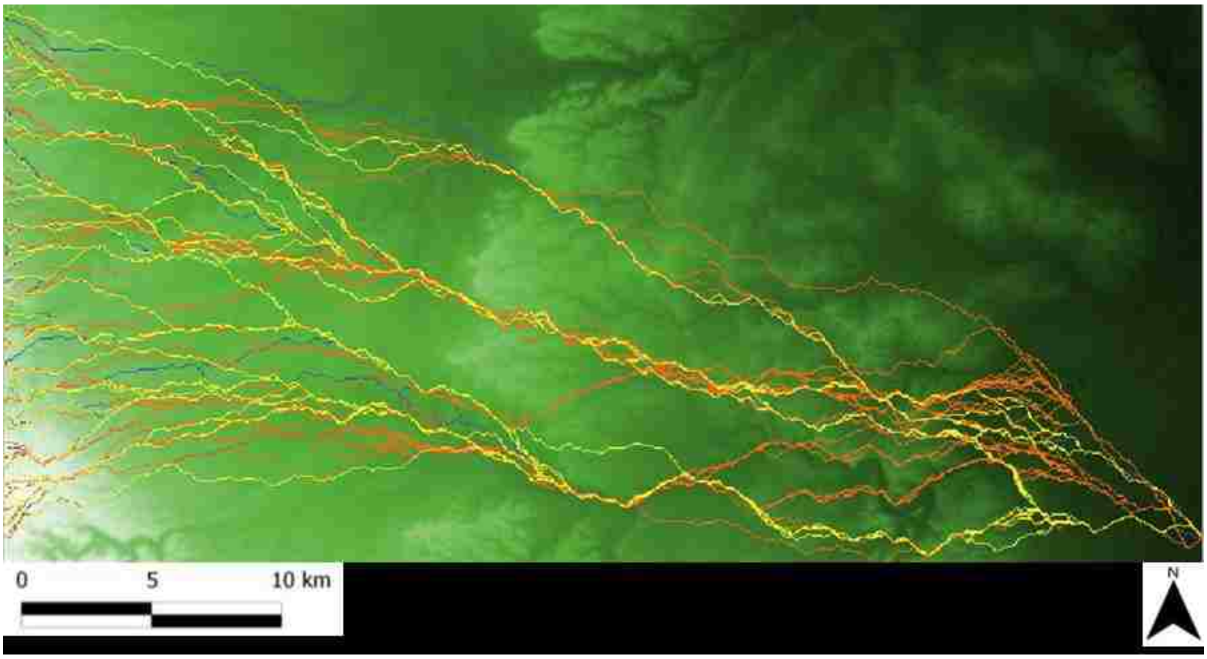


Figure 44: Paths with a normal distribution error rate in calculating costs, mean set at 0. (Blue) standard deviation of 0. (Yellow) standard deviation of 10%. (Orange) standard deviation of 50%.

Another option explored was to use the random exponential NetLogo command to set the error rate to an exponential curve. Significantly, this did not result in a great divergence in paths as seen with some of the other distributions of error rates.

Figure 45: Normal cost path (Black) compared to a random exponential error rate (Yellow).

All of these rates assume a random miscalculation in rates but it may be the case that the problems with estimating costs are patterned. For example, a person could overemphasise costs associated with steeper slopes. However, this does not need to be a mistake in assumptions; it could be that a person has trouble walking up or down steep slopes. If that were the situation, then a person may just have a preference for shallow slopes. To over-emphasise the costs associated with slopes, the square and cube of travel costs were taken into account. The results show that the paths are very different from those seen with the normal calculation of costs.

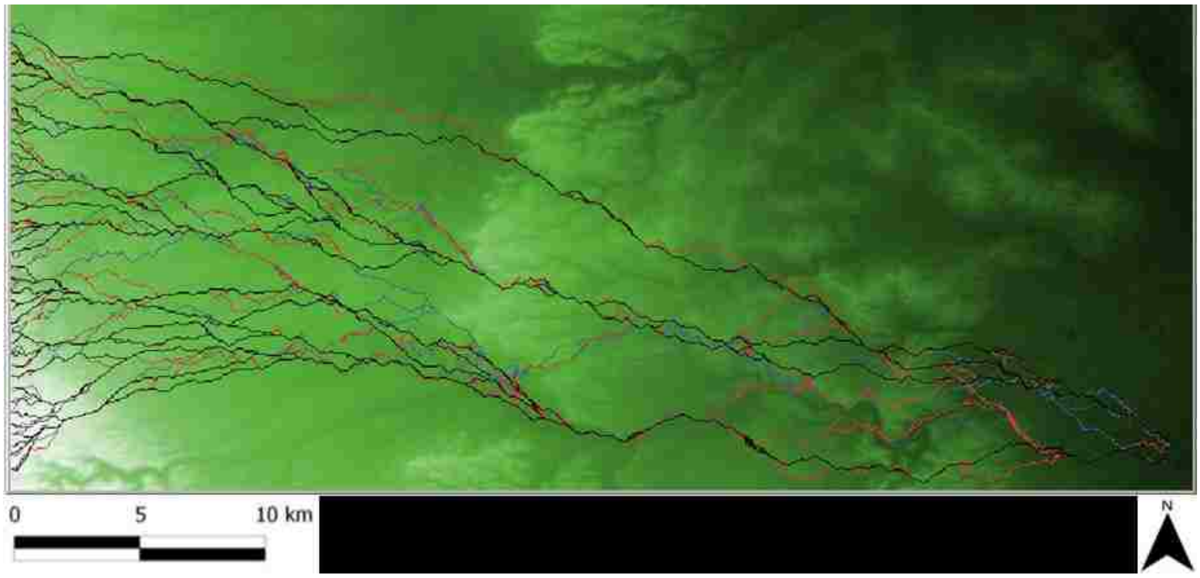


Figure 46: Cost paths using an exponential (Red) and cubed (Blue) calculation, to emphasise lower slopes, compared to the normal cost path (Black).

Discussion

All of this testing has shown that there can be significant changes in least cost paths due to a multitude of factors. These factors combined with other noted problems result in an extensive list of factors that alter the actual least cost path taken as explored with the agent based model:

- resolution of the datasets used
- the different programs used to create the paths, e.g. ArcGIS, GRASS GIS, NetLogo
- the directional application of forces: single, multiple, full 360 degrees
- dataset creation and interpretation by the programs used
- neighbour sampling: D8, D16, D24 etc.
- how costs are calculated: Tobler's Hiking, Backpacker's equation, Minetti's cost path equation etc.
- error in calculating/estimating travel cost by individuals
- knowledge of paths, i.e. viewshed
- edge effect: paths next to data set edges are not taken into account

In some of examples shown above the differences in paths were several kilometres apart. For site predictive modelling that is too great an error rate to be of any use in analysis. It could even be argued that this is too great an error rate for any other applications. Certainly, when comparing the past least cost path models with actual travel data there are significant differences in the results. Figure 47 shows least cost paths compared against known trails in Michigan: the results show that least cost paths can be tens and hundreds of miles/kilometres off from the paths they are supposed to model. One has to question how, if at all, least cost path analysis in its current form could be of any use to archaeology.

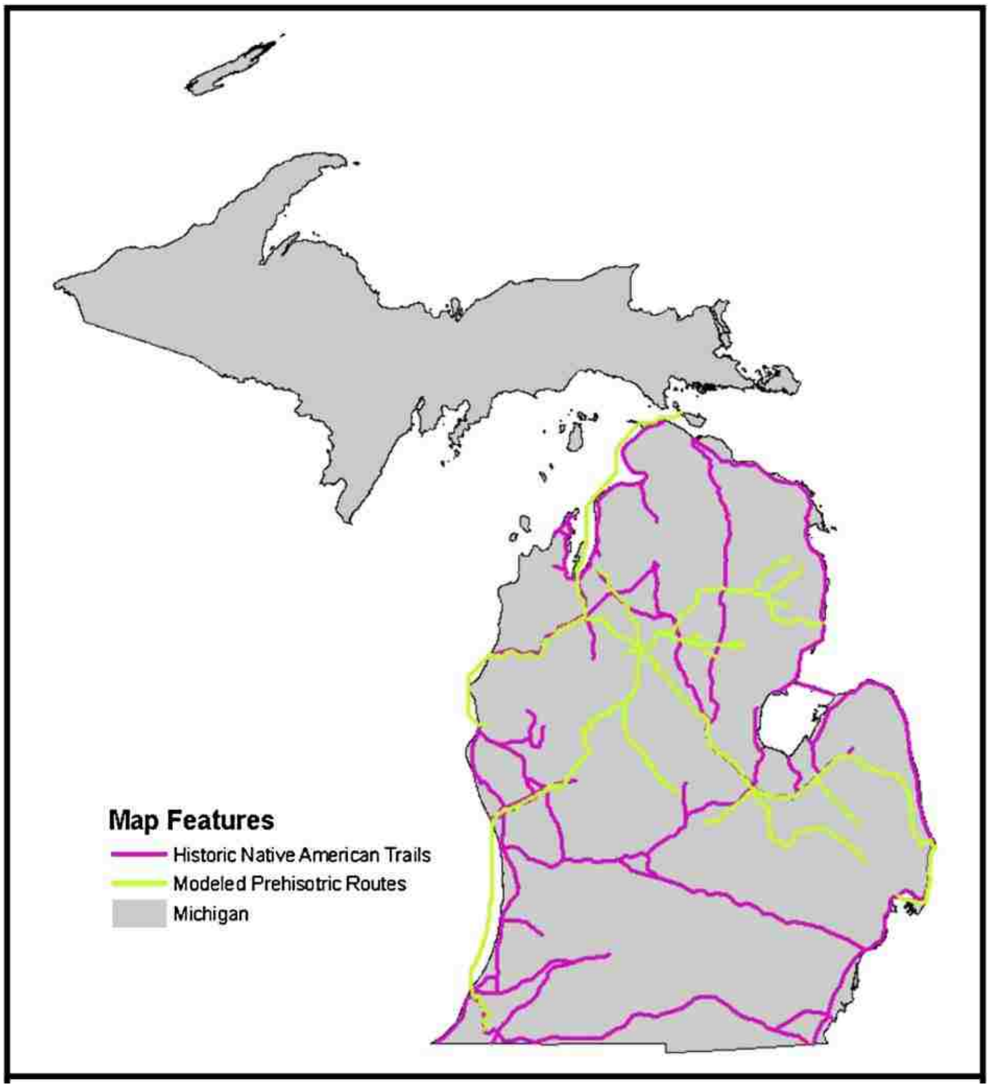


Figure 47: Historic trails and modelled routes compared in Michigan (Howey 2007 Figure 8 p. 1841).

Does Least Cost Path Even Matter?

The underlying concept of least cost path is to find the 'optimal' route. However, that assumes that an optimal route actually makes a difference in the results. Traditional least cost path analysis only allows one to check the 'optimal' route and not to investigate alternative pathways. However, with the agent based model it is possible to take into account alternative routes. This allows the possibility to examine alternative routes and the costs associated with them.

By limiting the possible range that the agent could take it was possible to create alternative routes. This was done with the optimisation code, which caused the route taken to be straighter. The results of these investigations show that the difference between the distance and time taken by the optimal route and other routes are minimal (Table 13). However, the differences between these routes are significant. For example, the furthest points between the 100 patch and 20 patch path were 5.25 km away while the difference in distance travelled in only .05 km and the time difference is a little more than two hours.

Limiting range-patches (distance km)	Total Distance Travelled (km)	Total Time (Hr.)	Difference From Optimal Total Distance (km)	Difference From Optimal Total Time	Colour
300 (18.52)	69.0211	19.4518	-	-	orange
100 (6.17)	68.2822	19.5608	0.73895	0.109 (6.54 min)	red
50 (3.09)	68.5565	19.9596	0.4646	0.5078 (30.47 min)	black
20 (1.23)	68.2398	21.8760	0.7813	2.4242 (2 hr. 25 min)	blue
10 (.62)	66.9127	23.7209	2.1084	4.2691 (4 hr. 15 min)	yellow
5 (.31)	65.4869	33.0143	3.5342	13.5625 (13 hr. 34 min)	pink

Table 13: Travel times and distances for different cost paths.

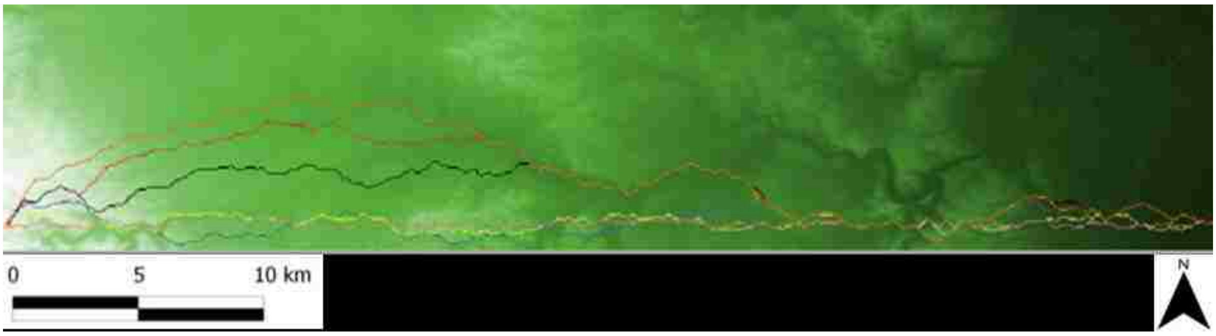


Figure 48: Different cost path routes for Table 13.

These results are similar because, for all the complex calculations taking into account direction of travel, aspect of the landscape and different cost calculations, the actual cumulative costs are similar. Figure 49 shows the spread of costs from a single point across the landscape at different scales. While there are areas where one can travel quicker than in others, the cumulative difference is minimal.

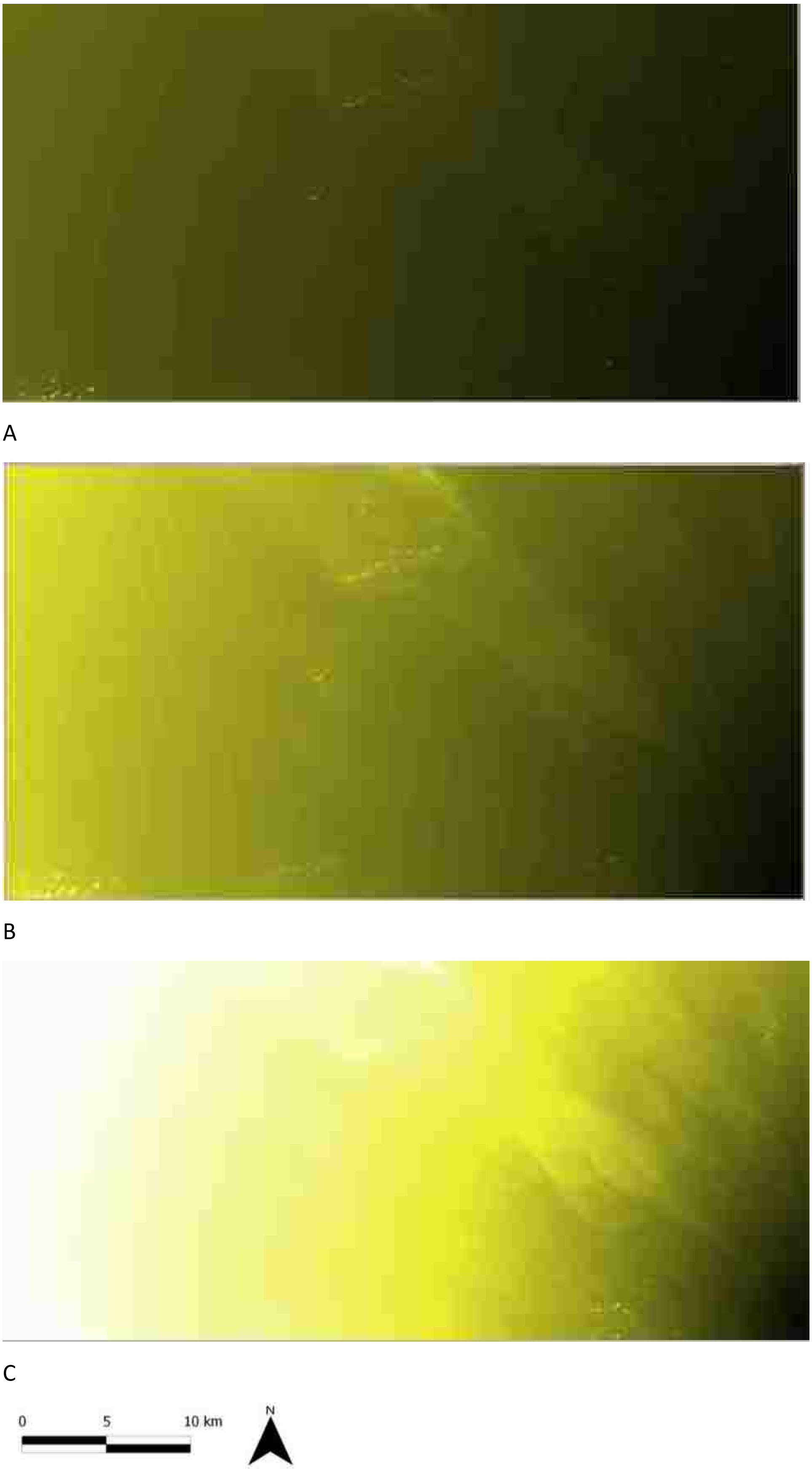


Figure 49: Cumulative cost path analysis at different scales. A: max 1500 B: max 700 C: max 300.

Time Difference

Cumulative totals blunt the effects of optimal routes. The yellow path in Figure 29 only takes 22% more time than the optimal route and yet deviates by more than 6 km in some areas. The blue path only takes 12.5% more time and still has a difference of over 6 km. This raises a question that has yet to be addressed in archaeological literature: how much did time actually matter to past people? Would half an hour have made a difference?

Costs in this example are measured in time, but these costs could easily be measured in something else, such as calories used. Yet, the difference is going to be very small and one has to question where people, past or present, are able to tell the difference in calories burned with any sort of accuracy. Would they be able to tell the difference between 10, 20, 50, 100, 1500 calories?

Regardless of how cost is calculated the same problem remains: least cost path analysis gives results in the form of a single detailed line when there are multiple paths, all similar in costs.

Corridor Analysis

In other projects, least cost paths have been used to create line density analysis of landscapes. This process involves looking at multiple least cost paths between various locations to find areas that receive a high and low number of least costs paths. Those areas that receive a higher density of paths could be labelled as travel corridors. These areas can then be compared against known sites to see correlations (Figure 50), a use that could be applied to a predictive model. One could use 'Morphometric Analysis' as a tool to find natural corridors in a landscape. Moreover, others have used tools like 'Corridor Analysis' in ArcGIS can, and has been used, to determine natural corridors (Figure 51) (Doneus et al. 2008).

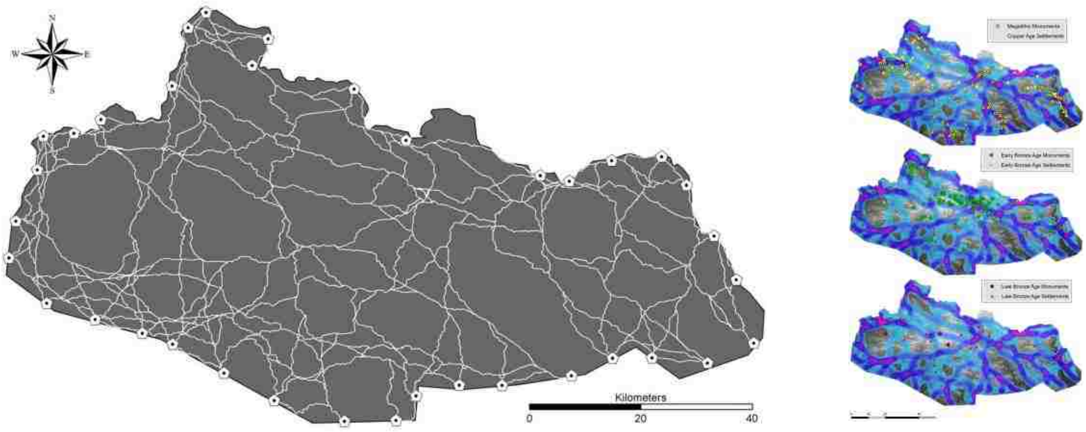


Figure 50: Least cost past used to find natural travel corridors (Murrieta-Flores 2012 p. 112-13 Figures 8 and 9).

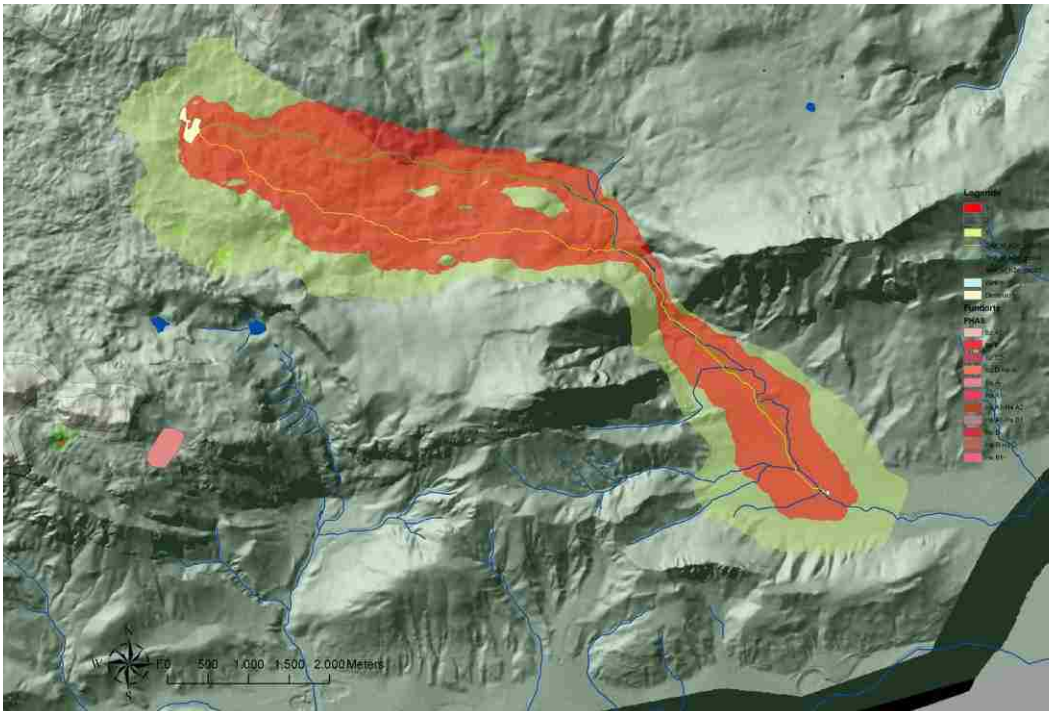


Figure 51: Example of Corridor Analysis in ArcGIS (Doneus et al. 2008 p. 5 Figure 4).

Corridors and Intensity of Travel

It was possible to apply a type of corridor analysis to the agent based model by looking at the intensities of areas travelled. The number of times the different paths intercepted were recorded creating a math of intensity of travel along the route (Figure 52). This was done with the standard model and none of the additional ones such as the viewsheds or imperfect knowledge. Figure 53 shows those paths in relation to the elevation model. This analysis did identify several potential main throughways across the landscape; areas that also correlated with known archaeological sites (**Confidential Appendix Figure 97**). But, there were also many sites that did not correlate at all to these corridors of travel.

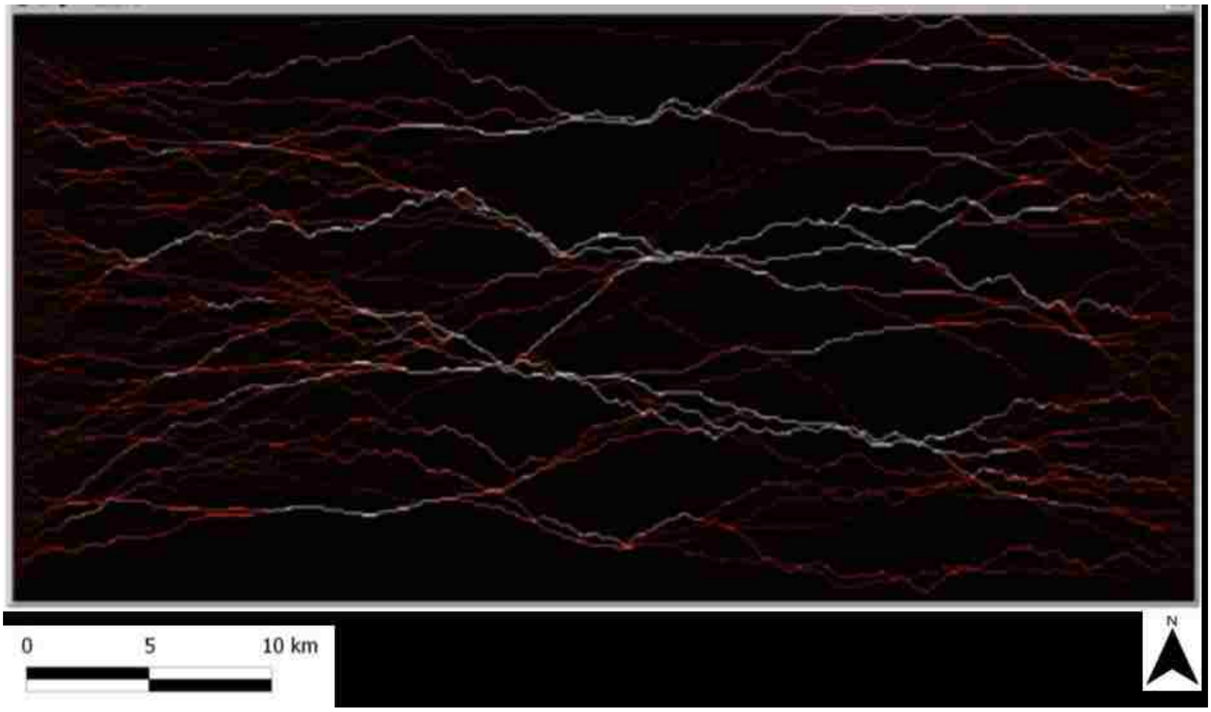


Figure 52: Frequency of path travelled for points five patches apart. White= higher frequency of travel.

Figure 53: Same map as Figure 52 but overlaid over an elevation map of the project area. Not to scale. White= higher frequency of travel.

Discussion

These investigations raise several key concerns about the use of least cost path analysis. On the technical level, there is concern that components of the agent based model may not accurately represent the real world through the processing of the data, i.e. convolution operators. The use of a Dijkstra algorithm causes distortions. However, that is a problem found with GIS programs and they have their own technical issues which do not make them a better tool. There are some potential technical issues that may invalidate the results.

On a theoretical level, this work has found serious flaws in least cost path analysis. Any deviation from an agent's perfect knowledge of the area leads to significantly different results. Incorporating a local viewshed and limiting knowledge, which is a closer representation of some situations, changes the paths taken. If a real person over or underestimates the effort involved in slopes, that can greatly change the outcomes.

The most damning indictment of least cost path analysis is that the optimal path makes very little difference to the results. Taking a less optimal path only increases the costs slightly, be it time or calories. One has to question if people, both past and present, would notice an extra six minutes in a 19.5 hour journey? Yet, that same journey takes them .7 km away from the optimal path. For predictive modelling or modelling travel in general, that is a significant difference.

At best, one ends up with corridors of travel. There are areas that the landscape funnels people into but not the neat lines seen with least cost path analysis. These corridors are limited in scope and subject to edge effect, not to mention all of the possible changes in paths laid out above. Moreover, these corridors are obvious and it would be questionable that a person looking at a topographic map would not have been able to spot and label them without the use of an agent based model.

This does not mean these results are not useful. We clearly have several areas with high potential for archaeological sites, assuming they result from travel to/from the mountains to/from the river. It terms of the hypothesis about the project area:

> 'We know that people came into the area… as travellers following a favoured route from the river to the uplands...' (Altschul, Hayden et al. 2005 p. 106)

The results of the modelling leave this as a possible contributing factor but not the main factor. This information alone is not enough to create a full predictive model but when combined with other datasets it can help modify the model. Doing such work will be discussed in the following chapters.

Overall, the results indicate that least cost path analysis has very little utility for predictive modelling or archaeology in general. Given the problems found here when testing human actions and the costs of less than optimal paths it seems as though least cost path may have limited application in the future.

Chapter 7: Water and Surviving

> 'We know that people came into the area, possibly in small, mobile groups that exploited locally available resources and then left … or possibly even as part-time agriculturalists establishing opportunistic fields at favourable locations to capture runoff' (Altschul, Hayden et al. 2005 p. 106)

After route analysis failed to account for most site locations, beyond several possible choke points that were associated with a few sites, the next pattern to be examined was one based on subsistence and living. It was suggested that the archaeological record reflects a pattern of subsistence in the project area. A hypothesis was tested and reviewed in this chapter.

Water in the Project Area

Chapter 6 outlined how the project area is and was an arid environment with little rainfall; a desert. Moreover, the studies of the paleoecology indicate that it has been this way for most of human occupation. Even if there was sufficient plant and animal life to support seasonal or long-term occupation in the subject area, without a constant and significant source of water habitation would have been impossible. Even limited foraging would have been constrained. Water was the foremost constraint on population exploitation of the project area and needed to be modelled to understand the possible foraging strategies employed.

Geology and Water

The geology and environment of the subject area limits where and how local populations could obtain water. For example, digging wells would be impossible in the majority of the subject area as the water table is below a restrictive layer (Table 14). A 'restrictive layer' is a nearly continuous layer that significantly impedes the movement of water and air through the soil or that restricts it. Examples are bedrock, cemented layers, dense layers and frozen layers (USDA 2012). It would not have been feasible for prehistoric people to reach the water table with hand-dug wells. Modern occupants have to dig wells over 100m through cemented layers (Hale 1945) and prehistoric occupants in the region are not known to have done this. This leaves running water from rivers, streams, springs and ponds as the primary source of water for the local populations.

Soil	Distance to Restrictive Layer	% of area	Soil	Distance to Restrictive Layer	% of area	Soil	Distance to Restrictive Layer	% of area
Aa	>200	0.13%	Ha	>200	0.01%	RM	>200	1.17%
At	84	0.01%	Hk	>200	0.02%	RPG	0	0.56%
AH	>200	0.01%	Ku	>200	0.01%	RTE	0	0.28%
Ah	>200	0.004%	LN	>200	0.18%	SG	48	0.21%
Ao	>40	0.02%	LT	>200*	8.97%	SM	48	0.07%
DP	>200	3.76%	MXC	>200	0.19%	TN	38	0.03%
DRG	36	0.04%	PD	>200	0.06%	TPE	25	0.70%
DYE	20	0.02%	Pe	>200	0.13%	UG	33	9.87%
EC	15	24.60%	PM	>200	0.70%	Uo	33	0.50%
EE	15	23.55%	RA	>200	2.45%	Up	14	0.02%
ER	15	8.45%	Rc	>200	0.13%	UR	33	2.32%
GA	>200	0.14%	Rd	>200	0.39%	Ut	30	0.04%
GC	>200	0.30%	RE	>200	9.34%	W	N/A	0.08%
GP	N/A	0.004%	RG	>200	0.55%			
Total Land > 200			28.73%		Total Land < 200			71.27%

Table 14: Distance in cm to a restrictive layer. *the distance for LT is listed as >200cm; however, this layer is itself a restrictive layer of limestone.

Several historical place names have the term ‘spring’ in their names (Table 15). Indian Big Spring was known to have occasional discharges (Motts 1968, Engineer 2004) but information on this spring was limited. Satellite photos on Google Earth indicate that the area around McKittrick Spring had significantly more greenery than the surrounding landscape, but it is unknown if this is connected to a natural spring. Further research showed that the local spring flow is determined by the height of the water table. This in turn is influenced by the rainfall in the project area (Cox 1967). Above average rainfall results in the slow trickle of water from the higher elevations into the lower gullies, which raises the water tables in those areas and results in springs forming. This explained the seasonality of Indian Big Springs and the fact that known springs are only found in the lower elevations.

Besides Indian Big Springs the majority of the springs listed for this project area did not have any record of use, except for the springs next to the Pecos River (Carlsbad Spring Number 16). Moreover, research on geology and hydrology indicated that only certain locations in the project area were suitable for springs. Part of the area lies on top of a reef escarpment that is not conductive to underground water flows (Brooke, Dawson et al. 1997, Goodbar and Rice-Snow 2012). This reef drops off at the Pecos River Valley at which point it is possible to have springs, and several do exist next to the Pecos River.

Name	Latitude	Longitude
Indian Big Spring	32.4553907	-104.4888505
Lancaster Spring (Lassiter Spring)	32.4117469	-104.3220966
Little Walt Spring	32.4165295	-104.4709678
McKittrick Spring (Mc Kitric Spring)	32.4136817	-104.3528780
McGruder Spring	32.4017753	-104.3331474
Walt Spring	32.4325679	-104.4811743
Yellow Jacket Spring	32.2878934	-104.3643942
Carlsbad Spring Number 16	32.4445617	-104.2643951

Table 15: List of Spring place names in the project area.

Figure 54: Location of Spring place names on Google Maps.

Past Predictive Models and Surface Runoff

Seasonality of springs meant that runoff from rain was the only possible consistent source of water for the majority of the project area. That was the next model to be considered.

Modelling surface runoff was not new in predictive modelling. Dalla Bona (2000) has stated that water is one of the key datasets employed by almost every site predictive model. Some predictive models have even been created solely based on sites' proximity to water (Altschul et al. 2004). Water played a key role in some of the previous predictive models created for the subject area (Altschul et al. 2005). There was a correlation between the areas predicted as having high probability of containing sites and proximity to the water features created on the map (Figure 55, Figure 56).

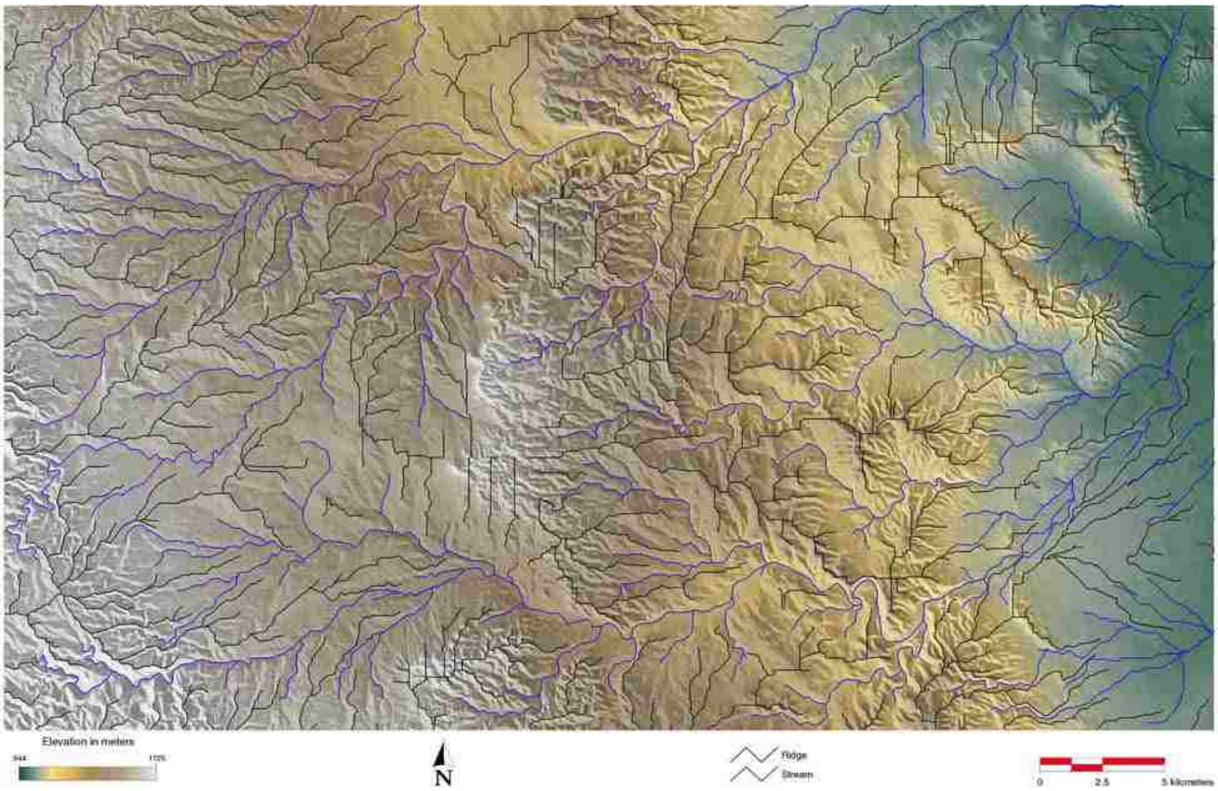


Figure 55: Dataset of drainages and ridges in the Azotea Mesa study area from the PUMP III project. (Altschul et al. 2005 p. 79 Figure 6.3)

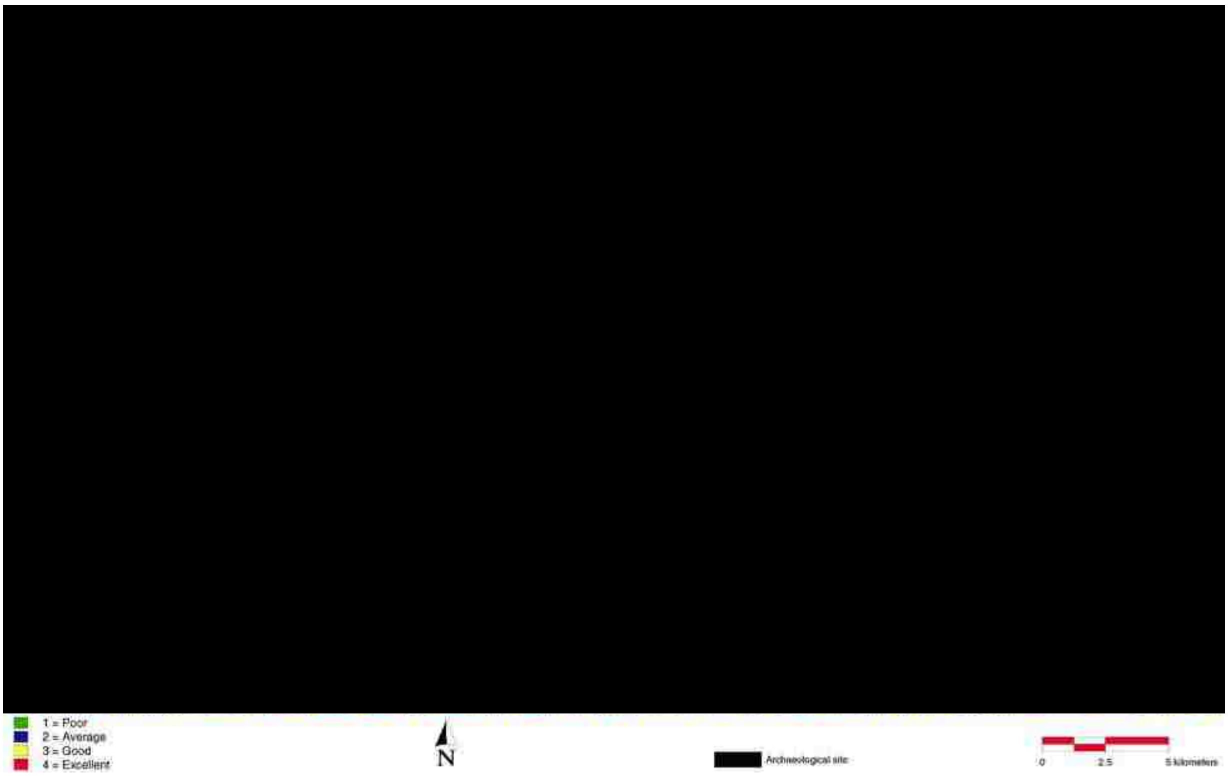


Figure 56: Results of PUMP III Regression Model. (Altschul et al. 2005 p.95 Figure 6.12) overlaid with an image (black lines) of the water system traced from (Altschul et al. 2005 p. 79 Figure 6.3).

Blue Line Features and Data Issues

A problem raised with predictive modelling is that 'blue line features' on maps may or may not actually contain water (Ebert 2000), an issue relevant to this project. Considering that, on average for the last hundred years, the city of Carlsbad had only received annual rainfall of a little less than 13 inches and the annual evaporation rate is 109.48 inches, it seemed unlikely that any of the waterways used in the previous project would actually contain water. Moreover, hydrology research of this and the surrounding areas indicated that, aside from some springs and the Pecos River, most of the channels listed are the result of flash floods (Hale 1945, Cox 1967) not perennial water ways. This raised the question if currently available datasets are in fact an accurate representation of the water systems.

The water systems used in the PUMP III model were not created from maps, but through the use of a GIS hydrology program. The PUMP III model used this data to represent distance to water, assuming that each and every one of the water features represented a usable continual source of water in the project area. Yet, this dataset was not checked against any other data to determine if it accurately represents the real distribution of water in the project area.

To confirm these suspicions that the GIS datasets did not represent the actual water system this project looked at the raw water flow data from the USGS. The USGS keeps records on water flows throughout the United States including areas in and adjacent to the subject area. There were a limited number of measurement stations in the drainages of the subject area but they did provide a picture of the information about the volume and frequency of water flowing out of the subject area (Table 16). These numbers are limited but they illustrate the fact that the vast majority of the precipitation in the study area is not converted into a consistent water flow. The images of long flowing streams crossing the subject area are clearly not reflected in any water data.

Test Point	Drainage Area	Period of Record	Annual Runoff (ac-ft) 2010	Annual Runoff (ac-ft) 1964-2010	Days with flow 2010	Avg. Annual Runoff Against Avg. Rainfall (assuming avg. 12 inches of rain)
08401900 Rocky Arroyo at Highway Bridge, Near Carlsbad, NM	285 mi^2, approximately.	November 1963 to current year.	939	4,040	5	2.1%
08405050 Last Chance Canyon, Near Carlsbad Caverns, NM	2 mi^2.	Water years 1959 to 1996, 2005 to current year	Not listed	Not listed	Not listed	
08405100 Mosley Canyon, Near White City, NM	14.6 mi^2.	Water year 1959 to current year.	Not listed	Not listed	Not listed	
08405105 Dark Canyon Draw Near White City, NM	327 mi^2, approximately.	February 2002 to current year.	2280	2560	7	1.2%
08405150 Dark Canyon Draw at Carlsbad, NM	451 mi^2, approximately.	January 1973 to current year.	260	3290 (+ 2100 for irrigation) 5390	1	1.9%

Table 16: USGS water data gathering stations in and next to the subject area. Full table with locations in Appendix D. Mi- miles, ac-ft- Acre-feet

Project Implications

In terms of the agent based model, this information had implications for how the model was constructed because it presented different behaviour issues. Limited water flows would have severely limited the ability of local populations to sustain themselves for any significant period of time. If water supplies had to be transported into the project area, then the time groups could spend hunting and gathering resources would be significantly limited. Limited water would have meant that the few springs that did exist in the area would not have had water either. On the other hand, the pooling of water into small reservoirs, depending on how deep they were, could have provided year-round water supplies. If such pools existed, they would have attracted people and thus the placement of archaeological sites.

Faced with these possibilities, this project set out to determine the conditions needed for runoff that will lead to pooling, water flows, or neither in the project area. The idea was to model these conditions to create an accurate picture of the natural flows of the water. These models were then calibrated with the actual climate conditions, current and historic, and the model adjusted accordingly to ensure accurate results. The knowledge it provided allowed me to make a determination on both the suitability of the areas for long-term occupation and the appropriate locations for such occupation.

Model Creation

Creating a hydrology model was more complex than making a travel model because a variety of variables needed to be taken into account. For one, precipitation can take different routes to become runoff (Garen and Moore 2005) that need to be modelled. There was Infiltration Excess Overland Flow, which occurs when the rainfall concentration exceeds the infiltration capacity of the soil, and the rainfall that cannot be absorbed by the soil runs down the land surface of the hillslope. It can easily be confused with Saturation Excess Overland Flow events, which occurs when the soil is already saturated before a rainfall and the water immediately runs off. Water from the saturated soils could also exfiltrate, called a return flow.

Saturation Excess Overland Flow typically takes place at the base of hillslopes, where soil moisture is high because of downslope movement of subsurface water. This process was first identified in the late 1960s (Dunne 1978) and is now well known among the hydrologic research community. However, it appears to not be well understood by many practising engineers and hydrologists (Garen and Moore 2005). This is the dominant streamflow generating process during most storms of ordinary intensity. Infiltration Excess Overland Flow typically occurs in high intensity storms while Saturation Excess Overland Flow can occur in low intensity storms.

Other routes that water could take were:

- **Shallow Subsurface Flow -** In some areas, water can flow downslope shallowly within the soil quickly enough to be considered part of the storm flow. This is often enhanced by the presence of macropores caused by earthworms, burrowing animals, tree roots, etc. Basically, a subsurface runoff flow.
- **Ground Water Flow/Base Flow -** This is the water that exfiltrates from the aquifer to the stream.
- **Direct Precipitation Onto Stream Surface -** Water is added directly to the flow without interacting with soil.
- **Percolation -** Water moves down through the soil into the aquifer. Though because this model was focused on runoff, events like percolation were not modelled.

Model Formula

To convert rainfall to runoff one uses an expression of conservation of mass where runoff is determined by the rainfall minus abstractive losses:

$$\mathbf{Q = P - L}$$

Equation 17: Conservation of mass for runoff. Q = runoff; P = rainfall; and L = abstractive losses.

These abstractive losses fall into five categories (Hawkins and Ponce 1996):

1. Interception storage by vegetation foliage, stems, litter or by cultural features e.g. roofs on houses, water storage facilities, etc.
2. Surface storage in ponds, puddles and other small temporary storage locations.
3. Infiltration to the subsurface to feed and replenish soil moisture, interflow, and ground-water flow.
4. Evaporation from water bodies such as lakes, reservoirs, streams and rivers as well as from moisture on the bare ground.
5. Evapotranspiration - the sum of evaporation and plant transpiration of water from the Earth's surface to its atmosphere.

Runoff Curve Number

The most common way to calculate these losses in the United States, and in other locations across the world, is through a Runoff Curve Number (RCN). The curve number procedure was developed in the 1950s by the Soil Conservation Service (SCS) as a simple procedure for estimating streamflow caused by rain storms. It should be noted that it only covers Direct Precipitation Onto Stream Surface, Infiltration Excess Overland Flow, and Saturation Excess Overland Flow and not any of the other flows mentioned. The primary documentation for the procedure is USDA-SCS (1972). The number takes into account vegetation and soil characteristics that can be looked up using tables. It does not take into account long-term losses caused by evaporation or evapotranspiration.

The advantages listed by Ponce and Hawkins (1996) for using a curve number are:

1. It is a simple, predictable, and stable conceptual method for the estimation of direct runoff depth based on storm rainfall depth, supported by empirical data.
2. It relies on only one parameter, the runoff curve number, which varies as a function of four major runoff-producing watershed properties:
    a. hydrologic soil group: A, B, C, and D
    b. land use and treatment classes
    c. hydrologic surface conditions of native pasture: poor, fair, good
    d. antecedent moisture condition: 1, 2, 3
3. It is the only agency methodology that features readily grasped and reasonably well-documented environmental inputs.
4. It is a well-established method, having been widely accepted for use in the United States and other countries.

The disadvantages listed were:

1. The method was originally developed using regional data, mostly from the Midwestern United States, and has since been extended by way of practice to the rest of the world. Some caution is recommended for its use in other geographic and climate regions.
2. In some instances, particularly for the lower curve numbers and/or rainfall depths, the method may be very sensitive to curve number and antecedent conditions. This is not necessarily a weak point, since it may be a reflection of the natural variability. There is , however, a lack of clear guidance on how to vary antecedent conditions.
3. The method performs poorly in forest sites and is best with negligible base flow found in arid and semiarid regions.
4. The use of curve numbers for areas greater than 100 sq mi or 250 sq km is unreliable.
5. The method fixes the initial abstraction ration as 0.2 but this can vary between climates.

Several of the disadvantages of this method were found to be applicable to the project area, i.e. different climate, low rainfall. However, as pointed out by Ponce and Hawkins (1996), replacement formulas presented up until the 1990s could not replace the curve number as a superior method. Garen and Moore (Garen and Moore 2005) have listed some additional methods but as of yet there is no clear replacement model for the curve number and as such was used in some of the models.

The SCS runoff equation that uses the runoff curve number is:

$$\mathbf{Q = (P - Ia)^2 / (P - Ia) + S}$$

Equation 18: SCS runoff equation: Q = runoff (in); P = rainfall (in); S = potential maximum retention after runoff begins (in); Ia = initial abstraction (in).

Ia can vary between values of 0.0 and 0.3 but has been generally found to be 0.2. Thus the equation can be converted to:

$$\mathbf{Q = (P - 0.2S)^2 / (P + 0.8S)}$$

Equation 19: Converted SCS runoff equation. Where S = (1000 / curve number) – 10

The formula results in the follow outputs:

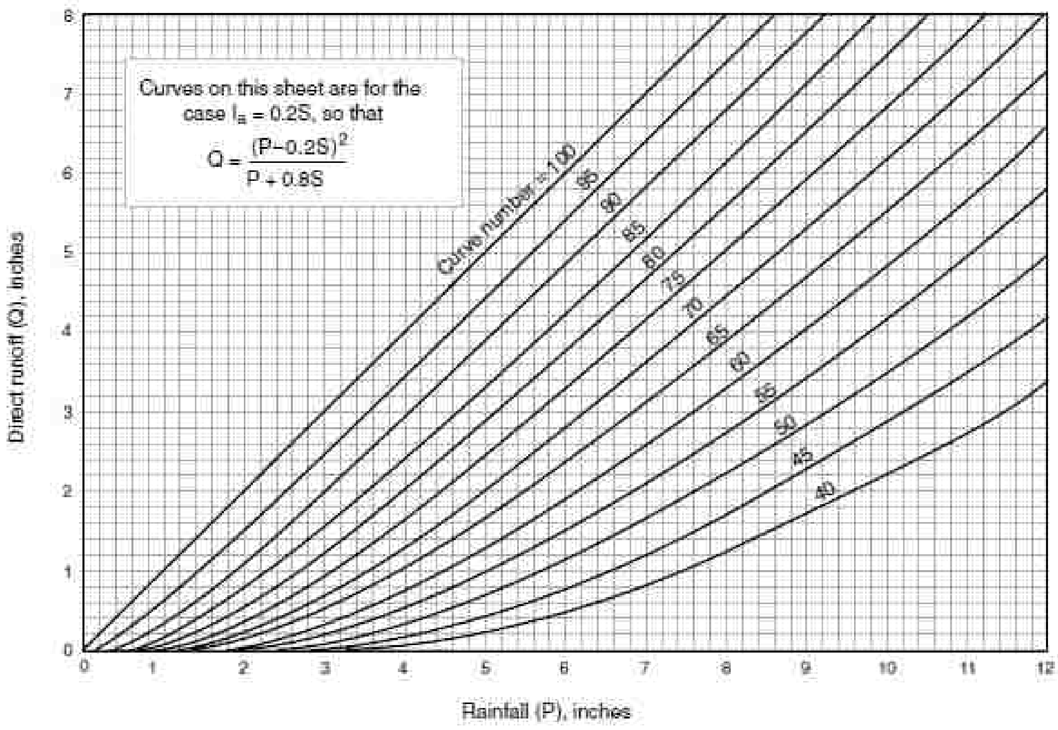


Figure 57: Rainfall to direct runoff chart using the runoff curve number (USDA 1986 p. 14 Figure 2-1).

Soil Groups

The curve number is primarily determined by soil groups, which are classified as A, B, C, D, and vegetation coverage. The soil classifications are:

> **'Group A:** Soils having high infiltration rates even when thoroughly wetted and a high rate of water transmission. Examples are deep, well to excessively drained sands or gravels.
>
> **Group B:** Soils having moderate infiltration rates when thoroughly wetted and a moderate rate of water transmission. Examples are moderately deep to deep, moderately well to well drained soils with moderately fine to moderately coarse textures.
>
> **Group C:** Soils having low infiltration rates when thoroughly wetted and a low rate of water transmission. Examples are soils with a layer that impedes the downward movement of water or soils of moderately fine to fine texture.
>
> **Group D:** Soils having very low infiltration rates when thoroughly wetted and a very low rate of water transmission. Examples are clay soils with a high swelling potential, soils with a permanently high water table, soils with a clay pan or clay layer.'
>
> (USDA 1986)

Vegetation works in combination with soil types to help determine the final runoff curve number. Several tables have been created to make determining the correct curve number from the combination of these two variables. The table for arid environments was used for this project (Table 17).

Cover description		Hydrologic soil group			
Cover type	Hydrologic condition	A	B	C	D
Herbaceous—mixture of grass, weeds, and low-growing brush, with brush the minor element.	Poor		80	87	93
	Fair		71	81	80
	Good		62	74	85
Oak-aspen—mountain brush mixture of oak brush, aspen, mountain mahogany, bitter brush, maple, and other brush.	Poor		66	74	79
	Fair		48	57	63
	Good		30	41	48
Pinyon-juniper—pinyon, juniper, or both; grass understory.	Poor		75	85	89
	Fair		58	73	80
	Good		41	61	71
Sagebrush with grass understory.	Poor		67	80	85
	Fair		51	63	70
	Good		35	47	55
Desert shrub—major plants include saltbush, greasewood, creosotebush, blackbrush, bursage, palo verde, mesquite, and cactus.	Poor	63	77	85	88
	Fair	55	72	82	86
	Good	49	68	79	84

Table 17: Vegetation conversion runoff curve numbers for arid environments. Poor: <30% ground cover (litter, grass, and brush overstory). Fair: 30 to 70% ground cover. Good: > 70% ground cover (USDA 1986).

As the curve number also models Saturation Excess Overland Flow one needs to take into account soil saturation. The SCS modified formula had three categories of soil moisture levels known as Antecedent Moisture Condition **(AMC)** Classes. AMC classes were used to adjust the runoff curve numbers (Table 39) (USDA 1986).

- AMC **I:** The soils in the drainage basin are practically dry (i.e. the soil moisture content is at wilting point).
- AMC **II:** Average condition.
- AMC **III**: The soils in the drainage basins are practically saturated from antecedent rainfalls.

Agent Based Model of Water

Because the runoff curve numbers were not meant to be temporal calculations but an estimation of the runoff that will occur from a rain event an agent based model needed to be created to determine what happened to the runoff once additional factors such as evaporation or evapotranspiration were taken into account. These are the factors that determine if pools of water existed in the project area or seasonal streams, answering the questions raised about site locations and their association with water features. Again, the decision was made to use the NetLogo program to model the full water system of the project area. GIS programs were used to manipulate some of the landscape datasets.

Soil Data

The soil quality the NRCS datasets mention in Chapter 6 were used in this model. The data had been collected in two surveys, one that captured the majority of Eddy County and was completed in 1966 (Anderson, Chugg et al. 1971). This data was originally at 1:20,000 and 1:31680 scales. It was gathered by first sending physical inspectors into the subject area to survey soil types. Then the soil types were mapped using aerial photography. The NRCS claims that field investigations and data collection are carried out in sufficient detail to identify, accurately and consistently, areas of about six acres. An evaluation was made of the soil survey in 1996 which determined that the soil map was accurate. Several minor updates were made and that updated version was digitalised and used in this project.

The Eddy County survey did not capture all of the project area, with a small portion of the south west corner missing. Data from a second soil survey, covering the missing portion of Eddy County and parts of Otero and Chavez counties, was included in the dataset (Derr 1981). This survey was finished in 1976. The scale of that report was at 1:24,000 and 1:63,360. This data was evaluated in 2003 and, like the other report, minor changes were made to some of the classifications before it was digitalised.

This data also contained information on water-carrying capacity, absorption rates and hydrology classifications used for calculating the runoff curve number for the different soils (Table 18).

Carrying Capacity (% of unit of measurement)	Percentage of Project Area	Soil Absorption Rates	Percentage of Project Area	Soil Hydrology Classification	Percentage of Project Area
0.06	23.55%	0.00 to 0.06 in/hr	3.67%	A	3.76%
0.07	0.02%	0.01 to 0.60 in/hr	12.75%	B	15.46%
0.08	0.74%	0.06 to 2.00 in/hr	56.60%	C	13.22%
0.09	24.67%	0.20 to 0.60 in/hr	1.21%	D	67.49%
0.1	10.37%	0.20 to 2.00 in/hr	0.44%	N/A	0.08%
0.11	0.21%	0.57 to 1.98 in/hr	0.01%		
0.12	0.13%	0.60 to 2.00 in/hr	12.33%		
0.13	2.24%	2.00 to 6.00 in/hr	3.95%		
0.14	2.45%	bedrock	8.97%		
0.15	0.33%	N/A	0.08%		
0.16	0.21%				
0.17	0.18%				
0.18	0.01%				
0.19	0.03%				
0.2	0.83%				
.07-0	0.84%				
.1-.09	0.01%				
.14-.09	8.45%				
.14-.11	2.32%				
.14-.13	9.40%				
.15-.11	0.07%				
.2-.12	3.76%				
N/A	9.19%				

Table 18: Characteristics of the soil on the project area from NCRS data. Detailed data in Appendix D.

Looking at the data for water carrying capacity we see that about two-thirds of the project area has none to very low carrying capacity and thus high potential for runoff. Also, there were low levels of absorption rates and lots of category D ratings, which also indicate a high potential for runoff.

Evaporation

Evaporation, like infiltration, is a complex system that has many components that influence it (Ritzema 1994), such as:

- dryness of the surrounding air
- exposure to wind
- composition of the air itself
- pressure of the surrounding atmosphere
- temperature of the water
- temperature of the air

Evaporation rates gathered from the Western Regional Climate Center showed that the project area has an annual evaporation rate of 109.48 inches and hourly rate of .012 inches. Adjusted for season variation the hourly evaporation rates ranged from 0.00 inches in the winter to .022 inches in the summer (Table 19). Compared against the estimates for soil absorption rates evaporation would only play a minor part in water loss. As such, it was decided to use only general hourly rates instead of trying to model complex systems like the dryness of the air and exposure to wind.

Location	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec	Year
Monthly	4.65	0.00	8.62	11.77	14.61	15.46	14.19	12.22	9.88	7.97	5.77	4.34	109.48
Daily*	.150	0.00	.278	.392	.471	.515	.458	.394	.329	.257	.192	.140	.300
Hourly*	.006	0.00	.012	.016	.019	.022	.019	.016	.014	.011	.008	.006	.012

Table 19: Evaporation rates from Brantley Dam 1987-2005. * calculated from the monthly averages.

Evapotranspiration

Evapotranspiration is a term used to describe the sum of evaporation and plant transpiration from the Earth's land surface to atmosphere. There are several formulae for determining evapotranspiration rates such as the Turc (1954), Jensen-Haise (1963), Blaney-Criddle (1950) and the Penman-Monteith (Monteith 1965) (Table 20). Each formula has its own data requirements and while some of the data required is available for the project area, in most cases it is not available. This forces the project to use a coefficient of .7 multiplied by the evaporation rate to set the evapotranspiration rate. This ratio has been found to be an accurate approximation of evapotranspiration rates (Linacre 1994).

Formula Name	Equation	Information
Turc's Formula (Turc 1954)	$ET_p = (P + 80) / \sqrt{}(1 + (P + 45 / L^{Tc})^2)$	ET_p = 10-day potential evapotranspiration (mm); P = 10-day precipitation (mm); L^{Tc} = evaporative demand of the atmosphere, calculated as $L^{Tc} = (T_a + 2) \sqrt{} R_s / 11.1$ were T_a = average air temperature at 2m and R_s = incoming short-wave radiation (W/M^2)
Jensen-Haise formula (Haise and Jensen 1963)	$ET_p = (0.025T_a + 0.08) R_s / 28.6$	ET_p = potential evapotranspiration (mm); T_a = average air temperature at 2m; R_s = incoming short-wave radiation (W/M^2)
Blaney-Criddle formula (Blaney and Criddle 1950)	$ET_p = k\ p\ (0.457T_{am} + 8.13)\ (0.031T_{aa} + 0.24)$	ET_p = monthly potential evapotranspiration (mm); k = crop coefficient; p = monthly percentage of annual daylight hours; T_{am} = monthly average air temperature; T_{aa} = annual average air temperature.
Penman-Monteith equation (Monteith 1965)	$ET = \Delta R_n + P_aC_p (\delta e) g_a / (\Delta + \gamma (1 + g_a/g_s)) \lambda_v$	ET_p = monthly potential evapotranspiration (mm) λ_v = latent heat of vapourisation (J/g); E = Mass water evapotranspiration rate (gsm^2); Δ = Rate of change of saturation specific humidity with air temperature (kPa); R_n = Net irradiance (Wm^2); C_p = Specific heat capacity of air (J/kg K); ρ_a = dry air density (kg/m^3); δe = vapour pressure deficit, or specific humidity (Pa); g_a = Conductivity of air, atmospheric conductance (mm/s); g_s = Conductivity of stoma, surface conductance (mm/1); γ = Psychrometric constant ($\gamma \approx$ 66 kPa).

Table 20: Equations for determining evapotranspiration rates.

Historical Data

Historical data was needed to both calibrate the model and compare the results against. There was over 100 years of historical data for parts of the project area. While there were fluctuations in weather as noted in the paleo-climate research for the general area the climate has stayed roughly the same since 5000 BC (see Chapter 6 discussion of paleoenvironment). It was thus possible to use modern climate data as a proxy for past weather events

The data for climate in the project area was obtained from the Western Regional Climate Center (WRCC), one of six regional climate centres in the United States (Center 2012). WRCC holds a variety of datasets on weather and climate in the Western United States. The specific datasets used were historical information on precipitation for stations in and around the project area but other datasets were also investigated such as temperature, snowfall, etc.

The historical data shows that precipitation occurs most frequently and in the highest concentrations in the middle to late summer months of July, August, and September (Table 21). A period locally referred to as the 'Monsoon Season'. August is the only month never to go without a precipitation event since data has been kept, though the data from the very early years is questionable. During the summer the primary source of rainfall is moist, warm air that pushes inland from the Gulf of Mexico. This moist air, combined with surface solar heating, results in localised afternoon and evening thunderstorms in the subject area. In the winter it is Pacific storm systems moving in from the west that provide the majority of the moisture. A problem is that the Guadalupe Mountains tend to block many of these systems from reaching the Azotea Mesa area (Sebastian, Altschul et al. 2005). No climate data was used in the PUMP III site predictive model so there is no conflicting datasets.

From Year=1900 To Year=2012											
	Precipitation										
	Mean (in.)	High (in.)	Year	Low (in.)	Year	1 Day Max. (in.)		>= 0.01 in.	>= 0.10 in.	>= 0.50 in.	>= 1.00 in.
								# Days			
January	0.4	2.31	1949	0	1912	0.79	1980	3	1	0	0
February	0.44	2.26	1997	0	1900	1.25	1997	2	1	0	0
March	0.48	4.39	1919	0	1903	2.41	1919	2	1	0	0
April	0.65	5.04	1915	0	1902	2.86	2004	2	1	0	0
May	1.19	12.28	1941	0	1903	3.41	1959	4	2	1	0
June	1.49	6.24	1948	0	1928	3.8	1972	4	3	1	0
July	1.86	10.5	1902	0	1903	3.8	1902	5	3	1	0
August	1.79	7.7	1984	0.01	1938	5.12	1916	5	3	1	0
September	2.14	12.27	1980	0	1907	4.6	1980	5	3	1	1
October	1.34	8.08	1907	0	1903	4.3	1945	4	2	1	0
November	0.58	4.58	2004	0	1915	2	2000	3	1	0	0
December	0.51	3.79	1991	0	1903	1.18	1986	3	1	0	0
Annual	12.87	33.94	1941	2.95	1924	5.12	1916	42	25	8	3
Winter	1.35	6.16	1992	0	1934	1.25	1997	7	4	1	0
Spring	2.32	17.99	1941	0	2011	3.41	1959	8	5	1	1
Summer	5.14	18.06	1902	0.74	1924	5.12	1916	15	9	3	1
Fall	4.05	16.01	1974	0.27	1951	4.6	1980	12	7	3	1

Table 21: Carlsbad historical climate data. For monthly and annual means, thresholds, and sums: months with five or more missing days are not considered, Years with one or more missing months are not considered, Seasons are climatological not calendar seasons (Winter = Dec., Jan., and Feb.; Spring = Mar., Apr., and May; Summer = Jun., Jul., and Aug.; Fall = Sep., Oct., and Nov.) (Center 2012).

Caution was taken when applying these numbers to the model as the Guadalupe Mountains received a higher annual precipitation then the surrounding areas (Figure 59). Unfortunately, the WRCC data stations were all located in the eastern portion of the project area around the city of Carlsbad (Figure 58). Luckily, there are several stations located next to the study area that served as proxies. On average the lower elevation stations have roughly (+ or – half an inch) the same average rainfall. The stations near the Guadalupe Mountains indicate a slightly higher amount of rain, around an inch or two each year (Table 22).

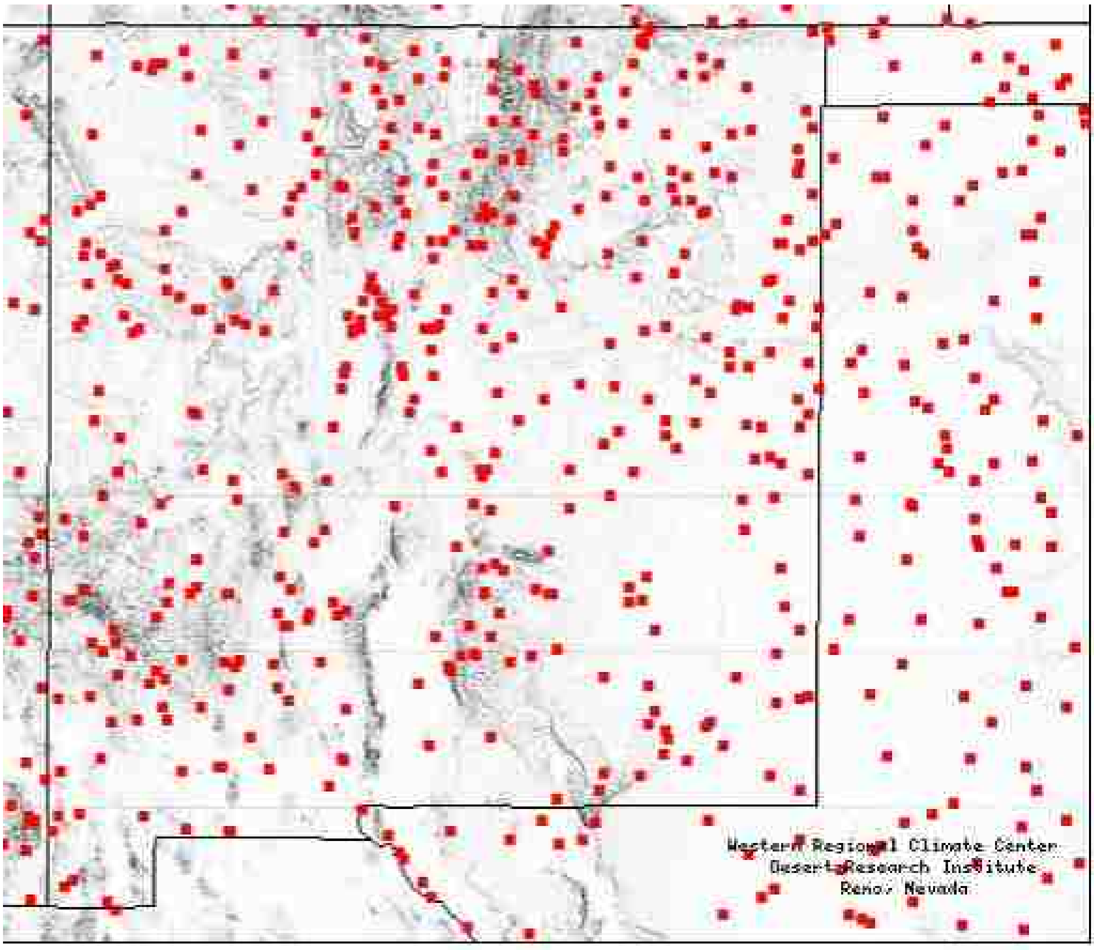


Figure 58: Data collection stations for the Western Regional Climate Center (Center 2012).

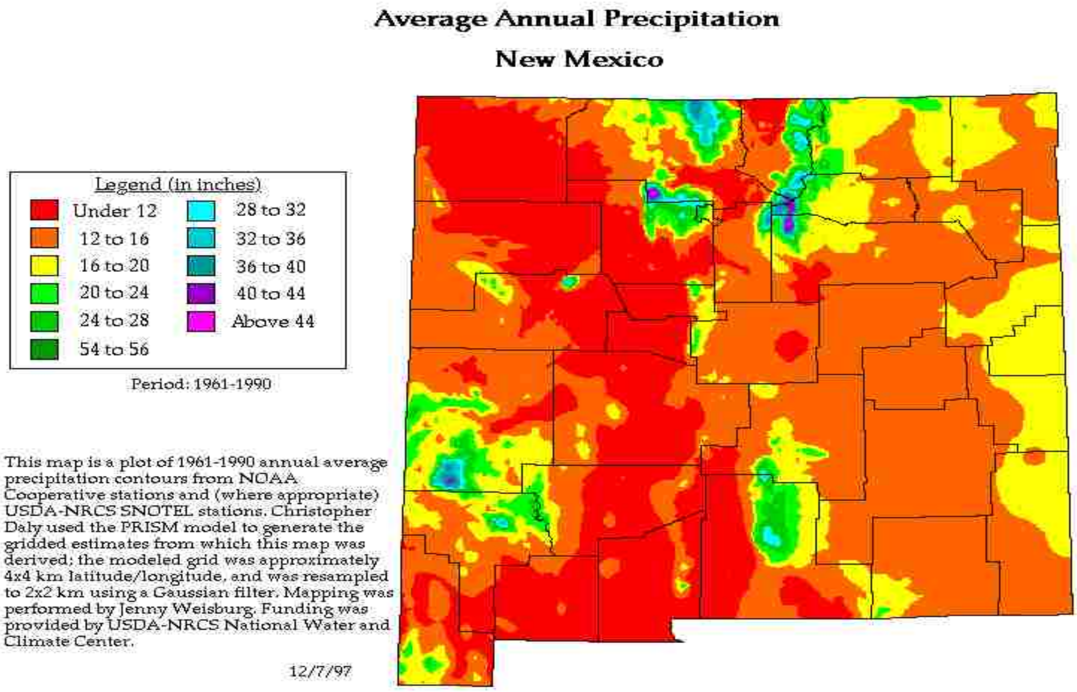


Figure 59: Precipitation map of New Mexico (Center 2012).

Station Location	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Brantley Dam	0.35	0.49	0.37	0.66	1.3	1.57	2.14	1.56	1.83	0.88	0.55	0.59	12.28
Carlsbad	0.4	0.44	0.47	0.65	1.2	1.49	1.86	1.79	2.14	1.34	0.58	0.51	12.86
Carlsbad Caverns	0.47	0.46	0.42	0.63	1.45	1.7	2.09	2.29	2.97	1.38	0.5	0.53	14.87
Carlsbad Airport	0.34	0.4	0.4	0.46	1.19	1.22	1.79	1.9	2.29	1.07	0.51	0.48	12.06
Lake Avalon	0.38	0.37	0.45	0.56	1.36	1.21	1.5	1.48	2.05	1.41	0.44	0.42	11.62
Queen Ranger Station	0.38	0.92	0.69	0.4	0.5	1.85	2.03	4.8	3.7	1.23	0.73	0.69	17.93
Difference Between Highest and Lowest- All Stations	0.13	0.55	0.32	0.26	0.95	0.64	0.64	3.32	1.87	0.53	0.29	0.27	6.31
Difference Between Highest and Lowest of Lower Elevation Stations	0.06	0.12	0.1	0.2	0.17	0.36	0.64	0.42	0.46	0.53	0.14	0.17	1.24

Table 22: Average precipitation for weather stations around project area: Period of Record: Brantley Dam (291153) 8/ 1/1987 to current; Carlsbad (Station number- 291469) 2/ 1/1900 to current; Carlsbad Caverns (291480) 2/ 1/1930 to current; Carlsbad FAA Airport (291475) 9/ 1/1942 to current; Lake Avalon, (294736) 8/ 1/1914 to 2/28/1979; Queen Ranger Station (297176) 1/ 1/1963 to 3/31/1975 (Center 2012).

Data was also obtained on the quality and frequency of the rain events at the different weather stations. Table 23 shows that a little less than half of rain events result in only .01-.099 inches of rain and the vast majority were under half an inch of rain.

	Carlsbad	Carlsbad Airport	Brantley Dam	Lake Avalon	Carlsbad Caverns	Queen Ranger
>= 0.01 in.	42	49	47	33	54	36
>= 0.10 in.	25	25	26	24	30	31
>= 0.50 in.	8	7	8	7	9	13
>= 1.00 in.	3	2	2	2	3	5

Table 23: Average annual frequency of precipitation event.

Model Creation

After the different methods of representing runoff were examined and the appropriate datasets gathered, a model for the project areas water system was constructed using NetLogo. The DEM was imported into the NetLogo world, like in the previous chapter, and this was used to create an aspect and slope dataset to determine the direction of water flows. The soil data was processed through ArcGIS first, to clip the dataset into the project area dimensions, and then imported into the system. These layers would serve as the environment for the model.

Turtle agents served as the water agents in this model. The agents would set their direction based on the two variables, either the aspect of the patch they were on or they would face a neighbour patch with the lowest elevation. Water levels were included in patch elevation but because the resolution of the elevation data was one metre this rarely made a difference in the model. The speed at which the ‘water’ moved was determined by the Manning formula (Equation 20) with took into account the slope of the patch and the roughness of the earth, represented by Manning’s roughness coefficient. This coefficient was set at .0035, which is the coefficient for rough, uneven, thin vegetation environments. Coefficients were not set for each patch as this information was not available.

$$V = Rh^{2/3} * S^{1/2} / n$$

Equation 20: Manning formula. Rh = hydraulic radius which is A/P where A = cross-sectional area. and P = wetted perimeter; S = sloped; n = Manning's roughness coefficient

The initial runoff, used as a variable of the water agents, was set using the runoff curve numbers most appropriate to the project area. There are only three possible variables for soil moisture rates, given this low number of variables all were tested in the model. The same was done with vegetation rates in the project area. Combinations of these rates for the different land forms were also investigated. However, the default levels were set as poor vegetation cover based both on the author's personal observations and from aerial photographs.

However, that only set the initial runoff and was not continuous. Additional reduction of water was taken into account through evaporation, evapotranspiration and soil absorption. More detailed discussion is presented later in this chapter. After the initial runoff was set the agents were let loose into the environment. When they lost all of their runoff, representing loss of water, they would disappear from the model. This could occur by either moving outside of the environment, i.e. draining out of the project area or through evaporation, evapotranspiration, or soil absorption.

Infiltration rates (Table 38) were taken from the soil data and applied to the project area. Also, max water holding capacity was included in this model. As a result, the model included the ability to reach saturation and thus stop absorbing water, should enough water be present in an area. This was modelled using the ratio of hydraulic conductivity to the water capacity of the soil. Research has shown that this relationship is not perfect (Marshall, Holmes et al. 1999) but it is a close enough approximation for the purposes of this study. Infiltration was set to take place after a brief rest, simulated five minutes, to take into account the Runoff Curve Number (RCN) would account for initial infiltration. With infiltration rates were set to two classifications, the lowest and highest rates listed (Table 38).

A problem with runoff curve numbers was that they are ineffective for rain below one inch (USDA 1986). Which is what the historical data indicates is the most common rain event in the project area (Table 23). Other agent based models that have modelled water systems all have used simple infiltration in combination with evapotranspiration or evaporation (McDonnell 2003, Hong, Lin et al. 2007, Reaney 2008, Bithell and Brasington 2009, Bohensky, Butler et al. 2009). Given that other researchers have found no drawbacks to using this method a model was created that used just infiltration and evapotranspiration to look at the lower rainfall events the RCN removed from the model as the initial starting point. It was also decided to include duration and intensity of the rain event. The RCN method does not account for this but seeing that the infiltration and evapotranspiration will not initially reduce the rainfall event it was decided to model this to compensate for the lack of initial reduction.

Adjustable time frames were created to represent the scale at which the agents moved. This scale also determined how much 'water' was lost every turn. While this scale was adjustable it was set to a five-minute scale for all of the models presented here. This was to ensure that the agents did not overshoot their targets when moving. It also gave an easy to use time frame that did not extend out the modelling time too significantly and was small enough to capture small changes in water levels.

There were two possible ways to determine the direction of flow: aspect and elevation. Choosing which variable the models followed had significant repercussions for the outcomes. If elevation is followed, the agent moves to the patch with the lowest elevation, where an agent can potentially run into dips. These are areas of the project area where the agent has moved to what is the lowest point relative to its surrounding so the agent must wait for evaporation, evapotranspiration, or soil absorption to cause it to be removed from the surface. Agents following aspect, in contrast, do not take into account these dips in the project area and ignore water traps. The code for this model can be found in Appendix C.

Model Results

The different models, i.e. those with RCN and without, aspect or elevation based paths etc., were tested with an assortment of different variables and settings; the results can be seen in Table 24. Those scenarios with the higher infiltration rates result in higher amounts of surface water being present longer. This may seem counter-intuitive, as higher infiltration rates would result in the water being absorbed faster and less likely to be on the surface. The reason this occurs is because higher absorption rates result in a higher initial absorption but this in turn results in lower longer term absorption rates as soil reaches its maximum capacity more quickly. Lower absorption rates result in more water moving outside of the project area in flows and higher absorption rates and capacity later to clear out the left over water.

	Rain Fall (in)	RCN	Direction	Infiltration Rate	hr 1	hr 6	hr 12	day 1	day 2	day 3	day 5	day 7	end
	.01	None	Elevation	High									35 mins
	.01	None	Elevation	Low									40 mins
	.01	None	Aspect	High									35 mins
	.01	None	Aspect	Low									40 mins
	.1	None	Elevation	High	7804	4197	3052	1677	543	177	19	1	8 days
% of total					16%	8%	6%	3%	1%	0%	0%	0%	
Max depth					4.89	4.07	3.94	3.68	3.16	2.64	1.60	0.57	
	.1	None	Elevation	Low	4055	1746	1392	862	311	110	12	1	9 days
% of total					8%	3%	3%	2%	1%	0%	0%	0%	
Max depth					4.36	4.47	4.34	4.08	3.56	3.04	2.00	0.97	
	.1	None	Aspect	High	3943	41	2						18 hrs
% of total					8%	0%	0%	0%	0%	0%	0%	0%	
Max depth					4.24	1.77	0.30						
	.1	None	Aspect	Low	1810	2	0						15 hrs
% of total					4%	0%	0%	0%	0%	0%	0%	0%	
Max depth					4.08	0.26	0.06						
	.5	None	Elevation	High	131558	88398	83305	74271	59069	47194	30489	19856	49 days
% of total					53%	35%	33%	30%	24%	19%	12%	8%	
Max depth					27.44	25.26	25.13	24.87	24.35	23.84	22.80	21.76	
	.5	None	Elevation	Low	76900	24355	22854	20060	15482	11950	7158	4328	35 days
% of total					31%	10%	9%	8%	6%	5%	3%	2%	
Max depth					18.91	18.15	18.02	17.76	17.24	16.72	15.69	14.65	
	.5	None	Aspect	High	99414	170	1						23 hrs
% of total					40%	0%	0%	0%	0%	0%	0%	0%	
Max depth					31.92	2.64	0.56						
	.5	None	Aspect	Low	63743	4	0						23 hrs

	Rain Fall (in)	RCN	Direction	Infiltration Rate	hr 1	hr 6	hr 12	day 1	day 2	day 3	day 5	day 7	end
% of total					25%	0%	0%	0%	0%	0%	0%	0%	
Max depth					24.07	0.66	0.24						
	1	None	Elevation	High	305299	206520	200689	189273	168163	149871	118924	94611	97 days
% of total					61%	41%	40%	38%	34%	30%	24%	19%	
Max depth					50.23	50.12	49.99	49.73	49.22	48.7	47.66	46.62	
	1	Normal	Elevation	High	63499	49439	44314	35775	23526	15646	7041	3236	29 days
% of total					13%	10%	9%	7%	5%	3%	1%	1%	
Max depth					14.89	15.09	14.96	14.7	14.18	13.66	12.63	11.59	
	1	Low	Elevation	High	6,809	3,052	1,971	977	333	152	36	9	11 days
% of total					1%	1%	0%	0%	0%	0%	0%	0%	
Max depth					5.79	5.68	5.55	5.29	4.77	4.26	3.22	2.18	
	1	Saturated	Elevation	High	148,244	109,891	104,305	93,583	76,038	62,041	31,734	28,246	60 days
% of total					30%	22%	21%	19%	15%	12%	6%	6%	
Max depth					31.19	31.08	30.95	30.69	30.17	29.65	28.61	27.58	
	1	None	Aspect	High	158981	107	1						1 day
% of total					32%	0%	0%	0%	0%	0%	0%	0%	
Max depth					52.1	3.45	0.74						
	1	None	Aspect	Low	139523	4.5	1						1 day
% of total					28%	0%	0%	0%	0%	0%	0%	0%	
Max depth					57.06	0.87	0.74						
	2	None	Elevation	High	602251	429004	422098	408326	381166	355188	310029	269231	144 days
% of total					60%	43%	42%	41%	38%	36%	31%	27%	
Max depth					78.68	74.47	74.34	74.08	73.56	73.04	72	70.97	
	2	Low	Elevation	High	104206	81806	76160	66058	50007	38038	22538	13475	41 days

	Rain Fall (in)	RCN	Direction	Infiltration Rate	hr 1	hr 6	hr 12	day 1	day 2	day 3	day 5	day 7	end
% of total					10%	8%	8%	7%	5%	4%	2%	1%	
Max depth					20.57	21.06	20.93	20.67	20.15	19.64	18.6	17.56	
	2	None	Aspect	High	143592	57	1						22 hrs
% of total					14%	0%	0%	0%	0%	0%	0%	0%	
Max depth					82.96	2.79	0.2						
	5	None	Elevation	High	1303238	1014590	1005776	988197	953027	917977	848196	779845	256 days
% of total					52%	41%	40%	40%	38%	37%	34%	31%	
Max depth					132.84	132.73	132.6	132.34	131.82	131.3	130.27	129.23	
	5	Low	Elevation	High	668829	528915	521461	506593	477004	447685	392016	347011	175 days
% of total					27%	21%	21%	20%	19%	18%	16%	14%	
Max depth					95.59	90.54	90.41	90.15	89.63	89.12	88.08	87.04	
	5	Normal	Aspect	High	71214	14							12 hrs
% of total					3%	0%	0%	0%	0%	0%	0%	0%	
Max depth					74.4	1.31							

Table 24: Results of agent based models. All with evaporation rates for June. Water left by days is measured in patches (500,000) by depth of water.

Possible Model Needs Influencing Model Results

The models where elevation was dictating water flow direction resulted in dips, where water pools and no longer forms stream flows. That meant that these water agents no longer need to take into account the soil absorption rates of the soils it might land in; a process that takes up both computing capacity and time, as the models need to make calculations for every patch at every tick of the model. Nor does it need to run through the entire simulation to determine when it will disappear, as these calculations can be done by dividing the water by the evaporation rate. To reduce the run time of the models when less than 1% of the water in the project area had stopped moving, the model would stop and compute the remaining factors with simple maths, i.e. amount of water divided by evaporation rates, instead of running out the model.

This significantly reduced the time it took the models to run. Yet, this also changed the outcomes of the models but only slightly. For example, results between the optimised model and one that ran all the way through was a difference of a few hours for max time water was still present in the project area: a difference of <5% in results. This difference did not matter to the final results and conclusions drawn but it is something to be aware of if trying to use the model for your own purposes.

Results Against Known Events

To confirm that these models are accurate representations of the actual hydrology of the project area the results were compared against the historical data collected. That data needed to be adjusted because it came from two different sources. The USGS had runoff collection stations (Table 16) but WWCR had precipitation data. There were only two USGS stations, 08405150 (Carlsbad) and 08405150 (Carlsbad Airport) in the subject area. Interestingly, it was found that when comparing the USGS data to the WWCR data there was limited correlation between rainfall and runoff: .46 and .49 for Carlsbad Airport and Carlsbad (Table 41). There is actually a correlation in the amount of rain and runoff experienced.

There are several reasons for this discrepancy between the datasets. One is that the stations are in different locations. Rainfall was not distributed evenly over the project area. Table 22 illustrated how different weather stations record different amounts of water; some of these stations were only a few miles away from each other. Another reason for discrepancy between datasets is that the intensity of the rainfall affects runoff flows. Using the agent based model it was possible to demonstrate how intensity, not just volume, affects runoff. The American Meteorological Society characterises rain events as 'light', 'moderate' or 'heavy'. Light rain is considered rain of between trace amounts and 0.25 cm (0.10 in.) per hour, with a maximum rate of fall being no more than 0.025 cm (0.01 in.) in six minutes. Moderate rain is between 0.26 and 0.76 cm (0.11 to 0.30 in.) per hour and a maximum rate of no more than 0.076 cm (0.03 in.) in six minutes. While heavy rain is considered any rain over 0.76 cm (0.30 in.) per hour or more than 0.076 cm (0.03 in.) in six minutes (Society 2012). Plugging those factors into the model demonstrated how intensity changes the outcomes (Table 25).

This was done by instead of having the water agent represent all the water from an event multiple water agents were created over a period of time. A process that created many more agents but that was also more reflective of real world conditions.

Rain	Rain	Hr. 1	Hr. 6	Hr. 12	Day 1	Day 2	Day 3	Day 5	Day 7	End
0.1	Light	12207	7747	5629	3002	855	243	25	3	8 days
% of total		24%	15%	11%	6%	2%	0%	0%	0%	
Max depth		3.84	4.25	4.12	3.86	3.34	2,82	1.79	0.75	
0.1	Moderate	8528	4447	3126	1633	469	136	11	1	9 days
% of total		17%	9%	6%	3%	1%	0%	0%	0%	
Max depth		5.08	4.49	4.37	4.11	3.59	3.07	2.03	1	
0.1	heavy (.5 per hr)	3770	1929	1506	920	321	108	11	0	7 days
% of total		8%	4%	3%	2%	1%	0%	0%	0%	
Max depth		4.63	3.55	3.42	3.16	2.64	2.12	1.09	0.05	
0.5	Light	114281	91267	85079	74066	56472	43505	26355	16196	45 days
% of total		46%	37%	34%	30%	23%	17%	11%	6%	
Max depth		23.39	23.29	23.16	22.9	22.38	21.86	20.83	19.79	
0.5	Moderate	104286	77689	72776	63777	49226	38159	23287	14332	55 days
% of total		42%	31%	29%	26%	20%	15%	9%	6%	
Max depth		28.69	28.58	28.45	28.19	27.67	27.16	26.12	25.08	
0.5	heavy (.5 per hr)	96263	71242	66489	57885	44101	33792	20288	12292	42 days
% of total		39%	28%	27%	23%	18%	14%	8%	5%	
Max depth		27	21.44	21.31	21.05	20.53	20.02	18.98	17.94	
1	Light	257436	228284	220250	204506	176803	152938	115772	87911	97 days
% of total		51%	46%	44%	41%	35%	31%	23%	18%	
Max depth		39.75	49.96	49.83	49.57	49.05	48.53	47.49	46.46	
1	Moderate	255848	201105	194564	181708	158369	138712	1066640	82533	81 days
% of total		51%	40%	39%	36%	32%	28%	213%	17%	
Max depth		42.18	42.07	41.94	41.68	41.16	40.64	39.61	38.57	
1	heavy (.5 per hr.)	240373	183367	177228	165207	143337	125062	95402	73260	79 days
% of total		48%	37%	35%	33%	29%	25%	19%	15%	
Max depth		41.01	40.9	40.78	40.52	40	39.48	38.44	37.41	
1	heavy (1 per hr.)	231026	173396	167438	155795	134610	116317	87320	65846	75 days
% of total		46%	35%	33%	31%	27%	23%	17%	13%	
Max depth		37.97	38.86	38.74	38.48	37.96	37.44	36.4	35.37	
1	heavy (2 per hr.)	242119	180381	174341	162441	140196	121040	90536	68227	78 days
% of total		48%	36%	35%	32%	28%	24%	18%	14%	
Max depth		35.24	40.05	39.92	39.66	39.14	38.62	37.59	36.55	

Rain	Rain	Hr. 1	Hr. 6	Hr. 12	Day 1	Day 2	Day 3	Day 5	Day 7	End
2	Light	508906	464717	455756	437863	402427	368616	312848	264709	120 days
% of total		51%	46%	46%	44%	40%	37%	31%	26%	
Max depth		62.38	62.27	62.14	61.88	61.36	60.84	59.81	58.77	
2	Moderate	516220	456780	448534	432091	399687	368519	315927	269288	146 days
% of total		52%	46%	45%	43%	40%	37%	32%	27%	
Max depth		75.53	75.43	75.3	75.04	74.52	74	72.97	71.93	
2	heavy (.5 per hr)	502661	418633	411052	395980	366286	337443	288745	245434	179 days
% of total		50%	42%	41%	40%	37%	34%	29%	25%	
Max depth		69.98	69.87	69.74	69.48	68.97	68.45	67.41	66.37	

Table 25: Surface water in the project area resulting from different rain intensities. Time counted from after rainfall stops.

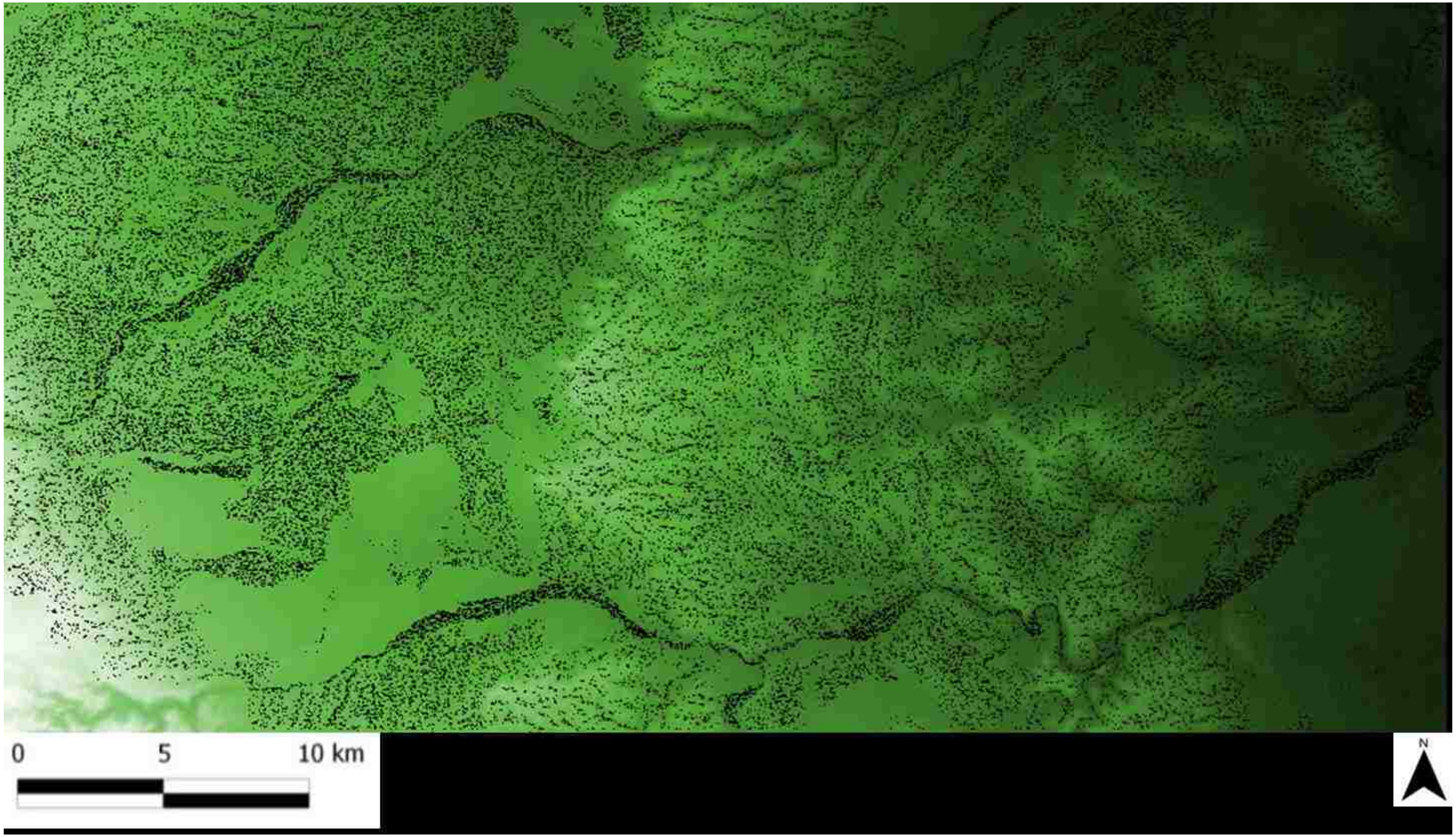
Unfortunately, the historical distribution of rainfall over the project area was only measured by a few weather stations (Figure 58), making it impossible to determine rain distribution accurately for each historic rain event. The same goes for rainfall intensity; there was no dataset for the project area that gives historical rainfall intensity. As such, the project was left with a crude approximation of how rainfall correlates with runoff – when rainfall is over an inch there is usually a runoff event. However, most of the time precipitation translates into very little runoff (Table 25).

Year	Runoff ac-ft	Rain (in)	% runoff out of total
2011	135	5.06	0.11092%
2010	2280	17.32	0.54728%
2009	917	11.96	0.31876%
2008	0.06	9.39	0.00003%
2007	845	19.02	0.18470%
2006	668	8.67	0.32032%

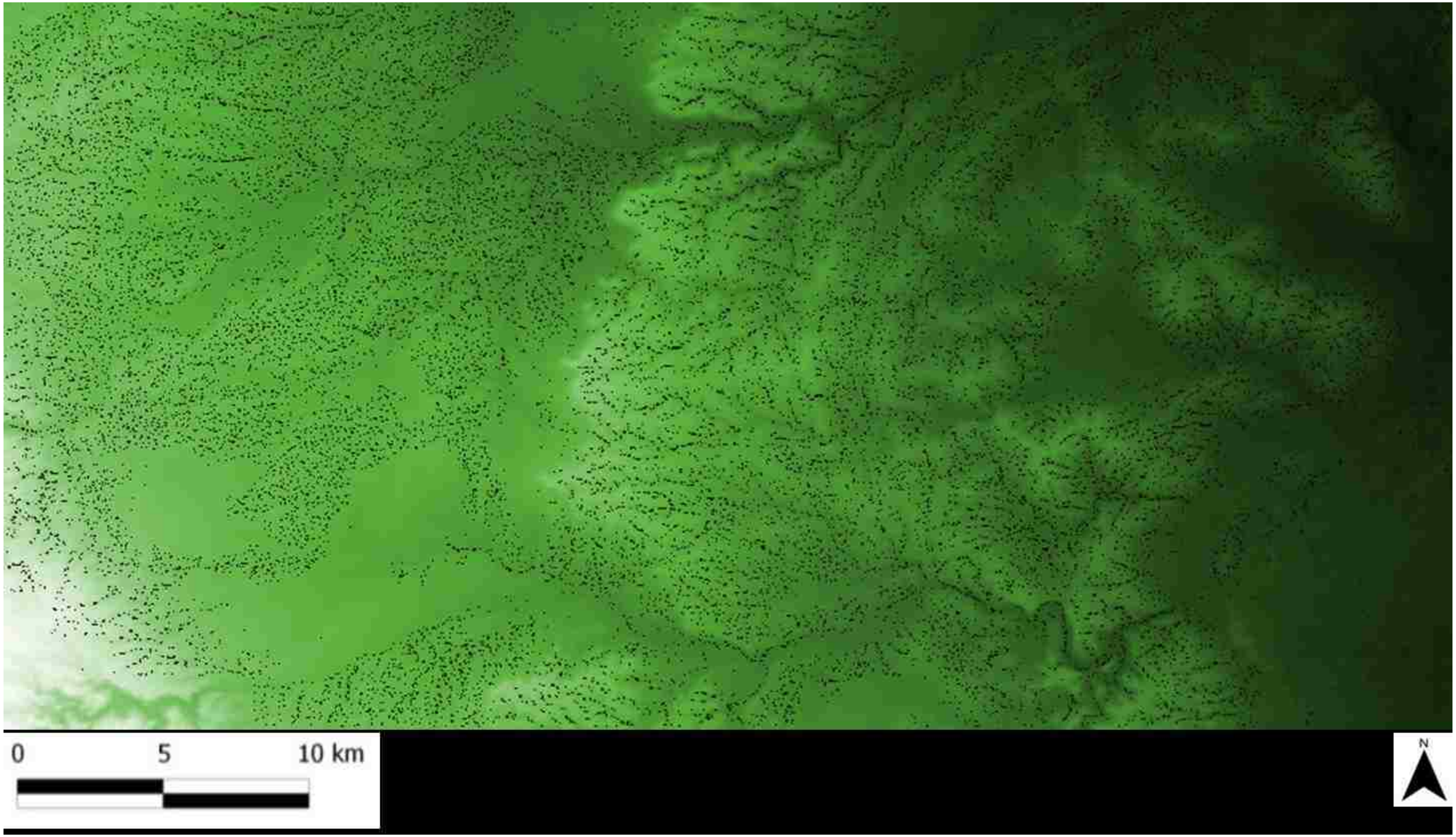
Table 26: Runoff and rain per year for Carlsbad.

This historical approximation is what is seen in the agent based model. For small rain events there is almost no runoff and water is almost completely gone from the project area within hours (Figure 60). Moreover, the runoff is so weak that only a few major waterways appear to attract what could be called flows. Most of the water ends up being absorbed and not draining away. Coalescing into ponds or lakes also does not happen, as there is not enough water to pool into anything significant. A few small pools/puddles form but they quickly dissipate. Only large rainfall results in anything that could be considered streams or large pools and those are rare events in the project area (Table 21).

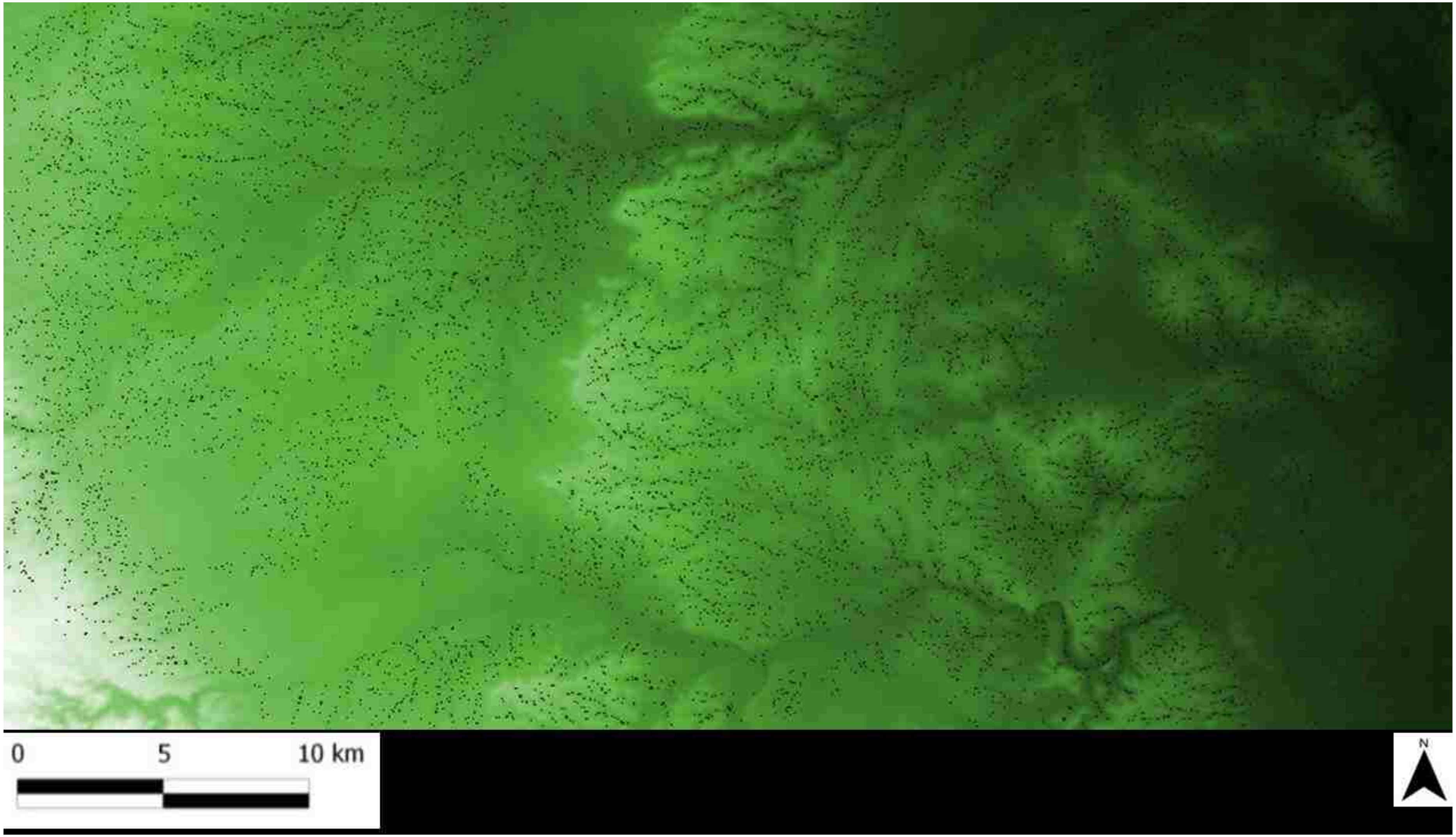
Figure 60: Water (black) in project area for 0.1 inch rainfall event with low intensity rain at A. 1 hour B. 6 hours C. 1 day D. 3 days.



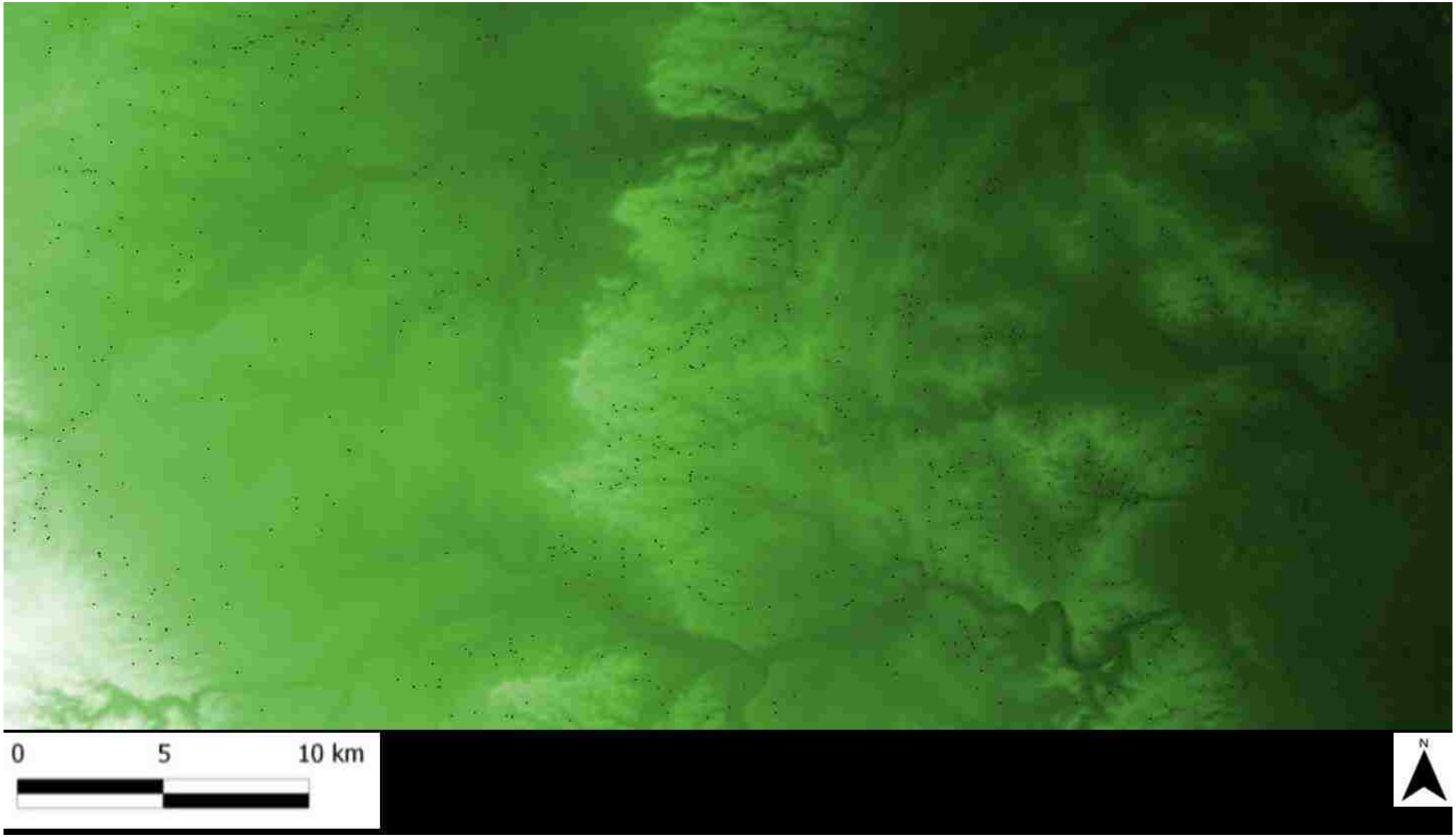
A.



B.



c.



D.

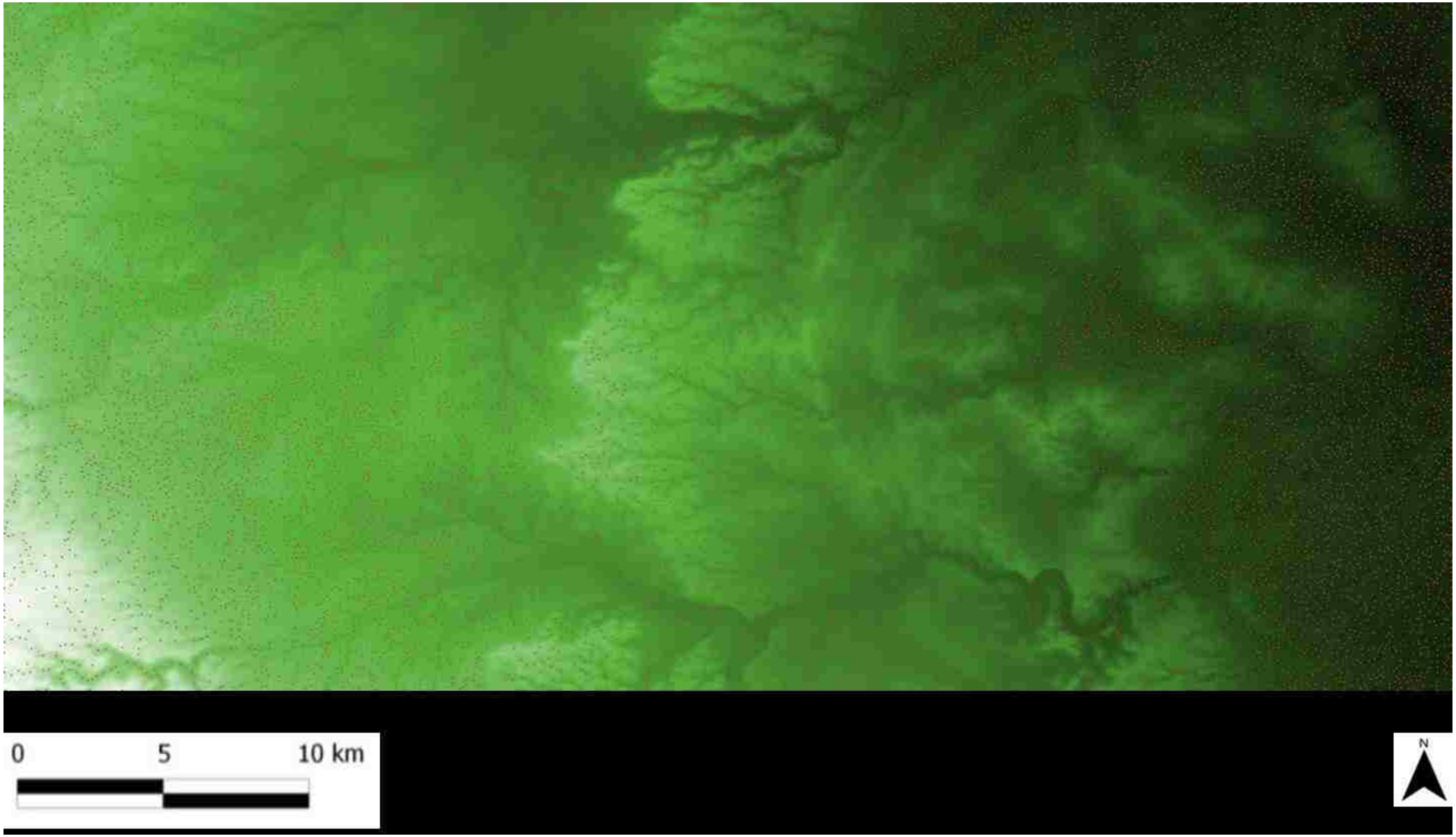


Figure 61: Sinks in the project area. Red dots represent patches where none of the neighbours have a lower elevation.

Sinks and Pools

Even if the large rainfall events were rare, the models showed that there were many areas that would have had pools (Figure 61). Most models showed that flows in the form of streams and flash floods subside after only a few hours. All water still available in the project area after a precipitation event was in the form of these pools. The models demonstrated that any available water would have been in the form of pools, possibly some springs, and not streams.

There were limits to the conclusions that could be reached from this sort of model. The sinks in the project area were the result of the DEM data used, the accuracy of that data and its resolution. The resolution is only a vertical metre, which meant that any difference in elevation, which dictated movement, of less than a metre will not have been registered. The accuracy of the DEM data is only 17 metres at a 95% confidence. Different convolution functions used to interpolate the data were found to change the placement of sinks in the project area. Moreover, a sink in the DEM could actually be the result of a data error. This error could result from a problem with the radar that gathered the points. It could have occurred with the transfer of the raw data or even with the reading of the data by NetLogo. That made it questionable to try to ascertain the exact location of these sinks in the project area from the model.

Determining which water sinks were real, or not, was not possible with the model and data. Nevertheless, the model did provide some data as to where these sinks may exist on a larger scale. Slopes have to be present to funnel surface water into the sinks, but not too steep, as that reduces the probability of having sinks. Figure 61 best illustrates the regions that are most likely to contain sinks in the project area.

Duration of Habitation

Results for the duration of the runoff and the intensity of that runoff over time indicate that the majority of the project area was not suitable for long-term habitation. When the rain intensity is significant enough surface water can be present for several days, even weeks, in small amounts. Historical data conversely indicates that the probability of such rainfall is low (Figure 62), with a less than 10% chance of 5 inches (12.7 cm) of rain during any given year. At best, water would be present for brief periods of time, in the late summer and early fall.

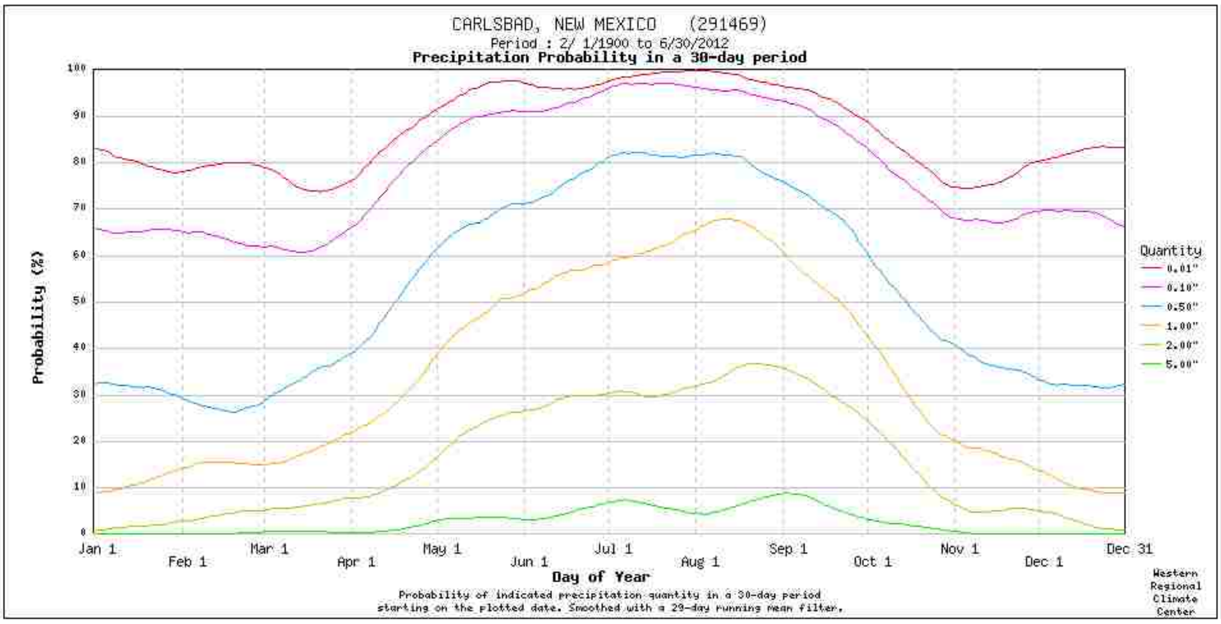


Figure 62: Precipitation probability in a 30 day period for Carlsbad, New Mexico (Center 2012).

This meant that people could either:

1. Bring in water to the centre of the project area from the Pecos River or sources in the Guadalupe Mountains during periods of low perception. Limiting the time spent in the area to how much water can be carried into the area.
2. Wait until a rain storm of intense enough downpour occurs (> 1 inch) and then move into the project area to use the residual water to survive for several days or weeks, assuming sinks occur.
3. Move into the project area during months with high probability of rain events and survive off of the rainfall and/or surface water in sinks that are likely to occur after a rainfall.

The first option would handicap the amount of time a person could spend in the project area. For example, if a person drinks three litres of water a day they would have to carry with them 3.03 kg of water. A week would require 21.21 kg of water; a significant amount of weight to carry into the project area. In that scenario there are obvious limits to the amount of time one could stay away from a source of water.

Discussion of Water Model

While the models provide very interesting results, it also left the project with still unanswered questions. The PUMP III project had run statistical analysis and found that site location had the strongest correlation with the drainages. However, the model showed that these drainages would not have contained water for 99.9% of any given year. Even during wetter years it was unlikely that the drainages would have water. The landscape did not have the capacity to maintain perennial water sources.

There were possible exceptions to this. Knowledge of the geology and water table indicates that during wet years the Indian Big Spring might have been a potential stable water source. Moreover, there was an edge effect occurring in the model. Because the project area did not include outside areas, like the Guadalupe Mountains, that could have fed water into the project area, there is the potential that outside sources might have provided water. Indeed, the PUMP III project had problems with its models because it forgot to take into account the Pecos River.

However, the Indian Big Spring and outside water sources are funnelled into either one of the two major drainages. Even if those drainages did contain water for any matter of time, more than a few hours, it could not account for the correlation between site locations and all drainages (**Confidential Appendix Figure 98**). Sites are associated with many of the minor drainages. The same drainages that the model shows did not contain water, even during extreme water events. Leaving the mystery of why people were attracted to these draining locations if it was not for the water?

Chapter 8: Attractors

> ‘We know that people came into the area, possibly in small, mobile groups that exploited locally available resources and then left...’ (Altschul et al. 2005 p. 106)

The first two models created for this project indicated that it was unlikely that the area could have supported part-time agricultural subsistence, except for areas near the Pecos River. Furthermore, the idea that the area was used as a thoroughfare was not possible to prove and that theory could not explain many of the locations of different sites but it may contribute to some sites locations. With these activities ruled out as explanations for site location, the project explored the subject of attractors bringing people into the project area i.e. ‘people came into the area, possibly in small, mobile groups that exploited locally available resources and then left’. The models discussed in this chapter were created to explore possible attractors that would bring people into the project area and possibly created the archaeological record.

Lithic Material Procurement

One attractor that is rarely discussed in predictive modelling literature is quarry sites to procure raw materials such as chert, obsidian, quartzite, etc. During the research review only two predictive models were found to have created predictive models that looked at quarry sites in Wyoming (Church 1996) and Iowa (Goings 2003). Models of the distribution of chert across Europe have also been produced (Duke and Steele 2009). For clarification, when this thesis refers to lithic sources it specifically is referencing material used for cutting tools such as chert or obsidian, as opposed to materials used for decoration or utilities such as building or food preparation, like grinding stones. The advantages of modelling lithic sources were best articulated by Church:

> ‘Lithic sources offer several distinct advantages for modelling. The first is that they are unmoving and relatively unchanging. They can only be physically depleted. This also means that they are present today and have been investigated and mapped to some extent by geological studies, thus providing a reliable base of information. Further, desired sources of chipped stone are a relatively rare occurrence in the landscape. Second, stone was a consistently sought resource until the introduction of metal. Prehistoric people could often decide to shift from one to another comparable subsistence resource, if necessary. They could shift from big game to small game or from a game emphasis to a gathering emphasis. No such option existed in terms of the stone needed to perform many day-to-day activities, although they might shift emphasis within the narrow set of types of stone.’ (Church 1996 p. 157)

Predicting quarry sites served two purposes: one was to predict the location of a single type of site. The other was to predict the location of associated sites. For example, if the source material is more than a day’s travel from the normal habitation zones then campsites, and associated sites, will be created along the routes. Examples of this sort of construction are discussed by Andrefsky Jr. (1994) or Bamforth (2006).

There were several lithic quarries listed in the NMCRIS archaeological site records for the project area. These were broadly categorised as modern quarries (LA 88108, LA 131359), some of which resulted from vandals digging up sites (LA 116471) of large prehistoric quarries with thousands of artefacts (LA 28752, LA 67513, LA 112620, LA129466) or smaller prehistoric sites where lithic materials are listed as present but in small enough quantities that they are not extensively mined (LA 116399).

Methodology for Modelling Lithic Material Procurement

The first step in creating a lithic sources model was to look at the geology of the area. This information was obtained from the USGS (Anderson, Jones.G.E. et al. 1997). The USGS data showed that there was a lack of volcanic activity in the area which ruled out obsidian as a source of lithic material. However, the three primary rock formations that make up the project area, the Yates and Tansill Formation, Seven Rivers Formation and Grayburg and Queen Formation, are composed of carbonated rock such as limestone and dolostone; some of the common materials chert and chalcedony deposits are found in (Rapp 2009). The data provided by the USGS on geological formations in the area indicated that, in fact, the project area was rich in potential for lithic quarries, almost everywhere except the alluvium plains (Figure 63).

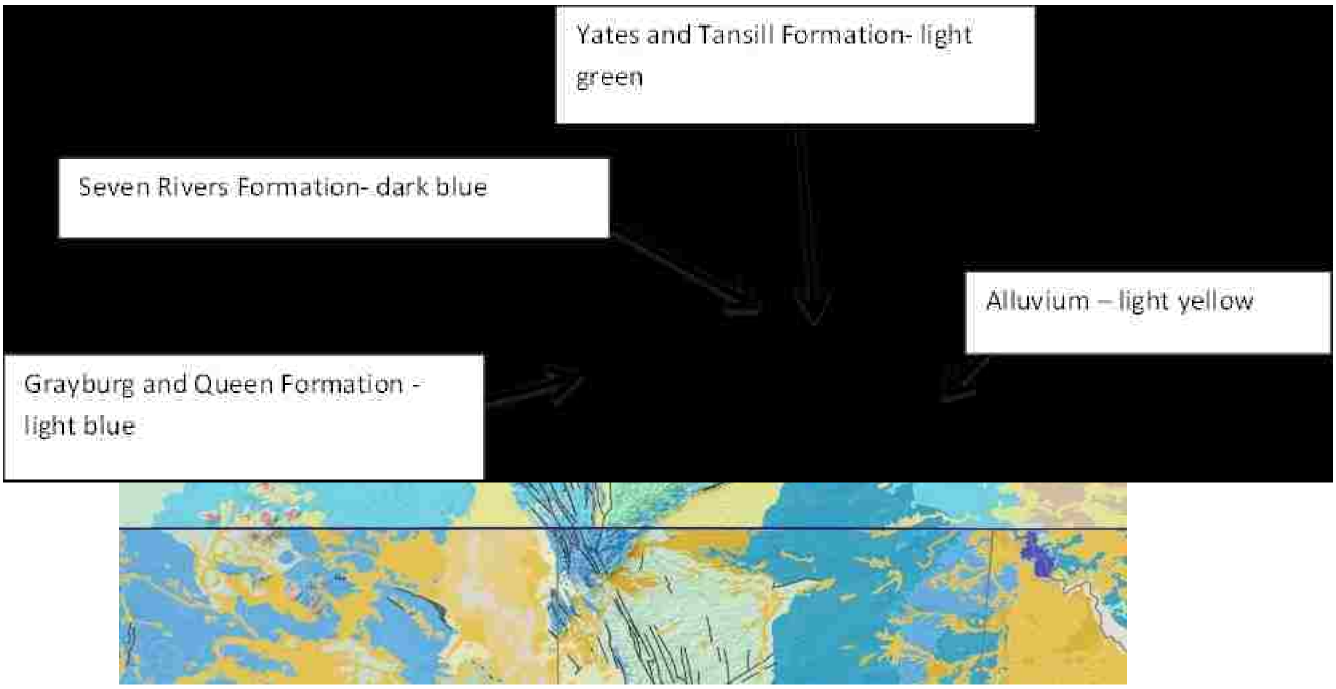


Figure 63: Geological map of southeaster New Mexico. Yates and Tansill – Permian (Age) Fine-grained mixed clastic rock (Primary) Carbonated rock (Secondary). Formation Seven Rivers Formation - Permian (Age) Evaporite (Primary) Fine-grained mixed clastic rock (Secondary). Grayburg and Queen Formation: Permian (Age) Carbonate rock (Primary) Sandstone (Secondary). Modified from (USGS 2012c)

The next set of variables examined were distance to water, slope, relief and depth to bedrock. These variables were chosen to find areas that the deposits of chert were likely to be exposed and accessible to people in the past. Hillsides and riverbeds are natural areas where erosion will expose the underlying bedrock. Uplifted areas of the landscape, where bedrock has been moved closer to the surface, are other areas of a landscape with higher likelihood of finding exposed deposits. Even if the majority of the area has geological potential it was thought areas with likelihood of finding quarries could narrow down the search area.

GIS over Agent Based Modelling

Given the static nature of the datasets i.e. geological, GIS was used to create this model. As discussed in previous chapters GIS is a better tool for dealing with geographic data and has greater functionality when dealing with such data than agent based modelling programs i.e. buffering, combining datasets, and a host of the other functions. GIS was used to combine different datasets to create a map of areas with little soil and with either rugged areas or stream beds i.e. areas with exposed bedrock and likely to contain quarries.

Most of the data had already been gathered for the previous models in this project. Soil depth was obtained from the NRCS data (Figure 64). The DEM data from the NASA provided all of the necessary base data to create datasets for slope and relief. For this model the ruggedness index created through the QuantumGIS (QGIS) tools was used instead (Figure 65). Ruggedness is a quantitative measurement of terrain heterogeneity. Details of Ruggedness is described by (DeGloria, Elliot et al. 1999). It is calculated for every location, by summarizing the change in elevation within the 3x3 pixel grid. The reason for using Ruggedness is because it acts as a proxy measurement of abrupt change in the landscape and thus the likelihood of finding exposed rock outcrops. The ruggedness index was based on a 3x3 moving window (DeGloria et al. 1999).

Water drainage data was the same USGS data discussed in Chapter 8 (Figure 66). There was no need to use the complex water model created previously because the model only needed to know where the water moves and not the absorption rates and other details of the model. The USGS data was sufficient for these purposes as it contained all the areas that potentially might have water erosion.

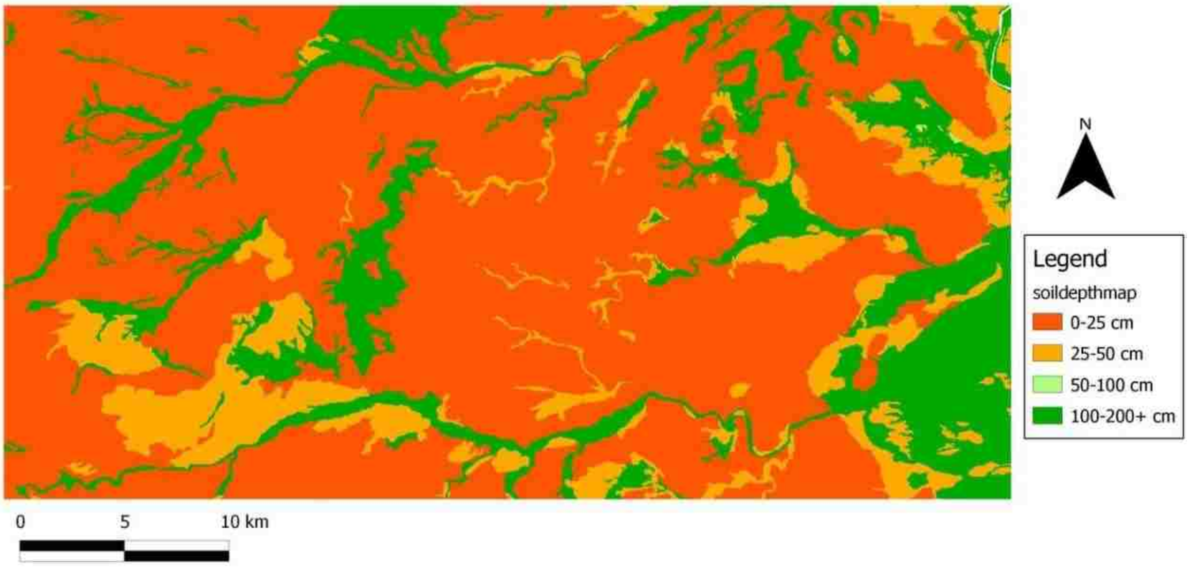


Figure 64: Soil depth map of project area.

Figure 65: Ruggedness index of the project area.

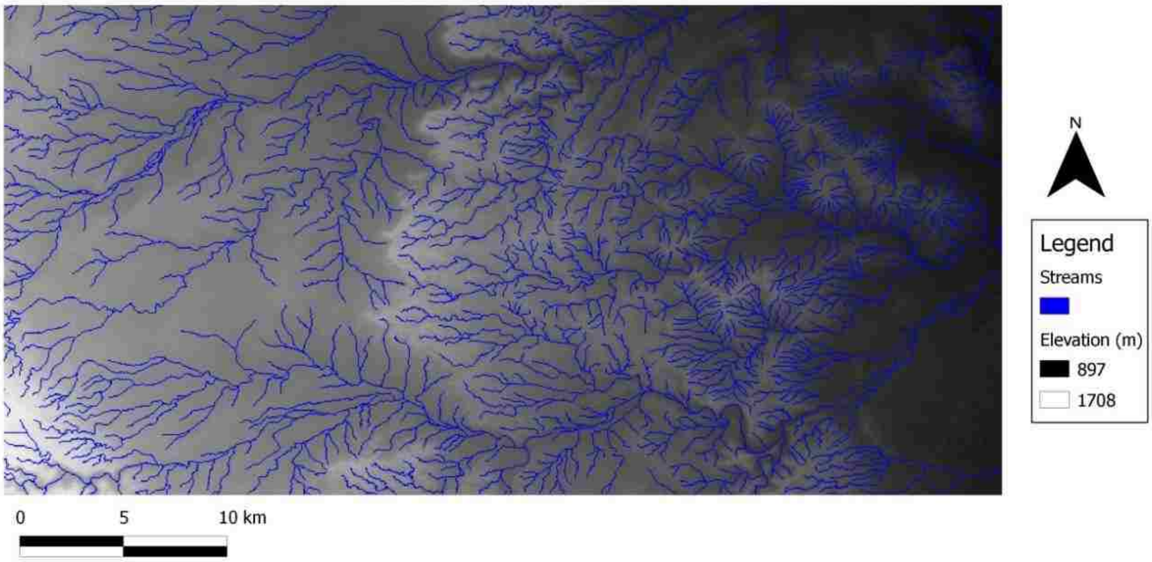


Figure 66: Water drainage paths in the project area.

All of the datasets were processed using QGIS and the GRASS GIS plugin for QGIS. The ruggedness application was also applied to the elevation model to create a ruggedness dataset (Figure 65). The soil depths were reclassified to those areas below and above 25 cm in depth. This was to eliminate areas with soil depths above 25 cm from the areas with potential for quarry sites. This is an arbitrary classification but, based on the resolution of data available it was felt this would eliminate areas with high deposits of soils that would obscure quarry sites. This dataset was then converted into positive and negative zones and combined with datasets of the water drainage features and ruggedness to produce the potential map (Figure 67). High soil depth was a negative factor while drainages and ruggedness were considered positive attractors and added together with equal weights.

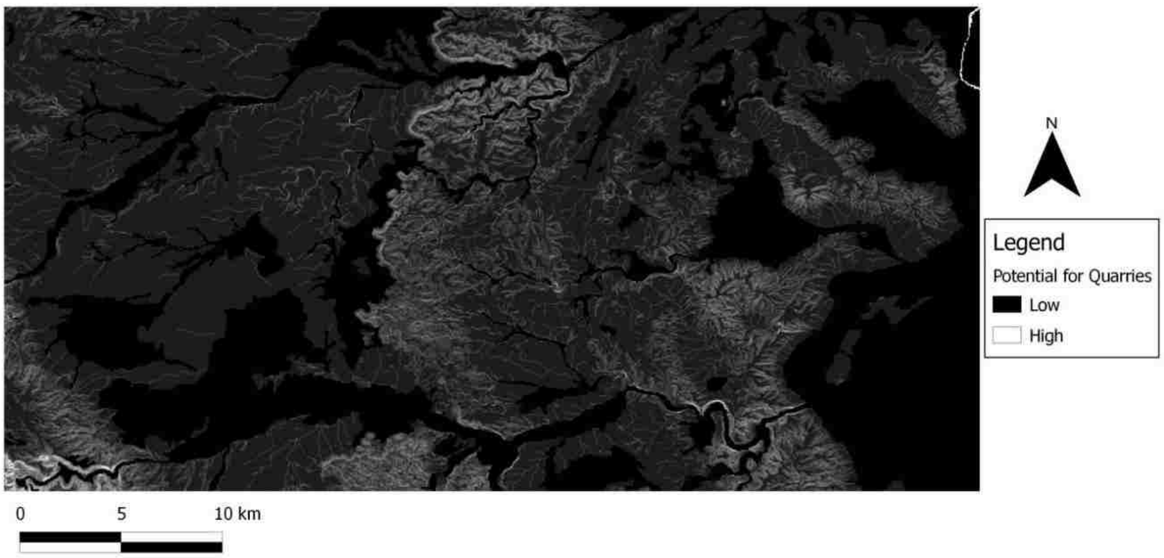


Figure 67: Probability map for quarry sites.

It was noted that all of the known quarry sites are above 3800 ft. (1159 m) (Table 27). It was not known if this is because no sites have been positively identified at lower elevations or because this elevation represents a natural seam of lithic materials. In case this did represent a natural seam of material, a filter of 1159 metres above sea level was added to the final model to narrow down the range of high potential areas (Figure 68). **Also see Figure 99 in Confidential Appendix.**

Site (LA)	Elevation (feet above sea level)
116399	4940
28752	4530
67513	4280
112620	3870
129466	4440

Table 27: Elevations of quarry sites in project area.

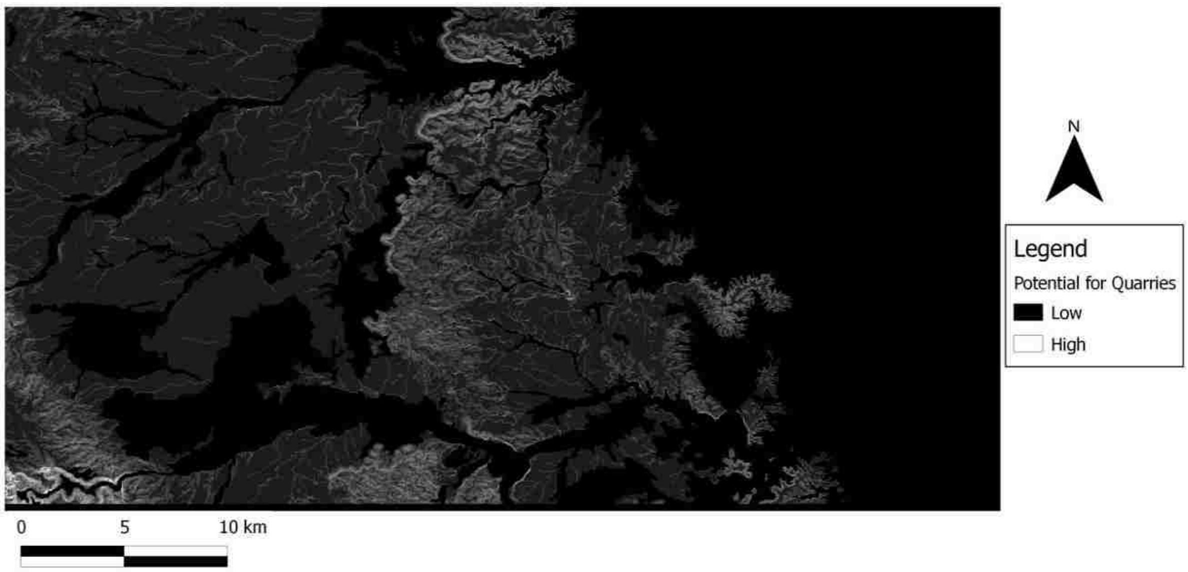


Figure 68: Quarry probability map for sites located above 1159m above sea level.

Discussion of Lithic Procurement

The result of the quarry predictive model showed that a significant portion of the project area could contain quarry sites. Not only was it a high percentage of the project area but these areas were also those that were labelled by the past site predictive models as not likely to contain sites, e.g. bedrock, steep slopes. This complicated the process of creating a predictive model, as quarry sites, while relatively rare in the project area, could potentially be located throughout the landscape. This eliminates that as a possible aspect for predictive modelling as too much of area would have to be labelled as possibly containing sites but only a few sites would be captured.

Depending on whether the above 1159m elevation model was an accurate representation of all quarries, two potential different resource procurement strategies were identified. One based on special journeys from the Pecos River area to obtain lithic materials from quarries because there were no resources near the river. The other potential strategy was to obtain materials during other activities undertaken in the project area, a ‘pick up the milk on your way home from work strategy’.

Both were possible but the later one does fit in with a pattern of food acquisition discussed later in this chapter. The special journey scenario is also realistic but cannot explain the location of most sites in the project area. Many sites were located in the Eastern portion of the project area.

Shelter in the Project Area as an Attractor

Without a satisfactory explanation for site locations based on quarrying, this project then looked at shelter. That was because, regardless of why people were in the project area, most journeys would have required an overnight stay. As shown in Figure 70, most of the project area is not within a day’s walk of either the foothills of the Guadalupe Mountains or the Pecos River. This meant that past people would have had to camp throughout the project area. One possible attractor of people looking for a place to stay was caves and rock shelters. Given that the temperature in the project area can rise above 100 F (38 C), cool caves and shady areas were a logical location to try to model.

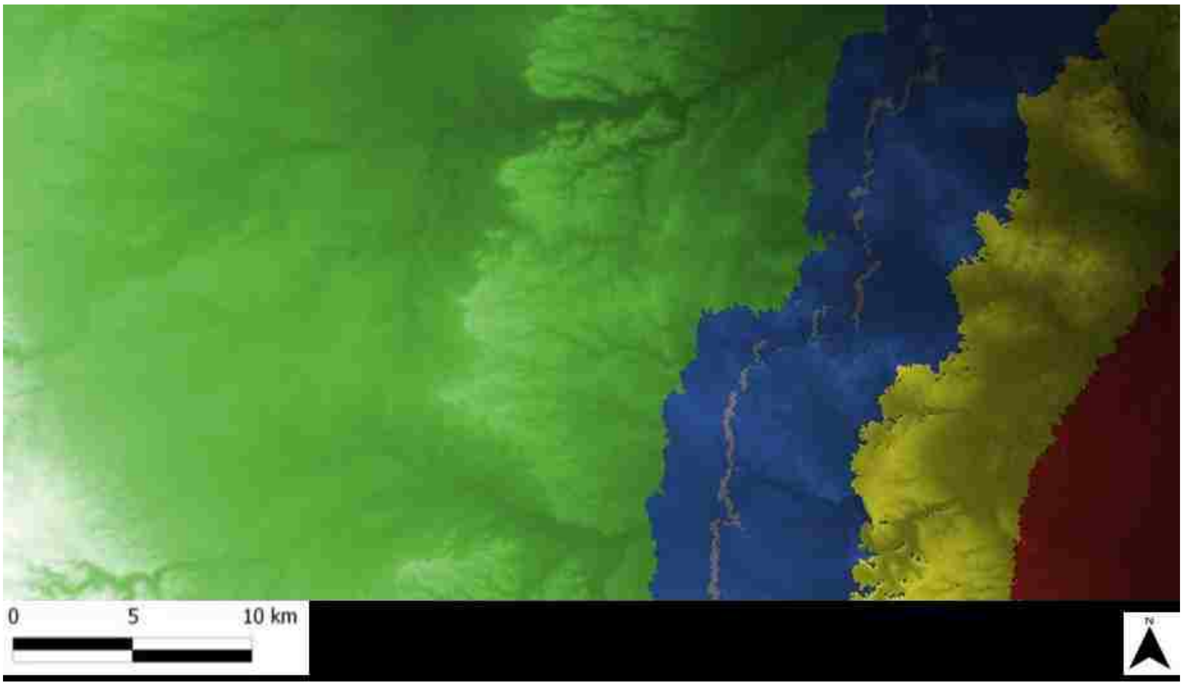


Figure 69: Travel time for south eastern corner of the project area: Red (0-3hrs.); Yellow (3-6 hrs.); Blue (6-9 hrs.). Most people experience a significant drop in pace after eight hours of walking (gray line).

Designing the Shelter Model

Similar to quarries, the project determined that these sites would best be modelled by the variables of slope and soil. Slope was used because rockshelters are formed by overhangs of bedrock which require vertical and near vertical rock outcrops (Figure 70). It would be very hard to find an overhang on level ground. Soil, on the other hand, does not cause overhangs to be formed but it does determine the likelihood of finding caves. Alluvium soils at the bottom of drainages and on flood plains would fill in any caves in those areas. Thus a cave needs to be above drainage zones where they could not be filled in by soil (Figure 71).

Figure 70: Overhang on the north face of Dark/Last Chance Canyon. Right photo is an enhanced view of the left photo.

Figure 71: Cave above the flood plain on south side of Dark/Last Chance Canyon.

As before, all of these datasets were already used in this project and were repurposed to find potential areas with rockshelters and caves. However, the dataset for water features was not used. While erosion through water would cause many of the features to be created, this would have to take place over millennia and it was felt by the author that the USGS dataset could not accurately reflect water features tens of thousands or even millions of years ago. Basically, the cave and rockshelter model used the same methods as the quarry model but removed the waterway dataset. The resulting model can be seen in Figure 72.

Discussion of Cave and Rockshelter model

The number of listed caves and rockshelter sites for the project area (Table 28) is significantly more than quarry sites. These known sites served as a dataset to check if, in fact, these type of sites are found in the predicted zone. The results (Figure 72) did, in fact, capture almost all of the known caves and rockshelters. Several were not captured but further investigations found that this was due to a data error. The soil data for some foothills near the entrance of Dark Canyon/Last Chance Canyon (purple rectangle in Figure 72) were labelled as deep alluvium soils as opposed to lightly covered outcrops of limestone outcrops. This fact was confirmed by a site visit in 2012. Changing this data resulted in correctly capturing the known sites within the model.

LA #	Name (if one is recorded)	LA #	Name (if one is recorded)
1770	Sheeo Draw Caves	82638	
9052	Honest Injun Cave	89375	Sunny Day Shelter, Sunny Day Smoked Shelter
14179	The Hobo Site	89376	Peek-a-Boo Shelter
14288	Robert's Indian Cave, Roberts Rockshelter, Roberts Rockshelter #1, Rockshelter #1	89377	Roberts Rockshelter 2
14289	Ellis Site	89378	Arrowhead Cave, Arrowhead Shelter
43426		89379	Riley's Folly Cave
43427		89380	Pancho's Cave
43431		89381	Dead Sheep Shelter
43432	Dark Canyon Shelter	89382	Water Hole Shelter
43434	Dark Canyon Cave	89383	Heidi's Cave
43435	Horizon Shelter	89529	Anderson's Cave
43440	Richard Brown Site	89530	Root Cave
43441		89531	Duplex Cave
43442		89532	Dogleg Cave
43446	Double Ended Cave	101494	Burial Cave
43449		101495	Cremation Cave, Barker Pen Cave
43671	Sacahuiste Draw	112611	
43673		113504	Maverick Canyon Site
43674	Roberts Cave, Robert's Indian Cave, Rockshelter #2	113606	Rattlesnake Cave
43676		116398	
43677		116399	
43679	Walt Canyon Site	116400	Elephant Head Shelter
43682		116404	
61348	Sitting Bull Falls Shelter	116405	
68252		140882	
81502	Kee Painted shelter, Kee Shelter, Key Shelter	140942	

Table 28: Known caves and rockshelters in the project area.

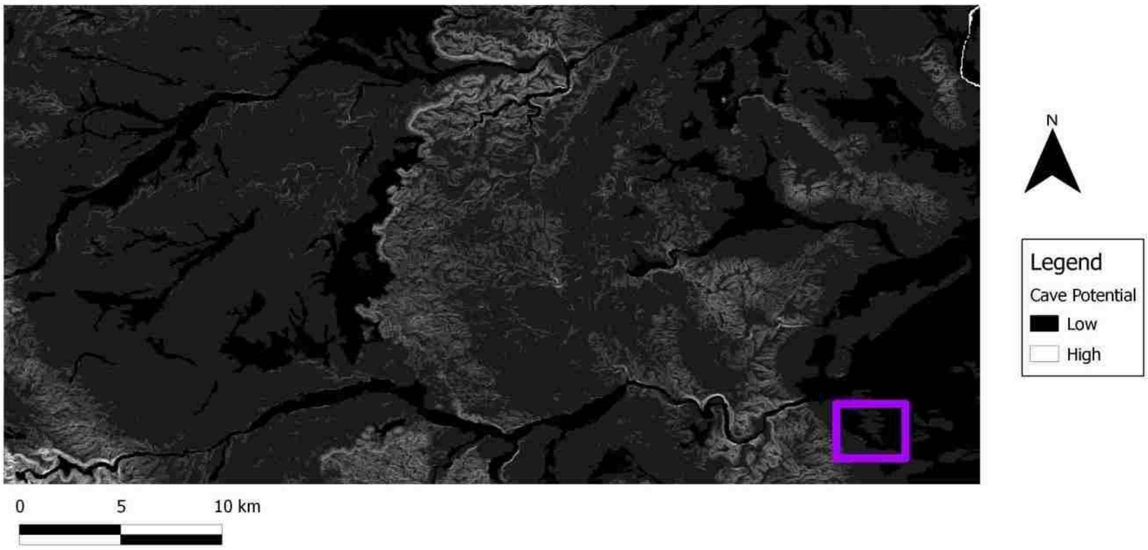


Figure 72: Probability of finding caves and rockshelters. Purple rectangle is mislabelled soil data.

Again, the results of the cave and rockshelter model indicated that there was high potential for 'red flag' sites throughout the project area. These were areas missed by the previous model. 7 % of these types of sites did not match the most significant characteristics used in the previous site predictive models, which could also explain some of the poor performance of those models. Moreover, the archaeological record indicates that these sites are more prevalent than quarry sites, with 52 known examples (Table 28 and **see Figure 100 in Confidential Appendix**). They were potentially a greater issue than quarry sites.

Subsistence in Determining Site Location

With travel routes, caves and rockshelters, approximation to water and quarries ruled out as the primary explanation for site location, additional avenues were explored. An examination of the known sites found that close to 60% of sites contained fire cracked rock (FCR) features. These FCR features were referred to in the records in several different ways, such as mescal pits, ring middens, hearths and roasters. These features were typically used to cook the edible agave plant (mescal) (Figure 73). These plants must be cooked because the sap of agave is toxic and can cause acute contact dermatitis (Crosby 2004).

This process has a long cultural history in that area of the US. The Mescalero Apache who inhabited the area during recent protohistory (see Chapter 6) practised agave roasting till very recently:

> 'Although not so extensively as formerly the Apache women still make expeditions to regions where agave grows in abundance for the purpose of collecting the edible portions of the plant. These visits are usually made in late May or early June when the reddish flower stalks begin to appear and the plants are most palatable.' (Castetter and Opler 1936 p. 35)

Neighbouring Apache bands called the Mescalero *Nadahéndé*, 'People of the Mescal' (Opler and Opler 1950). Mescal roasting still occurs in the city of Carlsbad to this day (Emery 2012).

Figure 73: Image of agave *lechuguilla* (Anderson 2004). The primary agave plant found in the project area.

Archaeological Data

A substance strategy of agave roasting and collecting was observed in the central Jornada cultural area, on the west side of the Guadalupe Mountains. It was found that from the Archaic up to the Early Pueblo period the inhabitants practised seasonal movement and resource collection (Whalen 1994). Past peoples would stay in the river valleys and mountains during the winter but venture into the highland desert areas during the rainy seasons to collect resources; primarily agave. This was in the Jornada cultural complex, which also includes the project area. Excavations in the surrounding region have yielded a similar pattern of subsistence and the importance of agave to past peoples (see Mera (1938) or Katz and Katz (1985)). On the basis of this data from the surrounding areas and the fact that 60% of the recorded sites had roasting features it was possible to focus on agave roasting as a key determinant activity in the study area.

Agave Roasting

Anthropologists working with the Mescalero in the 1930s described the process of roasting agave, which gave insights into how to model agave gathering and roasting:

> 'The crowns of the mescal plants are dug out with three foot sticks cut from oak branches *(Quercus* sp.) and flattened at the end. This end, when pounded with a rock into the stem of the plant just below the crown, permits the crown to be removed readily. A broad stone knife is used to chop off the leaves, two being left for tying the crowns together, thus making transportation more convenient. The naked crowns are bulbous, white in colour, and from one to two feet in circumference.
>
> Pits in which the crowns are baked are about ten to twelve feet in diameter and three or four feet deep, lined with large flat rocks. On the largest rock, which is placed in the centre, a cross is made with black ashes. Rocks are piled on the flat stones, but care is always taken that the top shall be level. Upon this, oak *(Quercus* sp.) and juniper wood *(Juniperus* sp.) are placed. Before the sun comes up this is set on fire and by noon the fire has died down.
>
> On these hot stones is laid moist grass, such as bunchgrass *(Sporobolus airoides),* side-oats grama *(Bouteloua curtipendu2a),* Texan crab-grass *(Schedonardus paniculatus),* big blue-stem *(Andropogon furcatus),* mesquite grass *(Muhlenbergia wrightii),* marsh foxtail *(Alopecurus aristulatus), Muhlenbergia neomexicana,* or the leaves of bear grass *(Nolina microcarpa),* but bear grass is usually preferred since it does not burn readily… After the mescal has been covered with the long leaves of bear grass and the whole with earth to a depth sufficient to prevent steam from escaping, the crowns are allowed to bake the rest of the day and all night. Early in the morning the pit is opened and a crown examined and eaten. The pit is again closed and the Indians refrain from drinking until noon of this day so as to prevent rain. The following morning all the mescal is removed.' (Castetter and Opler 1936 p. 35-36)

Creating roasting pits and roasting agave took at least two days of work, possibly longer, and required a variety of different materials, from rocks to plant materials. This investment in time meant they could not simply roast agave for lunch on their trip out to a quarry but that this was a planned out activity instead of an opportunistic one. People would have had to came into the project area to specifically undertake agave roasting, given the time and resource commitment involved.

This assumption was corroborated by the ethnographic data on such tasks. Castetter and Opler (1936) indicate that the procurement of agave was a significant event by itself. The other accounts of agave gathering and roasting mentioned that excursions usually lasted ten days to two weeks away from the primary camps, during which other resources such mesquite beans were also collected (Hodgson 2001). This was usually undertaken by groups of five to eight women (Buskirk 1986). While applying ethnographic accounts to the study of past peoples can be fraught with problems, it was noted that these accounts fit the line of evidence seen in Chapter 8 that excursions into the project area would be limited to only a few weeks due to a lack of water. Thus the ethnographic evidence matches the known environmental data for the project area.

Roasting Pit?

The pit method of cooking agave, described in the ethnographic accounts, does not reflect all of the archaeological remains found in the project area; a fact noted by the ethnographic accounts. Even in the 1930s the archaeological record indicated that the area near Carlsbad did not contain solely pit features (Bell, Castetter et al. 1938) but included roasters that were built on top of the surface instead of into a pit (Mera 1938). Moreover, Castetter and Opler (1936) also noted that there are earlier ethnographic accounts of Apache boiling agave near the Guadalupe Mountains. An account from Marcy in 1849 states:

> '… we have this evening for the first time seen the Maguey plant which constitutes almost the only vegetable food that the Apaches and southern Comanches get for a great portion of the year. They prepare it by boiling it until it is soft, then mash it into a paste, and I am told that in this form it makes a very palatable, nutritious food.' (Marcy 1850)

Castetter and Opler (1936) dismissed this as not being agave because the Mescalero Apaches they observed did not boil their agave, although this did not exclude it as a possible method. Boiling of agave can leave piles of fire cracked rock (FCR) as one method to boil water is to drop heated rocks into water. The process causes the rocks to split and, discarded after use, these can build up refuse piles of FCR. Thus the wide range of descriptor features in the archaeological record from ring midden to mescal pit to piles of FCR are likely to represent the different means of cooking agave in the project area.

Methodology

Based on the archaeological and ethnographic data several assumptions were made in creating the model. The first of these assumptions was that roasting features would be situated close to the actual resources. This was based on ethnographic accounts – one group would gather the agave from close by and another group would create the roasting pit (Buskirk 1986). Furthermore, estimations of distance travelled and time spent collecting plant resources varies from hunter-gatherer societies around the world and could suggest a time limit at around an hour or two and only a few kilometres from a base camp (Waguespack 2005). Base camps being the temporary camps set up during the agave roasting activities. Thus proximity to resources was emphasised in this model.

Another assumption made was that while agave grew, and still grows, on steep slopes the location of roasting pits requires more level ground. This assumption was based on an examination of the archaeological data of the project area (Appendix D). Sites with roasting features are located on level ground or areas with low slopes. This was probably to prevent the rock piles and the fires from sliding down the hillside. As a result, sites will be located not in areas of high potential for agave plants but flat areas close to those areas.

Agave habitation

The project area contained three species of agave that are edible, Agave *gracilipes*, Agave *neomexicana*, Agave *lechuguilla* and A. *gracilipes* which is a hybrid of the other two species

(Burgess 1979). Agave typically grows on dry ground at elevations between 3,000 and 8,000 feet (915-2440 m) (Bell et al. 1938, Buskirk 1986), well within the relief of the project area. Actual investigations into the distribution of the project area have shown that A. *lechuguilla* grows at lower elevations than the other two species, and while there is overlap, they tended to be segregated by elevation zones.

Modern Data

A first attempt was made using modern vegetation datasets to attempt to model agave distributions. Agave plants are considered a primary indication of the Chihuahuan desert habitation and with the GAP program (USGS 2012b) (Chapter 5) it was thought that one could identify these areas. The GAP data provides detailed information on a scale of one acre that clearly indicates the areas that encompass the Chihuahuan desert vegetation class; the areas in which agave was expected to be found. However, Chihuahuan desert vegetation accounted for 86.31 % of the project area (Figure 74). That is too high of a number to create a predictive model that can be precise.

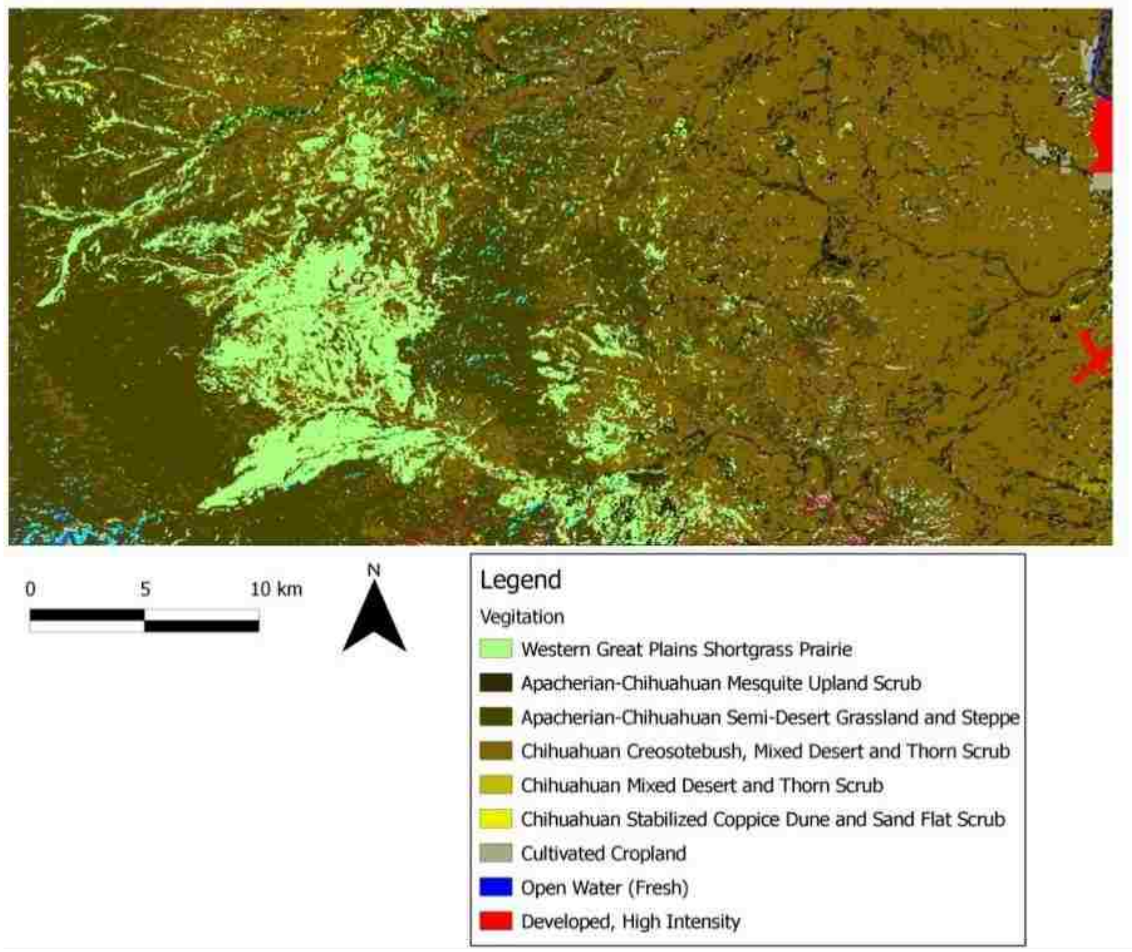


Figure 74: Gap data distribution of vegetation. Only those classes making up more than 1% of the project area are listed.

Distribution Studies

The next approach taken to determine distribution of agave was to examine flora studies of the Carlsbad area. A project had been conducted in the 1960s on the Guadalupe Escarpment to determine dominance and prevalence of plant species for different stratum and landscape conditions (Gehlbach 1967). As the Guadalupe Escarpment encompasses part of the project area it was decided to use this data to determine the optimal locations to find agave in the project area. Spot checks during visits to the project area confirmed the accuracy of this method.

Gehlbach's investigation found that Agave *lecheguilla* was the leading half-shrub between 3,800 and 4,600 ft. (1159-1402 m). It had higher IV numbers, a measurement of dominance, than any other species in any stratum below tree level (Figure 75). This is because the species forms extensive clones on coarse gravelly loams or weathered limestone outcrops. This clonal nature of A. *lecheguilla* suppresses other species in its range. The range of A. *lecheguilla* was found to be between a little over 3,700 ft. to around 5,200ft (1128-1585 m) in elevation and the highest point of dominance was around 4,200-4,600 ft. (1281-1402 m) (Gehlbach 1967) .

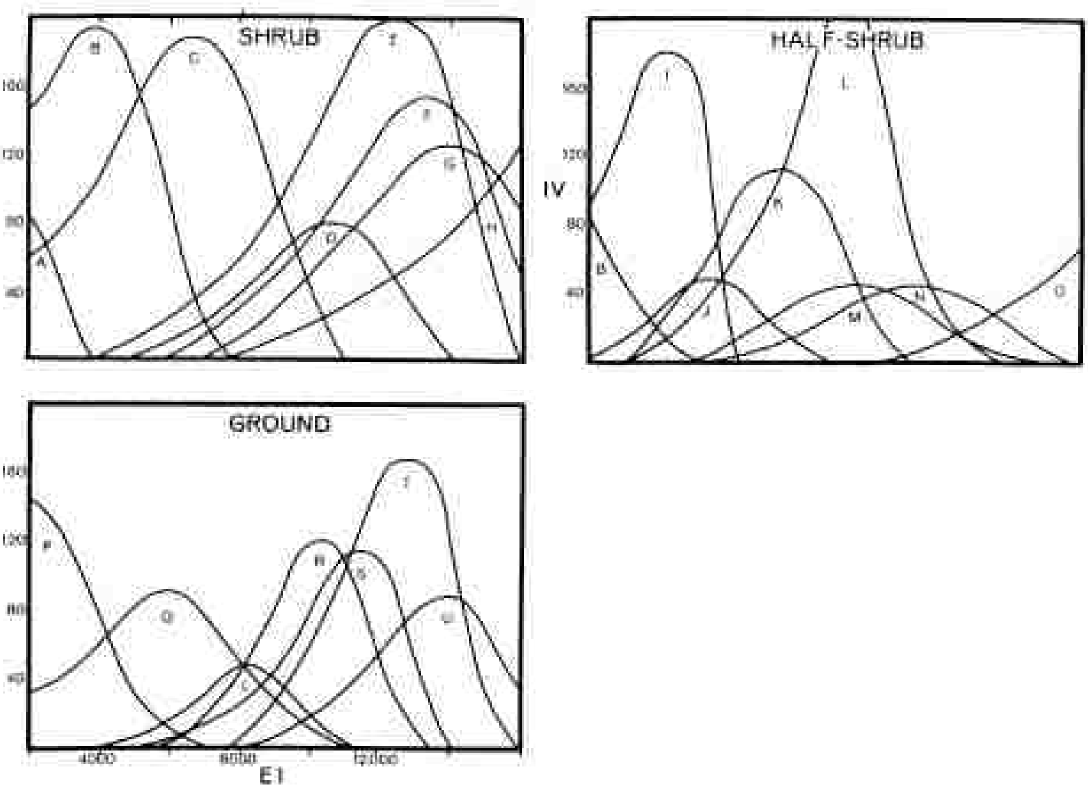


Figure 75: Ecologic amplitudes of some leading dominants in three strata on the Guadalupe Escarpment (A- Flourensia cernua, B-Larrea tridentata, C-Acacia constricta, D-Mimosa biuncifera, **E-Dasylirion leiophyllum (sotol)**, F-Juniperus monosperma, G-Quercus grisea, H-Rhu-s trilobata, I-Gutierrezia lucida, J-Parthenium incanum, K-Opuntia phaeacantha, **L-Agave lecheguilla**, M- Viguiera stenoloba, N Opuntia engelmanni, O-Nolina microcarpa, P-Tridens pulchellus, Q-Bouteloua eriopoda, R-Muhlenbergia setifolia, S-Tridens muticus, T- Tridens pilosus, U-Bouteloua curtipendula) (Gehlbach 1967 p.409 Figure 4).

In addition to dominance of agave by elevation the research also determined that there were only slight variations in density of agave between north and south facing slopes. Because of this lack of significant consistent difference this project assumed there was no difference in north and south facing slopes for agave. Moreover, it was found that agave is located on slopes in the project area and not in the valley floors of canyons. These slopes range from 10-20% gradient and above, but not sheer cliff faces (Gehlbach 1967). Further studies of agave on the Guadalupe Escarpment have also shown that while A*. lecheguilla* is dominant at lower elevations, A. *gracilipes* and A. *neomexicana* can be found at higher elevations. The range of agave does not stop at the 5,200 ft (1,585 m) limit of A. *lecheguilla* but continues higher in elevation, though with different species.

Results of Modelling Agave Distribution

Using the distribution of agave, described above, and QGIS, a map of potential distributions was constructed. Elevation was weighted in percentages, according to the distribution of A. *lecheguilla* in the literature*.* These involved weighting the elevations 4,200-4,600 ft (1,281-1,402 m) at 100 % likely to contain agave. 4,600-4,900 ft (1,402- 1,494 m) and 3,900-4,200 ft. (1,189- 1,280m) were weighted at 50% and the elevations of 3,700-3,900 ft (1,128- 1,189 m) and above 4,900 ft. (1,494 m) were weighted at 30%. All areas above 4,900 ft were given the 30% weight to account for A. *gracilipes* and A. *neomexicana* while everything below 3,700 ft was given a 10% weight. This was based on the distribution data. The ecological data also indicated that agave would not grow on slops of less than 10 degrees. To complete the model, the slope dataset was reclassified to eliminate those areas with slopes less than 10 degrees who were given a 0 weighting verses a 1 for the other areas. All of these layers were then multiplied together using the map algebra tool to create an estimation of the distribution of agave in the project area (Figure 76).

Figure 76: First attempt at mapping agave potential.

Accounting for Paleoenvironment

This first attempt at mapping did not correlate well with known sites that contain roasting features. Several possible explanations were then investigated to attempt to improve the model. Paleoenvironmental studies have shown that the distribution of plants can and does move up and down in elevations during different climate periods. One estimate is that during the end of the Pleistocene a 1,300-1,400m depression of life zones existed, extending the tree line all the way down to the Pecos River. However, others have disagreed with this assessment and believe that a xerophilous juniper woodland would have existed at this point in time instead of a dense forest. The general consensus, though, is that a 1,000m depression of ecological zones existed during this period (Van Devender 1995).

In terms of the model this meant that as the climate changed from the late Pleistocene into the early and middle Holocene there were significant movements in ecological zones upwards. The current desert scrub lands and grasslands that agave inhabits would have moved across the project area and would not have been present at all periods of time. In addition, there are micro-periods of wetter and drier conditions at various periods in the south west which would have expanded or shrunk the desert scrubs and grasslands. A wet period has been confirmed to exist at 3000 BP as indicated from faunal remains (Applegarth 1979) and if the conditions were the same across New Mexico others would have occurred at 3000 and 1000 BP. These would have been the time periods with possible seasonal springs (see Chapter 8) but less agave.

However, expanding the ranges of high potential for agave growing up and down the elevations did not improve the model greatly. Furthermore, not knowing the age of most sites, it was not known if there was, in fact, any correlation between climate and the location of roasting pits. Because of these issues other solutions were then examined.

Sotol

Agave is not the only plant that was roasted in the past; other plants include sotol (*Dasylirion leiophyllum) (*Figure 77). Sotol was another desert plant that is indicative of the Chihuahuan desert. It returns a similar number of calories as agave when roasted, though typically a little less (Dering 1999). Sotol, unlike agave, was in more environmentally diverse locations as found by researchers. It was on slopes of hills but also in water drainages. Furthermore, it was also more sensitive to aspect, with some southern facing slopes having a 320% increase compared with north facing slopes. But in most cases the difference is closer to 50% and that is the norm (Gehlbach 1967).

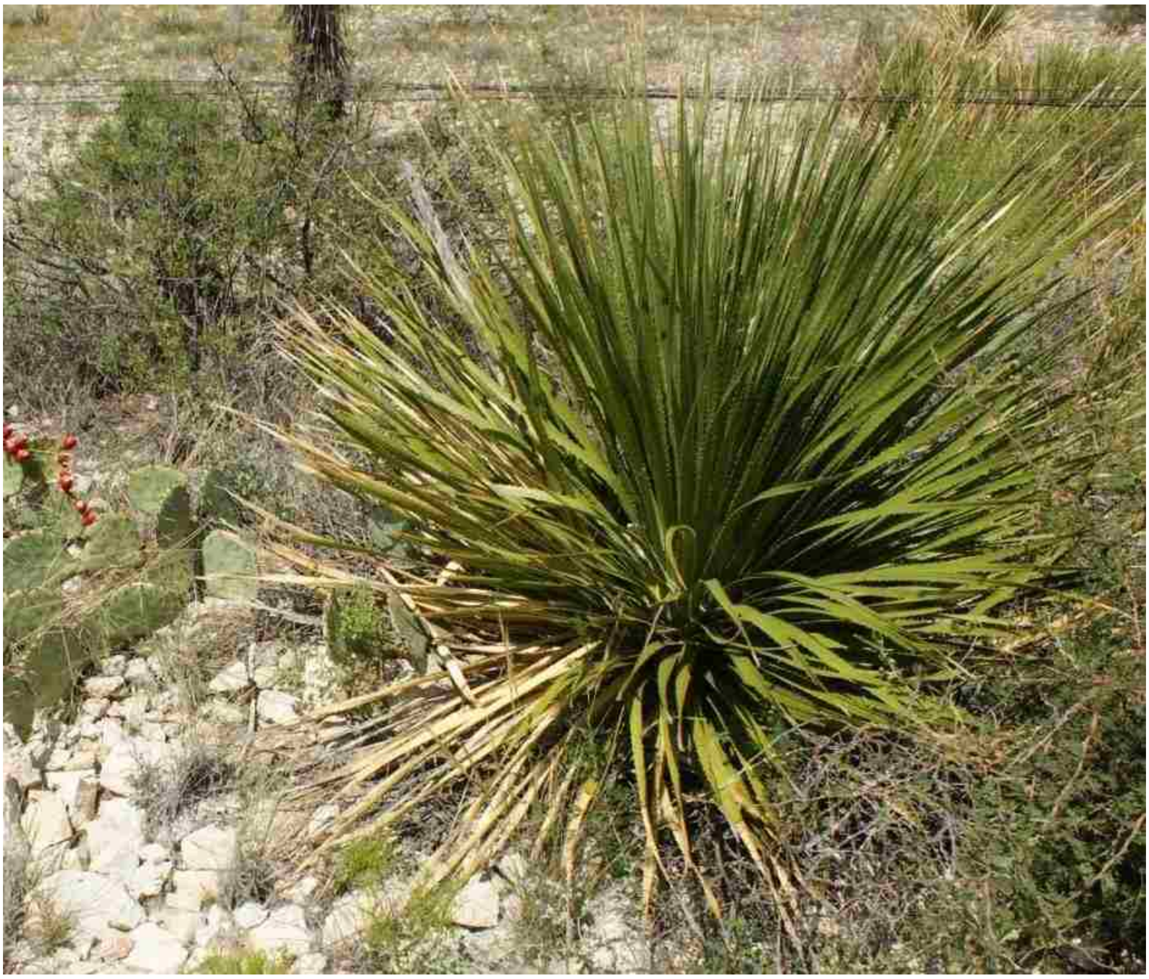


Figure 77: Sotol (*Dasylirion leiophyllum*). Documented in the project area by Dr. Ian Ralston, August 2012.

Final Model

Taking into account both paleoclimate change and the distribution of sotol a new distribution map was created. This involved stripping out the elevation weights from the previous model to account for paleoclimate, while sotol data was added in two ways. One was to use the DEM to create an aspect model. This model was then weighted so that south facing landscapes had a 50% increase in final score, east and west a 25% increase and north had none. In addition to that, the alluvium soil category was taken out of the soil dataset and added to the slope model as areas likely to contain the desired vegetation. The slope data itself was modified to eliminate slopes below 10%, as these areas were unlikely to contain the plants but not eliminating alluvium soils that overlapped with areas that were under the 10% slope. All of the datasets were finally recombined, resulting in the following distribution map:

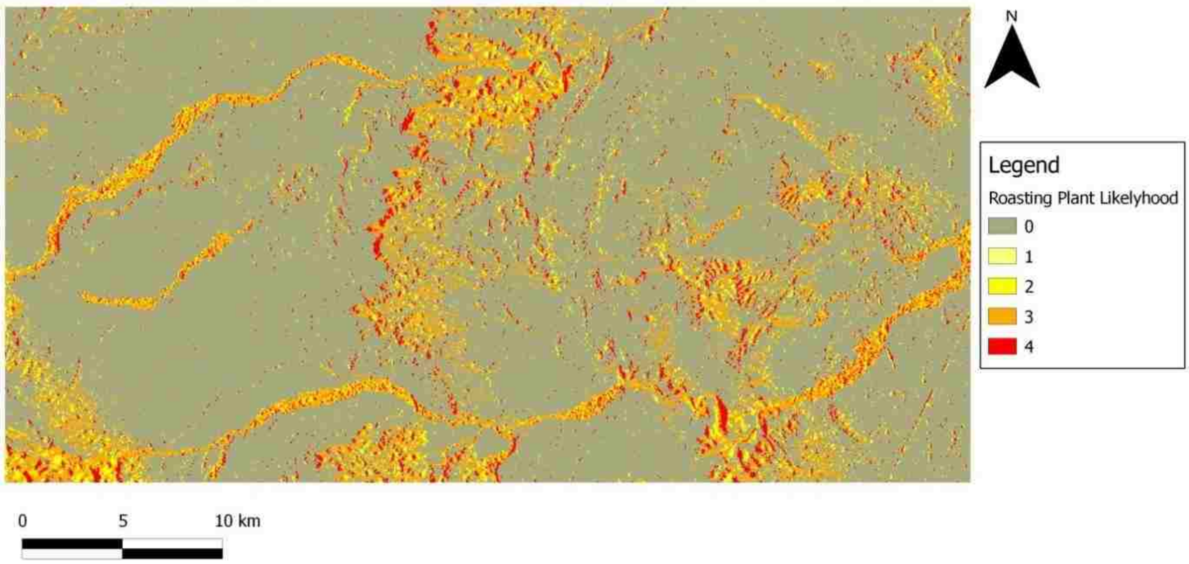


Figure 78: Areas with potential for roasting plants based on slope and soil. Red represents the highest likelihood for containing roasting plants. Grey almost zero likelihood of having the plants.

Again, this newer model did not result in the desired increase in performance when compared against known archaeological sites. But this problem was solved by replacing the soils dataset with the known water features dataset. This was to represent the stream beds that sotol was known to grow in. The early use of soil was determined to be too coarse a dataset. The resulting distribution greatly improved the performance of the model.

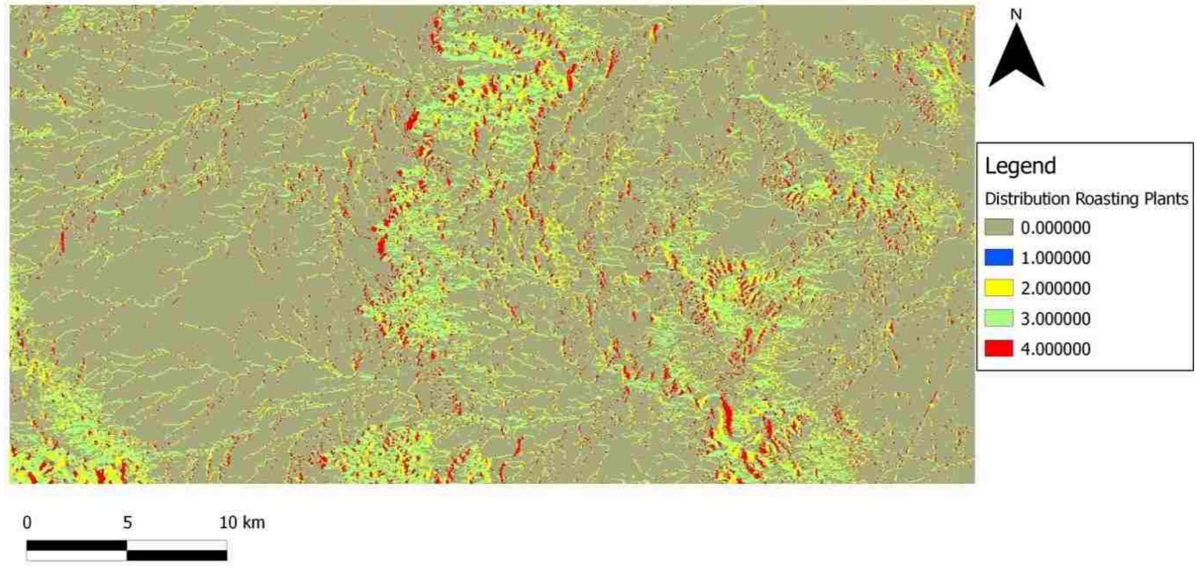


Figure 79: Distribution of potential areas with roasting plants with stream data. Red represents the highest likelihood for containing roasting plants. Grey almost zero likelihood of having the plants.

Maps in NetLogo

The next step was to find roasting sites, areas with low slope, that were located near these resources. The agave and sotol maps were then imported into NetLogo. Agents were sent out from low slope locations to see how many potential areas that might contain agave and sotol were within a set distance of that point. This distance was set at half an hour's walk, which gave total trip time of about one hour, in line with ethnographic accounts of foraging. The code used for this is a slightly modified version of the code used in Chapter 7 to determine full directional travel cost. The results of these models can be seen in Figure 80. This final predictive map correlates well with the known archaeological record (**See Confidential Appendix Figure 101**).

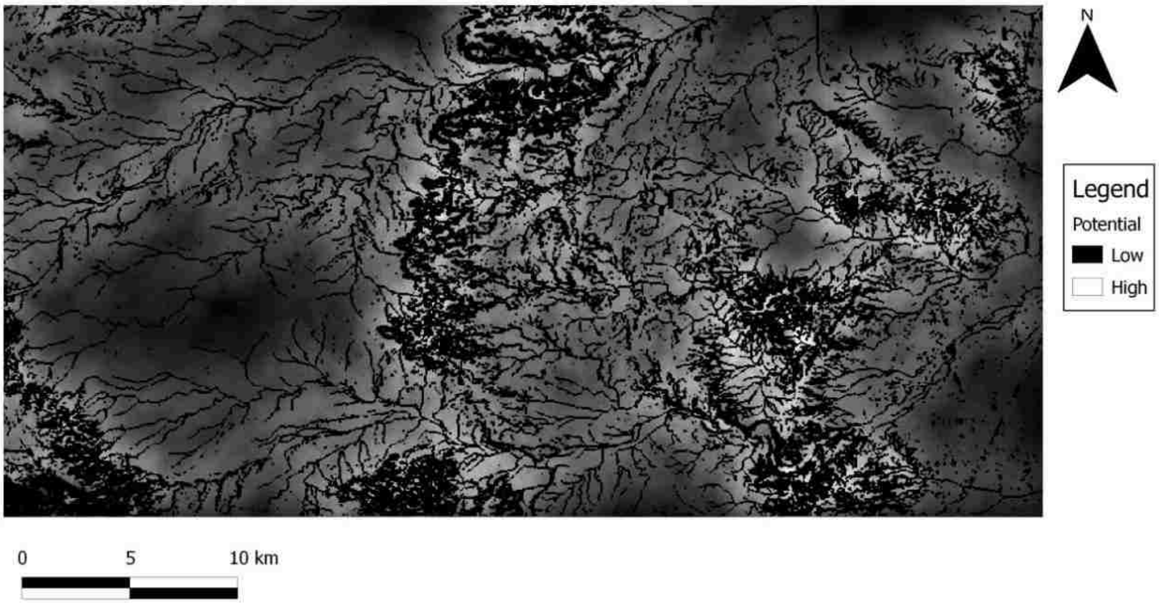


Figure 80: Areas for potential roasters.

Additional Attractors

Other attractors were considered for the project area but were eliminated for several different reasons. For example, large game hunting was considered as a possible attractor. Excavations in project area caves (Applegarth 1979) and in the surrounding areas (Katz and Katz 1985) have indicated a broad based subsistence strategy for those living in and near the mountains. Optimal foraging theory would dictate that priority should have been given to activities that return the highest caloric returns per hour of activity. According to research in the Trans-Pecos Region, which the study area is a part of, big game hunting has the best returns, significantly more calories per hour than agave and sotol roasting (Dering 1999).

However, when researching the potential areas where large game may have occurred it was found that big game hunting was unlikely in the project area. Big game found in this part of New Mexico, such as elk, deer, and American bison (buffalo), inhabit areas within several kilometres of water sources (mule deer (Brujes, Ballard et al. 2006, Esparza-Carlos, Laundre et al. 2011), elk (Truett 1996) buffalo (McHugh 1958, Truett 1996)). As Chapter 8 demonstrated, other than the Pecos River, water sources are highly unlikely to be present in the project area. There is the possibility that near Sitting Bull Falls and, depending on the intensity the yearly rain fall, Big Indian Spring may produce some water. Yet this left the vast majority of the project area as being unattractive to big game at any period of the year.

In addition to the lack of water in the project area, it was found that changes in the environment would further limit the availability of big game. Colder wetter climates are better for the growth of grasses while the drier and hotter climates are better for the desert scrub (McClaran and Van Devender 1995). Changes in climate are known to reduce the chances of finding large game associated with grasses (Parmenter and Van Devender 1995, Truett 1996). The archaeological evidence for bison hunting in the project area or even the surrounding areas is sparse (Hogan 2006). Historical evidence of these animals is also rare (Truett 1996). Finally, even during wetter periods sites near Big Indian Spring yielded no faunal remains of large game animals (Applegarth 1976). It may be those large game, which are heavily dependent on grasses for food stock and reliable water sources, could not have existed, and were not present in the project area for the periods of human occupation of the area.

Another activity considered was the hunting of rabbits, which had higher caloric rates of return than roasting plants (Dering 1999).Unfortunately, or fortunately if you were a prehistoric person, research into the habitat of the black-tailed jackrabbit (*Lepus Californicus*), found that these animals could range across a wide breadth of environmental zones. They could have been found in: desert shrublands; palouse, shortgrass, and mixed-grass prairies; desert grassland; open-canopy chaparral; oak (Quercus spp.) and pinyon-juniper (Pinus-Juniperus spp.) woodlands; and even low- to mid-elevation coniferous forests (Chapman, Dunn et al. 1982). These habitats cover the majority of the project area and thus it was not possible to limit the range of potential sites based on rabbit hunting. While this hunting or trapping probably did occur currently it is not possible to determine site location from it.

Additional site types and sites features were considered as another possibility however this was found to be problematic. A few sites have burials listed as some of the features but it was too few to draw any sort of conclusion about possible distribution. Mortors were other features found at sites but again too few to draw any conclusions from. Given the lack of excavations in the project area (Chapter 5) there were no other artefacts or features from which to try to investigate site patterns with. A lack of history in the project area also meant that cultural aspects like political boundaries could not be explored either.

Discussion

Despite the narrow field of attractors discussed the results were significant. 60% of sites could be accounted for by examining one single factor: optimal locations to place roasters. An additional 7% of sites could be accounted for by modelling locations that are likely to contain quarries, caves or rockshelters. This is a significant number of sites to be able to pinpoint while looking at only a few potential attractors.

More importantly, it was the quality of those sites captured that mattered the most. Caves, rockshelters and sites with roasting pits are all likely to contain dateable material. Because of that, they are likely to qualify for the National Register of Historical Places and possibly require excavation if a development were to be placed through the sites (Altschul et al. 2005). The remaining sites not captured by the roaster model and which were not a cave or rockshelter were very small lithic scatters (Table 29). Lithic scatters such as these rarely produce any sort of intact archaeological remains; most excavations of these sites end up removing the surface lithics and finding very little else (Hogan 2006).

Number of Lithics	Number of Sites	Percentage of Sites
0	20	6%
1	26	8%
10	153	49%
100	55	18%
1000	9	3%
10000	2 (both quarries)	
unknown	48	15%

Table 29: Prehistoric sites that do not have roasting features.

In essence, while not all sites are identified by these models, the ones most valued by cultural resource management were. These sites were also some of the rarer of the site types, caves and rockshelters, and are found in areas that would not have been expected in traditional predictive models. The method was able to demonstrate its value in terms of time and resource management in the planning stages of a CRM project. The implications of this and more detailed discussion will follow in the next chapter.

Chapter 9: Discussion and Conclusions

This thesis set out to improve the performance of site predictive modelling for CRM archaeology uses. To reach that aim the objective of increasing the explanatory abilities of predictive modelling was created, along with a plan of activities to reach that objective:

1. research causes of poor model performance and find cause to address;
2. create a methodology to solve the problem(s) that lead to poor model performance;
3. test proposed solution(s);
4. compare solution results against independent models to determine their effectiveness.

The first three activities were undertaken and reviewed in the preceding chapters. The groundwork was set for the last activity, testing performance, with the project area being specifically chosen so that the results could be compared against previous predictive modelling methods. So did the more explanatory model live up to its goal of improving model performance? The answer to that is more complicated than originally imagined, measuring Kvamme Gain Statistics for all sites between old and new predictive models.

Problem with Older Predictive Models

Initially this project undertook the comparison and found that the explanatory models performed better than the older models in terms of Kvamme's Gain Statistic, with some caveats which will be discussed later in this chapter, but problems developed. After this project had been completed in 2012, I set out to use some of the models in other research projects, specifically the water models. It was then that a serious flaw in the data used by the past predictive models was discovered. While the original PUMP III datasets were no longer available, it was possible to trace the images to semi-accurately recreate the dataset. A comparison of this data against USGS water data demonstrated clear differences between the two datasets (See Figure 55 and Figure 82). Satellite imagery (and physical spot checks undertaken in the summer of 2013) confirmed that the USGS data was accurate and that the PUMP III water systems were incorrect.

Projection Issues

There are issues with taking the data from the PUMP III report images. Figure 82 appears to indicate that there is a problem with different projections. Attempts were made to use different projections in GIS but were not successful in aligning the two datasets. An examination of Figure 56 in Chapter 7 shows that the water system does not match up perfectly with an image from the same report. It is suspected that some image editing occurred in the report creation that has changed the dimensions of the image and thus ruined the projections. However, the key point to take away from the image is that the water pathways are significantly different in shape, even with projection issues.

Figure 81: USGS quad map of Carlsbad West with hydrology features (blue).

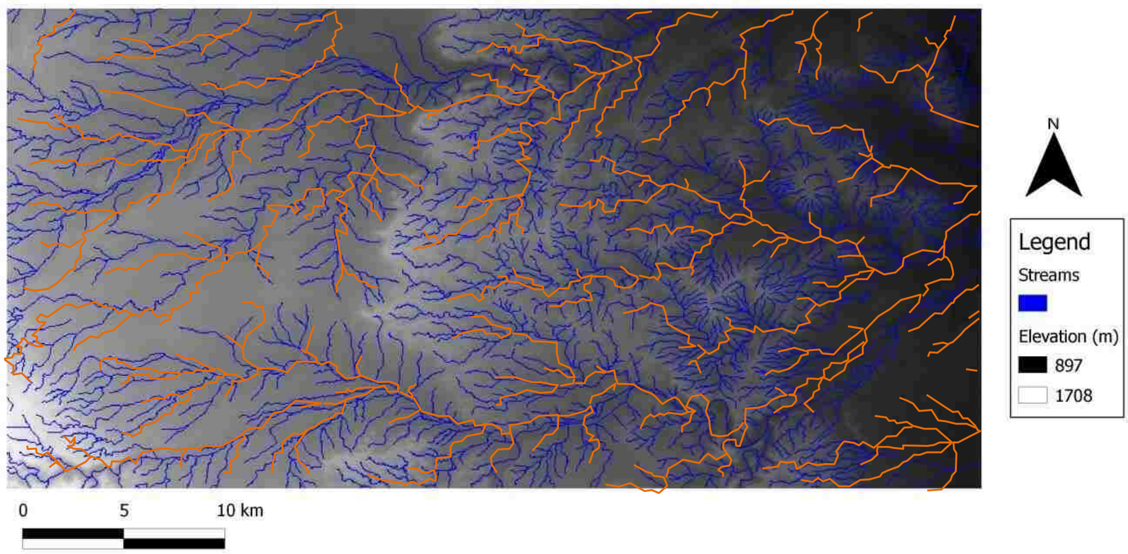


Figure 82: Water features in the project area from the USGS (2012) (blue) compared to PUMPIII water model (orange).

Why is the Data Wrong?

The GIS method of hydrology creation creates water systems by first using a ‘fill’ (terminology can vary between programs) to ‘fill’ in any small depressions or divots in the landscape; this is done with a raster DEM. Usually, this is undertaken by an interpolation program that ensures that each square is of a relative height to its neighbours. There are two reasons for this: one is to clean the data of any elevation errors that might exist in the dataset. The other reason is because the next step requires each square to feed into its neighbours, and depressions in the landscape interfere with this process.

After the dataset has been cleaned, each square is assigned an aspect or aspect-like number. The aspects are then used to assign numbers to a given square based on how many connected squares point at that particular square. The user then sets a threshold of how many ‘pointed connects’ a square should have before it is registered as containing a hydrology feature, flowing water. Theoretically the aspect and connection acts as a model for the direction in which runoff water would flow, and the threshold of connections sets a hypothetical point at which runoff would condense into a solid (liquid) body.

These techniques caused several problems. One is that this methodology assumes uniform connections and it fills in any depression in the landscape. That requirement was based on the system requirements of the GIS hydrology program and not tied to natural processes. A sink in the landscape would have caused water to pool at that location and not turn into the flowing linear features presented by GIS hydrology programs; a very likely scenario in the subject area, as the geomorphology report noted that the limestone in the area, ‘is karstic with several sinkhole depressions, especially in the western portion of the study area’ (Altschul et al. 2005 p. 81).

A further problem is the issue of edge effect. Because the method is based on cascading cells, those next to the edge of the dataset will have few cells to feed into each other. The real world areas that would have provided water flow to the edge are excluded from the dataset. Unless significant extra areas were included in the model the area is missing key data. This problem and the issue of ‘filling’ in parts of the DEM will create changes in stream location seen in the GIS waterways.

Implications of Poor Data

These problems explained the poor performance of the past predictive models and the ‘red flag’ sites observed. Proximity to the waterways was the highest weighted variable in the different PUMP III predictive models. A look at the ‘red flag’ sites finds that in the PUMP III model they are not near any waterway but an examination of the USGS data shows that in fact they are. Poor data has significantly affected the model’s performance.

With none of the original data available from the PUMP III models it was not possible to recreate the older predictive models with updated drainage data. However, using the drainage dataset from the USGS (discussed in Chapter 8) it was possible to create a traditional predictive model by placing buffers, capturing the area within a certain distance of the drainages. The results of the new model presented some very interesting findings. One is that good data is very important to data creation. The new simple model was able to double the Kvamme Gain Statistic of the PUMP III Boolean and Regression models and improve on the weighted model as well (Table 30). The 'red flag' sites mentioned in the PUMP III report were identified by this simple model as well. An adage popular among modellers is 'garbage in, garbage out'.

Method	Precision	Accuracy with site point data	Gain	Accuracy with site shape data	Gain
Buffer 100m	30%	37%	0.21	53%	0.44
Buffer 150m	42%	51%	0.18	66%	0.36
Buffer 200m	71%	81%	0.12	85%	0.17
Buffer 300m	81%	90%	0.09	92%	0.11
Buffer 400m	88%	93%	0.05	95%	0.07
PUMP III Boolean	46%			58%	0.21
PUMP III Weighed	43%			70%	0.39
PUMP III Regression	63%			85%	0.26

Table 30: Results of traditional models in the project area from this project (Buffered "streams") and the PUMP III project.

The second finding is that even with better data the results are not improved that much. The doubling of results is impressive, but less so when the bar was set so low. Moreover, it only captured 40-50% of the known sites before the performance collapsed down to the same level as the previous models. This greatly increases the chances of 'gross error', an unwelcomed outcome. The problem with this simple model, and the PUMP III models, is that it assumes that there is a correlation between sites and the streams. Which as presented by statistical analysis in the PUMP III project, and the results of the simple model here, there is.

The flaw with these methods is that there is the assumption that this relationship is direct and fairly linear. As demonstrated through this project it is most likely that this relationship is secondary with the real relationship being with agave and sotol resources. Figure 83 shows that this project's roaster model is able to capture those sites that do not appear to be correlated with the drainages and those that are as well, because it is capturing the correct relationship between sites and plant resources. The missing of a few sites in the image is due to using points to represent sites and the resolution of the data.

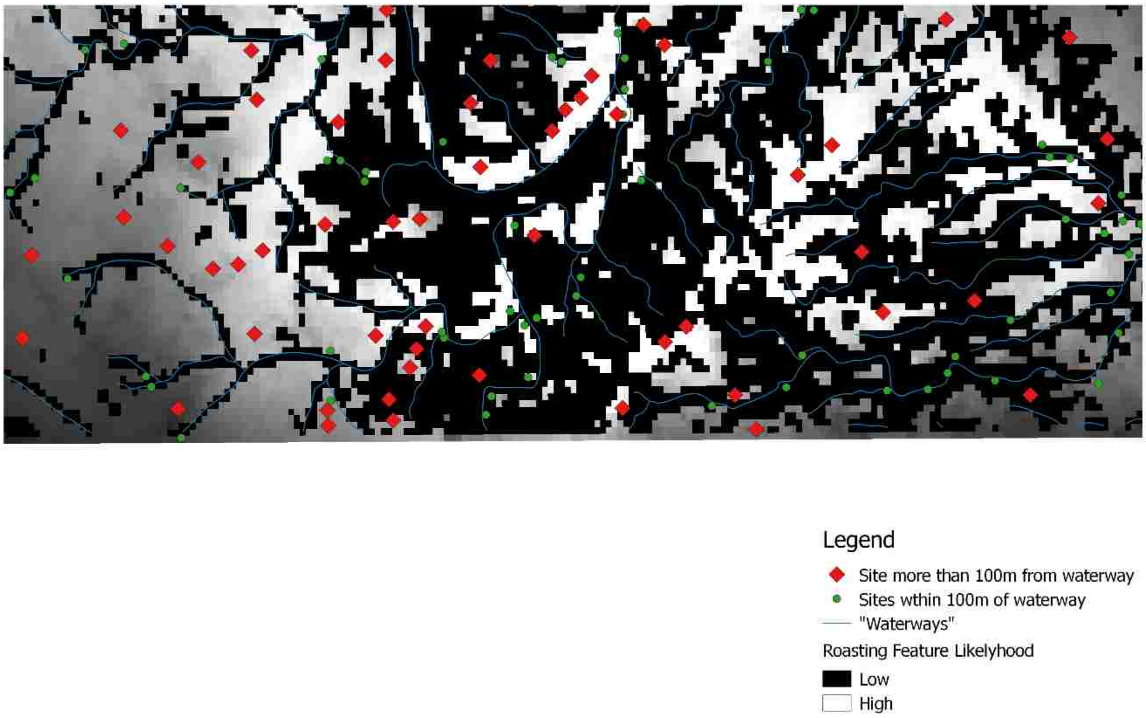


Figure 83: A section of the project area with sites within 100m of "waterways" and those further away with the roaster potential as a backdrop showing most sites not captured by distance to water is captured by roaster potential.

(See Figure 102 in Confidential Appendix for better presentation of the image.)

Was This the Right Performance?

Modelling for agave and sotol improved the performance of the predictive modelling in the project area, into the high 50s of the Gain Statistic depending on how the final results were adjusted. With a few tweaks to the model 71% of the sites could be captured with the labelling of 34% of the project area for a gain value of 0.52 or 29% and 68% for a gain of 0.57. That shows, at least, a roughly 20% increase in performance for the explanatory model over the other models. But it was this constant tinkering to improve the Gain Statistic that raised serious questions about how well this project measured performance.

It was anticipated at the beginning of this project that this goal would be measured by a simple metric, the Kvamme Gain Statistic. A project area was specifically chosen that had previous predictive models created by others that had this measurement. This allowed for the testing of an independent traditional method model's performance against the models created for this project without the worry of the author's skills/bias affecting the outcomes. The Kvamme Gain Statistic would have acted as an easy-to-compare metric of the performance between the two models.

However, it is now my belief that, after undertaking this project, reducing outcomes to a single number was not the best approach to measure predictive modelling results. Similar to some of the issues raised in the past about predictive modelling: 'functional, temporal, or cultural site types cannot be readily determined for most sites in an archaeological database, yet profound locational differences must exist between the types'; 'Grouping sites of many types into a single, site-present class creates too much variability to model' (Kvamme 2006 p. 6). The earlier criticism was based on the assumption that putting all different types of sites into a model as a single dataset would reduce performance. I would argue that this criticism has not been taken far enough and should include the presentation and measurement of results.

There has been significant effort to find methods to measure predictive model performance beyond the Kvamme Gain Statistic. Kvamme (2006) has published a list of different measurements of predictive modelling performance:

Statistic	Derivation	Interpretation
p(S)	From survey data: (total area of site class)/ (total area field surveyed)	Base rate or chance probability of archaeological site class in study region
p(M)	Determined exactly by GIS: (total area of model)/(total study area)	Base rate or chance probability that a model will indicate a site; proportion of study region mapped to M
p(M')	1−p(M)	Model precision; high values indicate high precision
p(M)/p(S)	Ratio	Model fit; indicates how many times larger a model mapping is than the total site-class area
p(M|S)	Estimated by proportion of known archaeological sites correctly specified by model	Model accuracy; probability that a model will correctly indicate a site: 100×p(M|S)= percent correct
p(S|M)	Estimated by proportion of locations in M that contain archaeological sites	Probability of archaeological site presence when model specifies a site
p(S|M')	Estimated by proportion of locations in *M'* that contain archaeological sites	Probability of archaeological site presence when model does *not* specify a site
p(M|S)−p(M)	Subtraction	Improvement that model offers over chance in specifying known archaeological sites
p(S|M)/p(S)	Ratio	Model improvement ratio; indicates how many times more likely a site is in M than the base-rate site probability
p(S)/p(S|M')	Ratio	Model improvement ratio; indicates how many times less likely a site is in M' than the base-rate site probability
p(S|M)/p(S|M')	Ratio	Model improvement ratio; indicates how many times more likely a site is in M versus M'

Table 31: List of different measurements of model performance.

Verhagen (2007c) has also listed multiple methods, many listed by others. There is the Attmell-Fletcher test (Kamermans 2006), which is similar to the author's max gain and Verhagen (2007c). Also, there is the S statistic (Altschul et al. 2004). But, regardless of the methods used to measure performance they almost always focus on measuring all sites, regardless of differences between them, with few exceptions (see Figure 84 for an example).

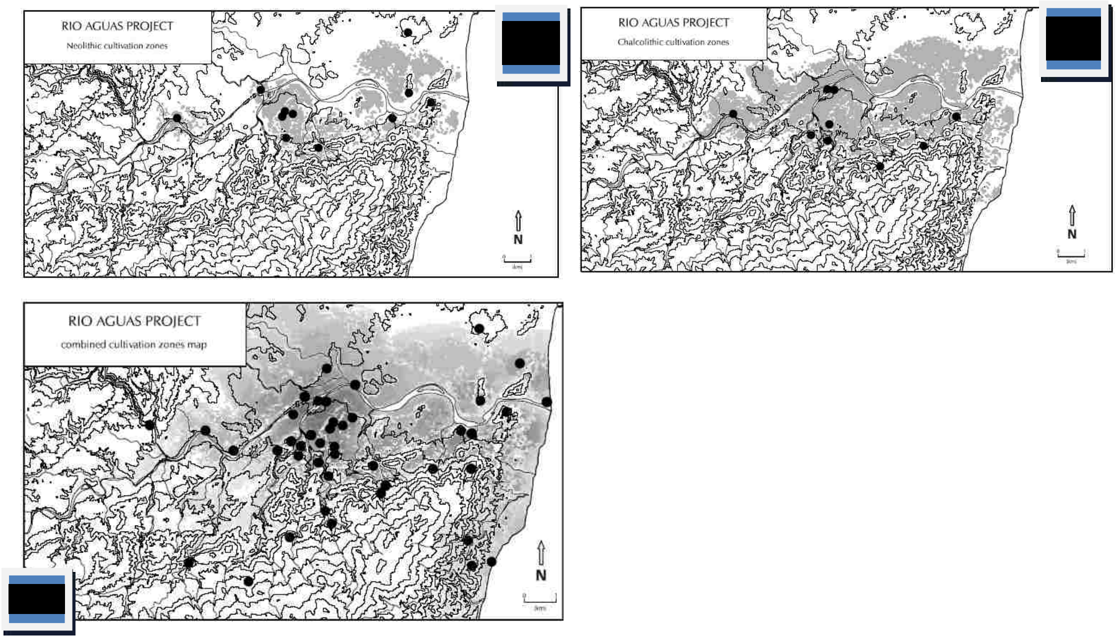


Figure 84**:** Predicted distribution of sites by time period: A - Neolithic; B - Chalcolithic; C - all periods, from Gili et al. (2007) p. 181-2 Figures 8.5(A) & 8.6(B) & p. 190 Figure 8.17 (C).

Yet the separation of site types is a key activity for, and has a significant impact on, CRM archaeology, because not all sites are created equal in terms of the law and resources to manage them. The time and resources required to document and/or excavate a small lithic scatter consisting of a few flakes on the surface and no buried deposits is significantly different from a site that has a hundred roasting features and many buried artefacts. For project planning in CRM it is not just the number of sites one will encounter but the quality of those sites as well. As pointed out by others, at least in the US, predictive models are most helpful if they can predict which sites might qualify for inclusion in the National Registry and thus require more investigation (Altschul et al. 2004).

The models created for this project showed several types of sites, caves and overhangs, quarries and roasting sites, that are connected and complementary but, in many areas, mutually exclusive, i.e. roasting features won't be on cliff sides but caves and overhangs will. Combining these different datasets resulted in very poor quality predictive models because of this lack of overlap. Kvamme's Gain Statistic is based on labelling the least amount of area likely to contain sites. A predictive model that labels a majority of the project area as likely to contain sites will perform poorly in terms of the Gain Statistic.

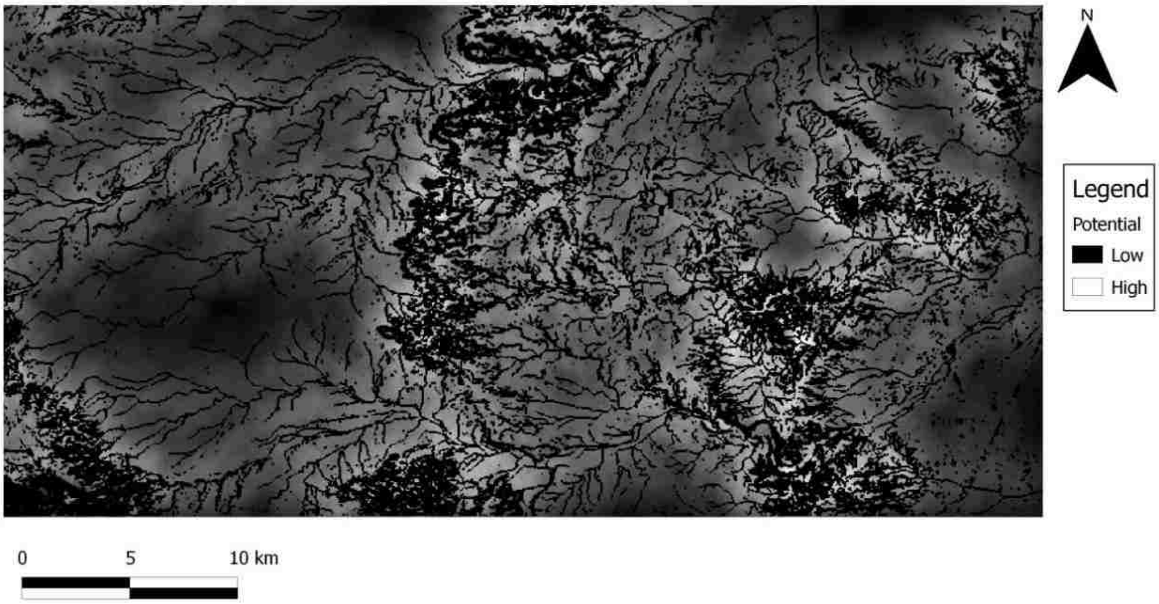


Figure 85: Model of areas likely to contain roasting features.

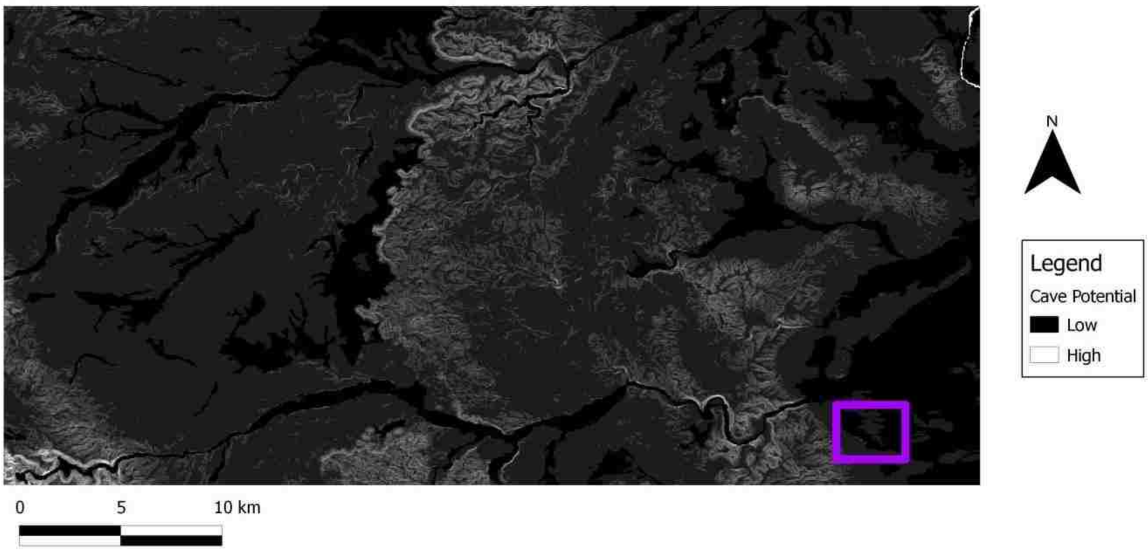


Figure 86: Model of areas likely to contain caves and rock shelters.

Taking each behaviour model/site type on its own provides more information for CRM managers. The roasters model accounts for the placement of the majority of sites in the project area. For CRM workers this can serve as a guide for where most sites are likely to be found and allow them to plan accordingly. The quarry and cave models represent only a fraction of the sites likely to be encountered, about 6%, but also ones that are most likely to be the most expensive to excavate. The quarries in the project area can have hundreds of thousands of lithics that would need to be recorded (Appendix D). Caves have the potential to have buried deposits and if excavations in the surrounding areas are any indication they will require expensive preservation of organic materials and analysis of human and/or animal remains.

From the perspective of a CRM manager, this information can help better estimate costs for projects. They will know what type of site is likely to be located in which locations. They would know that they will most likely encounter roasting features on flat surfaces near drainages and hills. If their project takes them near hills and limestone outcrops they should anticipate quarries and caves/shelters but only a few. This is information that can be conveyed to clients before work begins to shape their expectations and to alert them to the risks. It also gives the clients the ability to decide if they want to develop areas with high likelihood of smaller sites or areas with less likelihood of sites but ones which have greater cost potential. This result is exactly the type of outcome that CRM managers have been asking for (Altschul et al. 2004, Altschul et al. 2005).

Wrong Project Area for Single Gain Statistic?

If the simulations are correct than the project area is disposed to not work well with a single Gain Statistic. A significant portion of the project area has agave and sotol growing potential. Moreover, those areas with less potential for roasting features are areas that are likely to contain caves and rock shelters. There are two different location strategies that do not overlap significantly with each other. Finally, there is not a strong attractor like water resources to draw people to limited areas of use. The project area has a dispersed use pattern which makes it very hard to increase the precision of the models. Even under the best scenario the performance of the model capturing all sites is unlikely to reach the .7 or .8 gain value of some 'high performing' models.

What is the Right Measurement?

This project started out with the vision that models could simply be measured using the Kvamme Gain Statistic and the results would speak for themselves. It did show that the explanatory models could improve performance by about 20%, however, that metric does not accurately capture the quality of the outputs for CRM purposes, site type, and is unlikely to work well in the project area with such wide difference in locations between site types. Reducing the results to one number obscures the complexity of the archaeological record. Predictive models should discern site time when presenting results, using Gain Statistic, Attmell-Fletcher test, or any other measurement. In the case of the models created in this project the agave model had a Gain Statistic of the high 0.5s, depending on how it is tweaked. The cave and rock shelter model had higher performance in the 0.6s, again the results can be tweaked up or down. As will be discussed later in this chapter it may be possible to increase the performance of this model with better data. This range of performances better reflects the archaeological record and meets the needs of CRM managers.

Did the Model Improve Our Understanding of the Past?

As discussed in Chapter 2, this project chose researching explanatory-based predictive models as it had the added benefit of potentially being of greater interest to academic archaeologists, because they could be used to explore past human behaviours. So did the models add to the archaeological understanding of the project area and thus broaden the appeal of predictive modelling? The results show a clear yes. These models helped answer some of the questions that have been raised about the area:

> 'What is particularly interesting about the project area, however, is that much of south-eastern New Mexico does not conform to the traditional image of Formative cultures in the American Southwest. Rather than being based predominantly on corn, beans, and squash, much of the post-Archaic subsistence intensification in this region seems to have been based on agave and shin oak, and in many ways it appears to have been a continuation of an otherwise largely Archaic lifeway.' (Altschul et al. 2005 p. 10)

The question about change in subsistence from the Archaic (hunter-gatherer) to the Formative/Pueblo/Ceramic period (traditionally farming and sedentary living) has been raised in the regional research framework (see Hogan (2006)) as a topic that needs investigation. The prevailing theory is that during the Ceramic period the region was occupied by both farmers and hunter-gatherers, with hunter-gatherers located in the project area away from the Pecos River (Larralde and Sebastian 1989). This is based on excavations of roasters that have Ceramic period C14 dates. Because of this theory the regional research framework asked questions about the subsistence patterns of these 'ceramic period hunter-gathers':

- 'What wild plant and animal resources were exploited by Ceramic period groups in each of these areas?'
- 'Which resources were dietary staples?'
- 'What seasonal variability is there in the availability of those resources?'
- 'Were any food resources stored?'
- 'Could the available mix of food resources in the area support a year round occupation, or would resources in other areas have to be exploited at some point in the annual round?'

(Hogan 2006)

The models for this project help to explain why we find both agriculture and hunter-gatherer type subsistence strategies alongside each other during the Ceramic period and can aid in the process of creating archaeological theory. The water model and paleoclimate data show that for the majority of the cultural history of the area there was no perennial water. Potentially, there would be seasonal streams fed by springs when enough rain fell to raise the water table. Otherwise the whole area could not support long-term habitation. People could only travel into the area for brief periods of time.

The roasting model shows that the majority of sites in the project area are associated with roasting activities. It is possible that agriculturalists or semi-agriculturalists were coming into the area during the Ceramic period to exploit the resources then leaving. Agave and sotol could have provided an alternative food source to diversify the diet. Excavations of sites in the region along river terraces dating from the Ceramic period indicate that the people would hunt and gather whatever resources were immediately available to them (Hogan 2006) and this fits the general pattern of the model. Agave and sotol are not just food resources; the leaves are used for producing mats, sandals, baskets, cloths, etc. People could have been collecting agave for other reasons, and using agave and sotol for food was a secondary activity. The different models created for this project help us understand that instead of having two different groups of people during the Ceramic period we have one group, who were able to exploit agave and sotol alongside their other activities. This is not to say that they did not also practise hunting-gathering at the same time. However, the project area would not support separate groups of people undertaking different subsistence strategies living next to each other year-round.

Returning to the regional research framework, the scenario created by the models makes it possible to answer part or all of the questions asked about subsistence in this region:

> 'What wild plant and animal resources were exploited by Ceramic period groups in each of these areas?' (Hogan 2006)

For the project area we can say that agave and sotol roasting were some of the resources exploited. Moreover, because deer, elk and buffalo require a water supply relatively close to where they roam, 5–10km, it is possible to exclude those resources for parts of the project area based on the water model.

> 'Could the available mix of food resources in the area support a year-round occupation, or would resources in other areas have to be exploited at some point in the annual round?' (Hogan 2006)

The models show that in most of the project area there would not be year-round occupation, except near the Pecos River and areas with reliable water sources. Based on the model and ethnographic research it would appear that food resources were exploited on an annual rotation or only occasionally.

Additional research and work will be required to confirm these results. Still, the models have created a plausible explanation to answer some of the questions from the regional research framework. Something that the previous predictive models were unable to do –

> ‘As the preceding discussion indicates, correlative predictive models may allow us to discern patterns in settlement. They do not explain such patterns ...’ (Altschul et al. 2005 p. 97)

For this project the models were able to broaden their use beyond that of a basic CRM utility to find sites. In doing so it met the goals of trying to expand such modelling outside its narrow field of use.

Such positive outcomes need to be tempered with the knowledge that this is only a case study. This methodology has been demonstrated but further tests in different settings need to confirm the utility of using simulations for predictive modelling. Even then the full impact of this research will not be felt until such methods garner wider use in archaeology. At the moment this project stands as a positive first step in that process.

Did This Actually Increase Our Understanding?

ABM enables the exploration of ideas about the past, but that means it is not modelling the past itself but rather ideas about the past (Premo 2008). A simulation does not tell archaeologists how things actually were, even when a simulation’s results perfectly match the archaeological record. A match is not conclusive proof that that is how it happened in the past (Premo 2008, Premo 2010). It is worth reiterating these points as this review has discussed the knowledge gain as though the simulations prove what happened. As discussed earlier in this thesis, that is archaeology. We take the best evidence we have about the past and create narratives about what we think happened. As new evidence comes to light we change those narratives. Barring the creation of a time machine we will never know if we are 100% correct, but we move forward in the hope that were are at least close. So the results of this project indicate outcomes with the best evidence possible using the best tools available. Only more research will confirm or refute the theories created through simulations in this project.

Unexpected Outcomes and Observations

While this project started out with a very specific goal of improving model performance, it also contributed to other areas of research. What follows are some of those observed contributions outside of the stated goals and objectives.

Least Cost Path

ABM modelling of least-cost path analysis added several new insights into the practice in archaeology:

- The model provided a new tool for exploring travel costs, not limited to the single or few directional costs that can be modelled with most GIS programs. An agent-based model gives the ability to easily and quickly calculate 360 degree directional forces that change with each action taken by a person (agent). This is an attribute that may be of use to those investigating travel costs. For example, it could be used in site catchment analysis to produce an examination of resources within walking distance as demonstrated in Figure 70.

- Testing of the models found many issues revolving around interpolation and how NetLogo handles the DEM data. This indicated that interpolation will affect the least-cost path results.
- Problems were found when trying to interpret how a past person would view 'costs'. Most notable of these interpreting 'cost' issues was the combined viewshed and least-cost path analysis model, which showed that paths can dramatically change based on what an agent has knowledge of.
- A further issue, and one that questions the basic assumptions of least-cost paths, found that different paths taken by a person could diverge by several kilometres but arrive at the same location at almost the same time, negating cost as a determinant of paths. If costs are so small as to not be noticeable, then least-cost path loses its advantage.

Overall, this project resolved one of the problems with least-cost paths, single direction cost modelling, but raised many more problems, including one that questions the fundamental assumptions of such models, that least-cost routes matter. If multiple, significantly divergent, paths can lead to the same point with very similar costs then what is the utility of such a method? Archaeology needs to reconsider its use of least-cost paths as they probably are not demonstrating what most archaeologists are assuming: that people take the quickest/less calories burned/some other type of cost-saving path, and thus that the least-cost method shows the exact route taken.

Inaccurate Data

In Chapter 2 the different range of problems facing predictive modelling was reviewed and several of those that were not focused on still played a role in this project, such as 'GIS data are inaccurate' (Kvamme 2006 p. 6). As discussed in Chapter 8, some of the soil data used was found to be incorrectly labelled for soil depth. This project just adds one more example of the need to be aware of data quality as heavily discussed in any modelling literature or most literature that deals with data. Furthermore, this project demonstrated how ABM and GIS interpret the data, convert data or model actions that can dramatically change the outcomes. Chapter 6 reviewed tests undertaken by this project to see if the NetLogo program works and changes outcomes. Most predictive models do not undertake such work, or at least it is not reported. This raises a concern that not only can GIS data be inaccurate but the presentation of such data can alter the accuracy as well.

Blue Line Features

'Blue-line features on topographic maps are frequently arbitrary and unreliable indicators of water' (Kvamme 2006 p. 6). This project has demonstrated that it is not just maps that may have issues but that attempts to create water datasets using GIS tools have significant flaws. This will need to be replicated in different project areas but anyone using such datasets should re-evaluate them. As the water simulation confirmed that blue-line features on maps should be used with a sceptical eye to the intensity of water resources they represent, many will vary greatly in the amount of water present in them at any given time.

This project set out to improve predictive modelling performance, which it did, just not in the role envisioned; in the process it found serious issues that may affect model performance. Certainty the issue of 'blue lines' impacted the performance of past models. There is clearly more work that needs to be done to tackle the different problems with predictive modelling. These were just a few examples encountered in this project but the list from Chapter 2 needs to be more fully investigated.

Data Resolution

'GIS data have insufficient resolution and poorly represent the real world.' (Kvamme 2006 p. 6). This was dismissed as a concern at the beginning of this project but experiences have since shown that to be a wrong assumption, data resolution does matter. The New Mexico Bureau of Geology and Mineral Resources is in the process of creating 1:24000 geology maps of the state (Resources 2012). At the time of this project only a small percentage of the state was covered; however, that included the Carlsbad West Quad. The data from this map shows that only a small percentage of the geology actually contains the potential for chert or chalcedony concentrations as compared to the general geological formation information provided by the USGS (Figure 87). Potentially, if such a detailed dataset were applied to the whole project area it might be possible to significantly reduce the areas that are predicted to contain lithic resources.

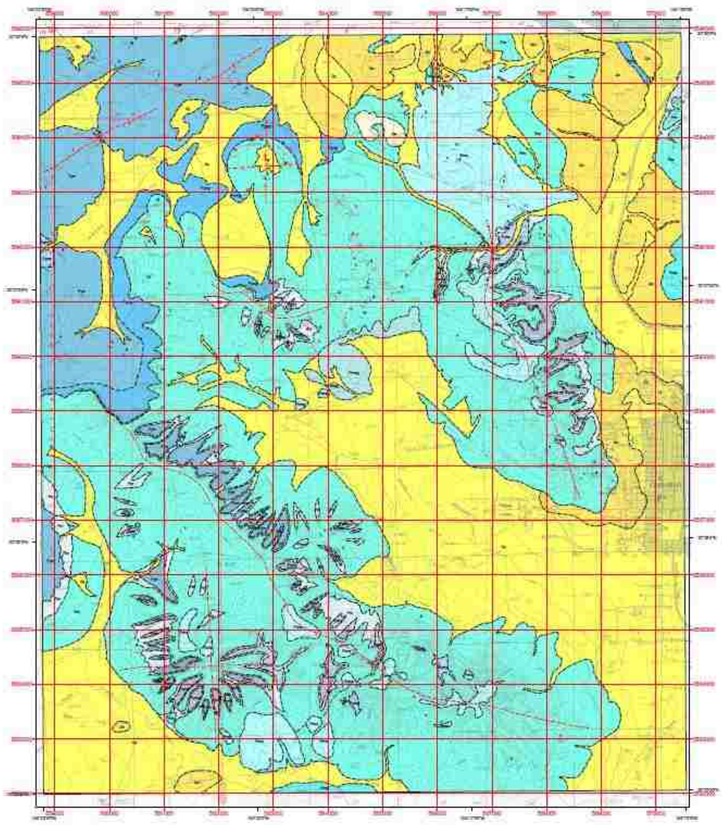


Figure 87: 1:24000 geological map for Carlsbad West Quad (Resources 2012).

The resolution of the DEM could have also affected the outcomes. The 30-metre resolution of the DEM used and how different software programs interpolate this data obscures the finer details of the landscape. Site visits to various locations in the project area for this project recorded multiple sheer faces of exposed bedrock (Figure 88). These geological formations would have been excellent locations to find exposed seams of lithic materials. Yet these formations, due to how GIS interprets the data, did not show up in the digital data available. Targeting only these formations could cut down the potential quarry and cave areas to a fraction of the model's current coverage.

Figure 88: South canyon side of Dark/Last Chance Canyon.

Did this project Choose the Right Goal?

Returning to Ebert's criticism, after undertaking this project, '... sixty to seventy per cent is not really bad but it is not very good either – certainly not good enough to justify spending a lot of money' (Ebert 2000 p. 142), I now question if focusing on the problem of model performance, specifically in terms of accuracy and precision, was the right choice. Of course, with the literature available at the beginning of this project this seemed like a logical problem to tackle. As discussed above, it was only through the exploration of model performance that issues arose with using a ratio of accuracy and precision to measure performance, the conclusion was reached that this is inadequate for the needs of CRM archaeologists.

Possibly more significant is Ebert's assumption that such performance is not enough to justify spending a lot of money. While the Minnesota and North Carolina models cost $6 million each to make, the vast majority of those costs came in obtaining the data, specifically, digitising the necessary data (Cole et al. 2006, MDOT 2009). Similarly, a little over 10 years ago the PUMP III project had to use aerial photography to create digital soil data, which is why it had a more than million dollar budget (Altschul et al. 2005). However, now most archaeological records are digitised, and increasingly so are most other datasets as discovered during the process of this project, and able to obtain DEM, soil, water and plant data for free and in higher resolution than was available to the PUMP III project. The cost of such data will vary between countries as not all countries make their data open as the US does. But if the Minnesota, North Carolina and PUMP III projects were to be conducted today they might only cost a few thousand dollars.

As accuracy of data increases, it is now possible to get sub-metre point cloud data of landscapes, and costs of obtaining such data continue to fall; drones can now collect LiDAR data, and the costs associated with creating predictive models will almost become a non-factor. One could probably spend an afternoon obtaining data from various websites and then working with a GIS program to create a fairly decent predictive model. Which raises the question, even if the accuracy and precision performance is poor, a dubious measurement, does it matter if the model cost next to nothing to create?

With costs being limited to the labour required to work a GIS program or other software there may be the potential for CRM archaeologists to focus on developing more explanatory models. Spending a few days attempting to better understand the placement of certain sites may be worth the cost for a better model for CRM purposes, i.e. a model that differentiates site types, something that would cost significantly less than the million dollars spent on other predictive models. It could be that we may be entering an era of explanatory-based predictive models, driven by CRM funding freed up from the need to obtain data.

Observations on ABM as a Method

Is there a role for agent-based modelling in explanatory predictive models? Based on the results of this project I believe so, at least the idea of more complex simulations being used to guide model creation. They may not necessarily take the form of agent-based modelling, sometimes mathematical modelling is more appropriate for simulations. As demonstrated by Verhagen and Whitley (2011) it is also possible to use GIS but with an agent-like construction.

The agent-based modelling allowed for the exploration of key concepts at different stages of the project. At the beginning ABM made it possible to examine ideas of least cost path analysis, adding to a growing literature on the problems with least cost paths. Critically, it allowed for the creation of a model that could show that multiple diverging paths could result in almost the same costs. Could this have been accomplished with GIS? Possibly, with a custom plugin designed to measure travel costs along a route but it would not have been an elegant system and without the ability to contribute to more complex models.

ABM really demonstrated its usefulness with the water model. GIS methods of water analysis were shown to be flawed, as the PUMP III models demonstrated. In this case, GIS is not currently able to replicate what ABM was able to accomplish. GIS programs did not have the tools to calculate absorption rates and such absorption rates' interaction with the movement of water. Understanding water systems set the stage for much of the theory developed about the project area.

GIS or ABM?

These observations of ABMs abilities are not to say it is superior to GIS in all ways. Each are separate programs that have their strengths. Much of the work on attractors involved GIS because it was the better tool for the modelling. As discussed in Chapter 4 some GIS programs have ABM plugins. It may be at some point that ABM and GIS become merged into a single computer program. While this project set out to test ABM this was only in service to the main goals of the project. The best tools for explanatory predictive modelling were going to be used, regardless if that was ABM or something else.

More complex models?

The ABM models created for this project were self-contained models, each focusing on a specific theory about the project area. Why were the models not combined together into a more complex model? Because there was no need. Each model addressed a specific theory that was different in its simulation needs. A person walking across the project area had no bearing on the absorption rates of soils. It made little sense to combine the models into a larger model when the individual simulations provided the required explanation about travel, water systems and attractors.

Yet the real potential of ABM may best be exploited in more complex models. After exploring the different avenues for site creation, a model utilising aspects of the resource gathering and travel simulations was created. This model combined viewsheds and travel paths. Past projects have combined least-cost paths and viewsheds to model in GIS what people would see as they travelled (see Harris and Lock (2006)). Yet these modelling attempts have been very static. A path is created first and then viewsheds are taken from points along that path. What people would have seen along that path does not influence their decision on the route they take.

ABM modelling allows the modelling of what a person sees to influence the route they take. For example, an agent could represent a person looking to reach a far-off destination but be travelling through an area they are not familiar with. Thus instead of assuming perfect knowledge of costs of travel, as happens in GIS, the agent decisions could be based on the knowledge available to them. They could end up travelling a less than optimal route because they are unware they will eventually run into thick vegetation or an unscalable cliff. Such a model was created post project.

The walking model was improved with a viewshed component. The agent would run a viewshed analysis from where it was positioned. Then only the areas that were visible were taken into account when running the cost path analysis. Based on that analysis, with imperfect knowledge of the project area – only what could be seen or had been seen, the agent would move on the first step of the path. At which point the process repeats itself with the agent looking around and reassessing the best route as new information is made available. Similar to how most travellers would plan their trips, a series of micro changes take account of changing environments. This creates a very different travel path (Figure 89).

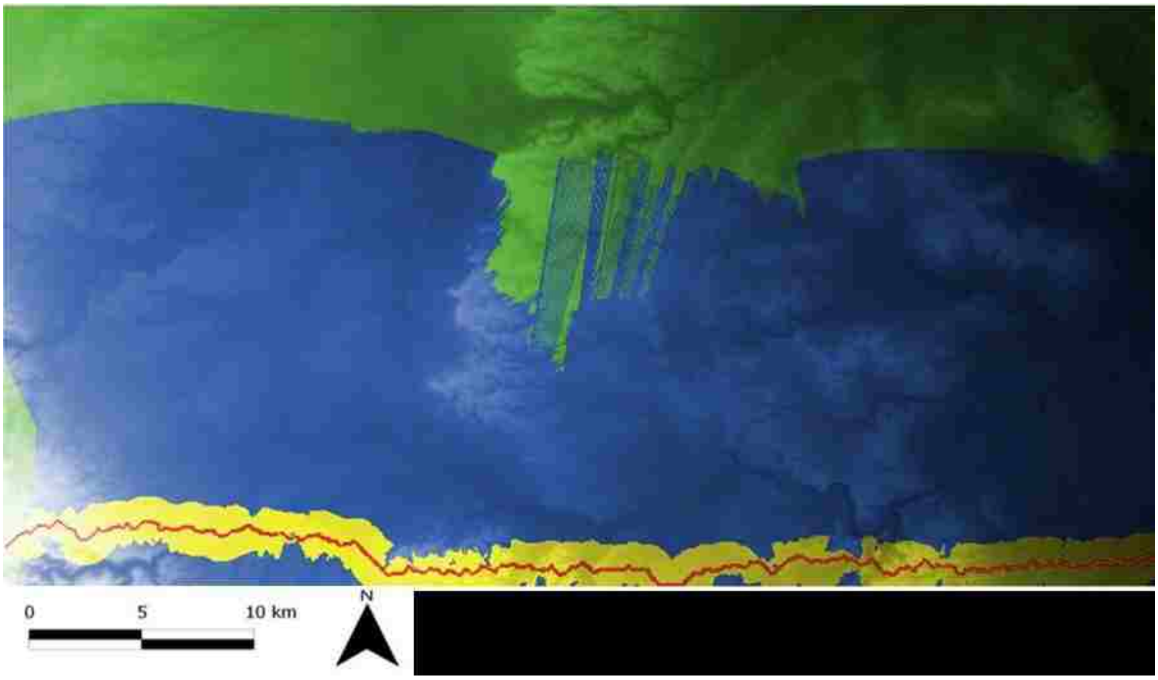


Figure 89: Higuchi viewsheds: (red) short distance (trees are recognised as individual entities with leaves and branches; distance of 60× tallest plant), (yellow) middle distance (the outline of trees are visible but not individual trees, 1,100× tallest plant) and (blue) long distance (can tell vegetation is present and little else, max distance 18km).

This only begins to scratch the surface of what such a model can do. Adding a Higuchi viewshed (see Figure 89 caption for detail) one could create a map of how one sees the landscape based on plants. In the example given for the project area (Figure 89), the red represents areas on the person's journey where they could distinguish individual leaves on a sotol plant and the yellow is areas where they can distinguish individual plants, while the blue shows areas of vegetation where they cannot identify individual species but their vision shows a blending of vegetation. This type of model then starts to move in phenomenological approaches to landscapes. It is now modelling what people would have experienced, at least visually, as they moved throughout the landscape. Combine this model with the water model and then you have a model of how past people would have seen a rain storm on a journey throughout the area.

While that is interesting, it adds very little to the goals of the project and that is why it was not further explored here. It is an answer to different research questions. It does demonstrate the potential for ABM to model different and potentially more complex issues than those in this project.

This example is something that would be incredibly time-consuming to attempt in GIS. Every single position would have to be modelled, which would create hundreds, if not thousands, of layers. It is just not practical to attempt something like this in GIS without ABM. That is why in my opinion ABM has the potential to be a valuable tool in the predictive modeller's tool box. It can allow for very complex models, even if this project did not need to utilise them.

Observations on Explanatory Method

The point of this discussion is that whether using GIS or ABM it does not matter, they are tools. Some tools will work better than others in certain situations. The driving method of this project is the use of simulations to explore possible locations of sites, or more precisely exploring the behaviours that create those sites. It moves beyond the guess work of past inductive predictive models, i.e. people need water so sites must be next to water. It is about testing those assumptions about site placement using ABM and/or GIS. The results of those tests give modellers a greater understanding of the processes that lead to site creation and thus where to find sites. For this project, that knowledge was with regard to water resources, travel and plant resources that led to a 20% increase in overall performance, even though that is a poor measurement of outcomes as discussed. It is this process, explicitly testing theory, that is most important, not the tools.

This project has demonstrated that this process can create models that perform better than older methods. This is in line with the views of Whitely and Verhagen (2011). This is now one of several projects that have found explanatory models perform better than deductive and earlier crude 'inductive' models. That does not mean it will always perform better, but this project stands as a case study about its potential.

Closing Thoughts and Future Work

In summary, several benefits were identified from using explanatory, in this case agent/agency-based modelling methods:

- It provided more nuanced models for CRM workers to use when estimating future project costs and potential resources discovered in the project area.
- It was able to provide answers to many of the questions left unanswered by the previous models that used older methods:
    - the pattern of sites does not match area usage as a travel route;
    - agricultural activity would not be possible in the majority of the project area;
    - people were moving into the project for brief periods of time to gather resources.
- The model adds to the regional research framework and presents evidence against the prevailing theory that the Ceramic period had two groups of people undertaking two different subsistence patterns.

- Discoveries during the process add to the growing literature on concerns about data accuracy in modelling.
- Results from one of the models added to the growing concern about using least-cost path in archaeology.
- A new sort of model using path choice and viewshed was created that could be used for future research by others.
- Increased performance as measured by the Kvamme Gain Statistic.

It has improved our knowledge at the local (site location), regional (understanding of subsistence and habitation), and international (new models and methods) levels. It added new insights to the discussions in several fields of archaeological research, such as predictive modelling, least-cost path and hydrology. It also tested a new method of predictive model creation. These results exceed the original expectations of the project.

Project Problems

This project was not without its setbacks. There were many issues with data, ranging from accuracy with the data used in this project to significant deficiencies with the data used for previous projects. The latter issue made this project's initial attempts at measuring modelling performance unattainable. That was not necessarily a bad thing, as the project highlighted that attempting to quantify performance with a single number does not capture the nuanced outputs required for CRM work. My initial dismissal of some of the problems facing predictive modelling when deciding on which issue to tackle were shown to be incorrect and more work is needed to address these problems.

Future Work

These problems requiring work highlight opportunities for future research. This thesis began with placing the history of predictive modelling into the Gartner Hype Cycle and this project has highlighted that predictive modelling has yet to reach the 'plateau of productivity'. There is still significant work that can be undertaken on predictive modelling. Specifically, areas that have been highlighted through this project:

- More exploration of agent-based modelling. This project served as a case study and demonstrated its usefulness but it is far from a comprehensive review of all its advantages and disadvantages. More comprehensive research needs to be undertaken to fully flesh out the benefits and problems with using such a method for predictive modelling.

- Data issues: rarely have such problems been explored in the predictive modelling literature. This project has found several problems with data and it should be explored more. A future project should look at how data problems affect predictive models.
- Further to exploring data issues, testing older predictive models against ones with new methods and data would be an interesting exercise and may give a better indication of issues with model quality. Could older models' performance, however one measures that, be improved with better data? Was the discovery of poor water data a one-off occurrence or do most predictive models suffer from data issues?
- Measuring predictive models' outputs needs revising. I started this project as a proponent of the Gain Statistic but now think it is a poor measurement of performance. Moreover, I do not see the other methods of measuring performance as being much better. A project to determine optimal methods for measuring the usefulness of predictive modelling would greatly benefit the uptake of predictive modelling. Aligning the outputs better with users' needs is likely to increase use.
- Fully examining all of the issues reviewed in Chapter 2: data problems, clumping together all sites regardless of site type, not creating explanatory frameworks, etc. This project encountered many more of the issues raised by others than was originally anticipated. There are a whole range of topics that could be researched further.

This project met its aim but in doing so created more questions than it answered. There will be much work to undertake in the areas of predictive modelling and agent-based modelling in the future.

Appendices

Appendix A- Full list of agent based modelling programs investigated.

Table 32: List of ABM programs reviewed in this project.

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
1	ABLE (Agent Building and Learning Environment)	http://www.alphaworks.ibm.com/tech/able	6/19/2005	Open source (free for academic use)	Yes	No	Yes	Yes	FAQ section, tutorials, examples, discussion forum, emailing developers, selected publications, API, documentation	No	No
2	ADK (Tryllian Agent Development Kit)	http://www.tryllian.com/adk.html			Defuct						
3	AgentBuilder	http://www.agentbuilder.com/Documentation/Lite/	6/4/2004	Proprietary-Fees	Yes	Yes	No?	Yes	Consulting, training, example, FAQ sectrion, users manuals, defect reporting, mailing list	No	No
4	AgentSheets	http://www.agentsheets.com/	2011	Proprietary-Fees	Yes	Yes	Yes	Yes	Manuals, tutorial movies, FAQ, Books on programming and simulation, personal contact with developers, elementary school training, teacher guides, Wiki	No	No
5	A-Global	http://agents.felk.cvut.cz/aglobe	2008	Free	Yes	Yes	Yes	Yes	Tutorials, Manual	Yes	No
6	Altreva Adaptive Modeler	http://www.altreva.com/	2011	Proprietary-Fees	Yes	No	No	No	FAQ, documentation, tutorial, examples, forum, email support	No	No
7	AMP (Agent Modeling Plantform)	http://www.metascapeabm.com/content/view/57/120/	2011	Open Source?	Yes	Yes	Yes	No	documents, forum, guide, wiki, bug report	coming soon	coming soon
8	AnyLogic	http://www.xjtek.com/	2011	Proprietary-Fees	Yes	Yes	Yes	Yes	Demos, training, consulting, knowledge base, online forum, ask a question, documentation, selected references, Book	Yes	Yes

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
9	AOR Simulation	https://oxygen.informatik.tu-cottbus.de/aor/	2010	Creative Commons License	Yes	Yes	Yes	Yes	Web site, selected references, documentation, examples	No	No
10	Ascape	http://ascape.sourceforge.net/index.html#Introduction	2010	BSD Open Source license	Yes	Yes	Yes	Yes	Online forum (emailing list), selected references, papers, manual, API	No	No
11	Brahms	http://www.agentisolutions.com/index.htm	2011	Free for research	Yes	Yes	Yes	Yes	Documentation, API, tutorials, discussion forums, email contacts	No	No
12	Breve	https://github.com/jonklein/breve, http://www.spiderland.org/breve/	2/25/2008	Open Source	Yes	Yes	Yes	No	Email developer, tutorials, FAQ, forums, defects section, API, documentation	No	Yes
13	Construct	http://www.casos.cs.cmu.edu/projects/construct/index.php	2011	Research Only purposes	N/A	N/A	N/A	N/A	Mailing list	No	No
14	Cormas (Common-pool Resources and Multi-Agent Systems)	http://cormas.cirad.fr/indexeng.htm	2011	Free to modify but not to distribute the modified version	Yes	Yes	Yes	No	Training, selected references, examples, online forum, email developers, documentation	Yes	No
15	Cougaar	http://cougaar.org/	2011	The Cougaar Open Source License (COSL) is a modified version	Yes	Yes	Yes	Yes	FAQ, tutorials, slide shows, documentation, selected references, email support, public forums, mailing lists	No	No

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
				of the OSI approve d BSD License.							
16	DeX/AndroMeta	http://dextk.org/Andro Meta/Home.html	2011	Propriet ary-Fees	No	Yes	Yes	No	Manual	No	Yes
17	D-OMAR (Distributed Operator Model Architecture)	http://omar.bbn.com/	Last updated 04/17/200 9 but most from 2004	Open Source	Yes	Yes	Yes	Yes	API, technical support from authors	No	No
18	ECHO	http://tuvalu.santafe.e du/~pth/echo/	2002	Free, Open Source	No	No	No	Yes	A few selected publications, one outdated publication on how to compile and use Echo	No	No
19	ECJ	http://cs.gmu.edu/~ecl ab/projects/ecj/	2009	Academi c Free License – open source	Yes	Yes	Yes	Yes	Tutorials, examples, API, documentation, online mailing list, Book	No	No
20	EcoLab	http://ecolab.sourcefor ge.net/	2011	Open Source	Yes	Yes	Yes	No	documentation,	No	No
21	Entorama	http://www.entorama. com/	10/3/2008	Unknow n	N/A	N/A	N/A	N/A	Contact authors, help files	No	Yes
22	EVO				Defunct						
23	FAMOJA (Framework for Agent-based MOdelling with JAva)	http://www.usf.uos.de/ projects/famoja/	2007	LGPL licences	Yes	Yes	Yes	Yes	Tutorial, API, wiki, documentation,	No	No
24	FLAME	http://www.flame.ac.u k/	2011	Unknow n	Yes	Yes	Yes	No	Tutorials, user guide, contact authors	No	No

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
25	Framsticks	http://www.framsticks.com/	2011	Depends on module-GPL/LGPL/Proprietary	Yes	Yes	Yes	No	Email developer, tutorials, manual, FAQ, forums, API, documentation, selected publications, examples	No	Yes
26	GAMA	http://gama.ifi.refer.org/mediawiki/index.php/GAMA	2010	LGPL	Yes	Yes	Yes	No	Contact authors, report bug, Tutorials, guide	Yes	No
27	GROWlab	http://www.icr.ethz.ch/research/growlab/	July 4, 2008	Freely available	Yes	Yes	Yes	Yes	Guid,, Publications,	No	Yes
28	ICARO-T	http://icaro.morfeo-project.org/	2008	Open Source	Yes	No	No	No	Code with examples and documentation.	No	No
29	iGen [18]	http://www.chisystems.com/cognitivemodel.html		no longer for public use							
30	JADE	http://jade.tilab.com/	2011	Open Source LGPL version 2				Yes	FAQ, mailing list, defect list, tutorials, API, documentation	No	No
31	Jade's sim++					defunt					
32	JAMEL (Java Agent-based MacroEconomic Laboratory)	http://p.seppecher.free.fr/jamel/	2010	General Public Licence (GPL)	Yes	Yes	Yes	Yes	examples, reference paper, contact developer	No	No
33	Janus	http://www.janus-project.org/Home	2011	GPLv3 for non-commercial use, or adhoc commer	Yes	Yes	Yes	No	FAQ, documentation, online forum, examples, defect list, tutorials	No	Yes

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
				cial license.							
34	JAS	http://jaslibrary.sourceforge.net/	3/18/2006	Open Source LGPL, associated third party licenses	Yes	Yes	Yes	Yes	API, documentation, tutorials, email authors	No	No
35	JASA (Java Auction Simulator API)	http://sourceforge.net/projects/jasa/	2011	GNU General Public License (GPL)	Yes	Yes	Yes	Yes	Public forum, not very well used, API, small set of selected readings, limited documentation	No	No
36	JCA-Sim	http://www.jweimar.de/jcasim/jcasim.html	4/11/2001	Free (closed source)	Yes	Yes	Yes	Yes	Examples, documentation, API, one contact listed	No	Yes
37	jEcho				Defunct						
38	jES (Java Enterprise Simulator)	http://web.econ.unito.it/terna/jes/	2006	Academic free license	Yes	Yes	Yes	Yes	limited documentation	No	No
39	JESS	http://herzberg.ca.sandia.gov/jess/	11/11/2008	Proprietary, free for academic use	Yes	Yes	Yes	Yes	FAQ, documentation, mailing list, examples, third party plug ins and libraries, wiki	No	No
40	JIAC	http://www.jiac.de/agent_frameworks/jiac_v/	2011	Unknown	Yes	Yes	Yes	Yes	Manual	No	No
41	LSD (Laboratory for Simulation Development)	http://www.labsimdev.org/Joomla_1-3/	4/24/2009	GPL	Yes	No	Yes	No	Documentation, a couple of examples, forum	No	No

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
42	Madkit (Multi Agent Development Kit)	http://www.madkit.org/	2011	LGPL for basic libraries, GPL for development and non- commercial applications				Yes	FAQ, documentation, online forum, examples, defect list	No	No
43	MAGSY	http://www-ags.dfki.uni-sb.de/~kuf/magsy.html				defunct					
44	MAML (Multi-Agent Modeling Language)	http://www.maml.hu/maml/initiative/index.html	10/20/2000	The compiler is freely downloadable for evaluation purposes (open source) Later the system will be put under GNU license	Yes	Yes	No	No	Tutorial, examples, reference papers, contact developers	No	No
45	MASON	http://cs.gmu.edu/~eclab/projects/mason/	2011	Academic Free License (open source)	Yes	Yes	Yes	Yes	Mailing list, documentation, Tutorials, third party extensions, reference papers, API	Yes	Yes
46	MASS (Multi-Agent Simulation Suit)	http://mass.aitia.ai/	2009	Proprietary, free version available	Yes	Yes	Yes	Yes	Manuals, tutorials, mailing lists, reference papers.	No	No
47	MAS-SOC (Multi-Agent Simulations for the SOCial Sciences)					defunct					

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
48	MASyV (Multi-Agent System Visualization)	http://masyv.sourceforge.net/	7/22/2008	Open Source	No	Yes	Yes	No	bug report, referneces	No	No
49	MIMOSE (Micro- und Multilevel Modelling Software)	http://www.uni-koblenz.de/~moeh/projekte/mimose.html	10/15/1999	Free (closed source)	Yes	Yes	Yes	Yes	User's manual	No	No
50	Moduleco					defunct					
51	MOOSE (Multimodeling Object-Oriented Simulation Environment)	http://www.cise.ufl.edu/~fishwick/moose.html	1997	Unknown	Yes	Yes	No	No	Selected references, user's manual in toolkit package	No	No
52	NetLogo	http://ccl.northwestern.edu/netlogo/	2011	Originally free, but not Open Source, when project began. Became Open Source during project.				Yes	Documentation, FAQ, selected references, tutorials, third party extensions, defect list, mailing lists	Yes	Yes
53	OBEUS (Object Based Environment for Urban Simulation)	http://www.enib.fr/~harrouet/oris.html	2000	Free (closed source)	Yes	No	No	No	User's manual	No	No
54	Omonia (previously Quicksilver)	http://www.xlog.ch/omonia/	1/13/2007	LGPL	Yes	Yes	Yes	Yes	Examples, little documentation	No	No
55	oRIS [40]	http://www.enib.fr/~harrouet/	2003?	Proprietary - (free for academic institutions)	Yes	Yes	No	No	Documentation, examples , API	No	No
56	PS-I (Political Science-Identity)	http://ps-i.sourceforge.net/	2011	GNU General	Yes	Yes	Yes	No	Documentation, selected publications	No	No

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
				Public License (GPL)							
57	Quicksilver				defunct						
58	Repast	http://repast.sourceforge.net/	2011	BSD	Yes	Yes	Yes	Yes	Documentation, mailing list, defect list, reference papers, external tools, tutorials, FAQ, examples	Yes	Yes
59	SDML (Strictly Declarative Modeling Language)	http://cfpm.org/sdml/	2000	GPL, third party license (for VisualWorks)	Yes	Yes	Yes	Yes	Mailing list, tutorial, selected references, limited documentation included with software package	No	No
60	SEAS (System Effectiveness Analysis Simulation)	http://teamseas.com/	2011	Free with government approval	Yes	No	No	No	User manual, examples, training, email, phone	Unknown	Yes
61	SeSAm (Shell for Simulated Agent Systems)	http://www.simsesam.de/	2010	LGPL	Yes	Yes	Yes	Yes	Tutorials, mailing list, FAQ, wiki, author contact	Yes	Plugin available
62	SimAgent (also sim agent)	http://www.cs.bham.ac.uk/research/projects/poplog/packages/simagent.html	5/30/2005	Free (open source), may be replaced by GPL	Yes	Yes	Yes	No	Tutorials, documentation, Selected publications, examples, author contact	No	No
63	SimBioSys	http://www.lucifer.com/~david/SimBioSys/	1994?	Artistic License	Yes	Yes	Yes	No	None	No	No

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
				Agreement							
64	SimPack				defunct						
65	SimPlusPlus	http://www.simplusplus.com/	early 2000's?	GPL	Yes	Yes	Yes	No	Contact authors	No	No
66	SimPy	http://simpy.sourceforge.net/	7/2/2005	GNU, LGPL	Yes	Yes	Yes	No	References, wiki, Tutorials, Manuals, Book, email list, contact programers, bug reporting	No	No
67	Soar	http://sitemaker.umich.edu/soar/home	2011	BSD	Yes	Yes	Yes	No	Documentation, FAQ, selected publications, defect list, third party extensions, mailing list, contact authors, tutorial, examples, wiki	No	No
68	Spark	http://www.pitt.edu/~cirm/spark/	2010	Unknown	N/A	N/A	N/A	N/A	Guide, Tutorials, user group, contact author	No	No
69	StarLogo	http://education.mit.edu/starlogo/	6/30/2005	Free (closed source) - Clearthought Software License, Version 1.0	Yes	Yes	Yes	Yes	Mailing list, tutorials, FAQ, bug list, documentation, developer contacts	No	Yes
70	StarLogo TNG	http://education.mit.edu/starlogo/	2011	StarLogo TNG License v1.0 - (closed source) - the code may be freed up	Yes	Yes	Yes	No	Tutorials, FAQ, documentation, mailing lists, API	No	Yes

	Software	Website	Last Updated	License	Windows	Linux	Mac	Java	User Support	GIS	3D
				eventually.							
71	Sugarscape	http://sugarscape.sourceforge.net/	???	GPL	Yes	Yes	Yes	Yes	API, walkthrough, tuturial	No	No
72	Swarm	http://www.swarm.org/index.php/Main_Page	2011	GPL	Yes	Yes	Yes	No	Wiki, tutorials, examples, documentation, FAQ, selected publications, mailing lists	No	No
73	TerraME	http://www.terrame.org/doku.php	2011	Open Source	Yes	No	No	No	Tutorials, Examples, Courses, references	Yes	No
74	VisualBots	http://www.visualbots.com/	1/5/2008	Free, Not Open Source	Yes	No	No	No	Object model documentation, tutorials, example projects	No	No
75	VSEit	http://www.vseit.de/	2001	Free (closed source)	Yes	Yes	Yes	Yes	Examples, users guide, defect list,	No	No
76	Xholon	http://sourceforge.net/apps/mediawiki/xholon/index.php?title=Main_Page	2001	LGPL	Yes	Yes	Yes	Yes	Tutorials, many examples, user guide, web sites	No	No
77	ZEUS			Defunct							

Appendix B- Data used in the project.

Table 33: Soil types from NRCS dataset (NMAES 1971, USDA 1981).

Code	Soil	Soil Description	Percentage of Project Area Covered
Aa	Anthony sandy loam, 0 to 1 percent slopes	Deep, light-colored, nearly level, calcareous soils that developed in stratified alluvium derived from crystalline an sedimentary rocks	0.13%
At	Atoka loam, 1 to 3 percent slopes	Well-drained, moderately dark colored, level to gently sloping soils that developed in moderately deep old alluvium derived from calcareous sedimentary rocks	0.01%
Ah	Anthony sandy loam, 0 to 1 percent slopes, eroded	Soil has been eroded by wind but it the same as Aa	0.01%
AH	Arno-Harkey complex, saline, 0 to 1 percent slopes	Soils affected by salinity and a fluctuating high water table	0.004%
Ao	Atoka Loam, o to 1 percent slopes	Well-drained soils that developed in moderately deep old alluvium, derived from calcareous sedimentary rocks	0.02%
DP	Dev-Pima complex, 0 to 3 percent slopes	Nearly level, moderately dark colored, gravelly soils that developed in alluvium.	3.76%
DRG	Deama-Rock outcrop complex, 50 to 150 percent slopes	Shallow, well drained soils that formed in residuum from limestone bedrock.	0.04%
DYE	Dye-Encierro complex, 5 to 30 percent slopes	Shallow, well drained soils on hills and tops of messas	0.02%
EC	Ector stony loam, 0 to 9 percent slopes	Very shallow to shallow, well-drained, calcareous, stony and extremely rocky soils that are underlain by limestone.	24.60%
EE	Ector extremely rocky loam, 9 to 25 percent slopes		23.55%
ER	Ector-Reagan association, 0 to 9 percent slopes		8.45%
GA	Gypsum land	Very steep and steep broken, or eroded exposures of gypsiferous rocks and earths and very shallow soils	0.14%
GC	Gypsum land-Cottonwood complex, 0 to 3 percent slopes		0.30%
GP	Gravel Pit		0.004%
Ha	Harkey sandy loam, 0 to 1 percent slopes	Deep, well-drained, strongly calcareous, moderately dark colored soils that developed in mixed alluvium.	0.01%
Hk	Harkey very fine sandy loam, 0 to 1 percent slopes		0.02%
Ku	Karro loam, 1 to 3 percent slopes	A stongly calcareous, loamy soils that developed in deep, old alluvium derived from calcareous sedimentary rocks	0.01%
LN	Largo-Stony land complex, 0 to 25 percent slopes	Deep, reddish-brown, calcareous, gently sloping soils that developed in alluvium derived from upland sedimentary material.	0.18%
LT	Limestone rock land	Steep to very steep canyon walls and escarpments	8.97%
MXC	Montecito loam, 0 to 10 percent slopes	Deep well-drained soils formed in mixed alluvium and eolian material	0.19%

Code	Soil	Soil Description	Percentage of Project Area Covered
PD	Pajarito-Dune land complex, 0 to 3 percent slopes	Deep, well-drained, weakly calcareous to noncalcareous soils that developed in wind worked material and alluvium derived from mixed, sandy sediments of the uplands	0.06%
Pe	Pima silt loam, 0 to 1 percent slopes	Deep, well-drained, moderately dark colored, calcareous soils that developed in alluvium derived limestone.	0.13%
PM	Pima silt loam, 0 to 1 percent slopes		0.70%
RA	Reagan loam, 0 to 3 percent slopes	Deep, well-drained, moderately dark colored, calcareous loams that developed in old alluvium derived from calcareous, sedimentary rocks of the uplands.	2.45%
Rc	Reagan loam, 0 to 1 percent slopes		0.13%
Rd	Reagan loam, 1 to 3 percent slopes		0.39%
RE	Reagan-Upton association, 0 to 9 percent slopes		9.34%
RG	Reeves-Gypsum land complex, 0 to 3 percent slopes	Well-drained, calcareous soils that are shallow to moderately deep over gypsiferous earths or rocks.	0.55%
RM	Reeves-Reagan loams, 0 to 3 percent slopes		1.17%
RPG	Rock outcrop-Deama complex, 40 to 150 percent slopes		0.56%
RTE	Rock outcrop-Tortugas-Ustifluvents complex, 0 to 80 percent slopes		0.28%
SG	Simona gravelly fine sandy loam, 0 to 3 percent slopes	Well-drained, moderately dark colored soils that are calcareous and moderately course in texture.	0.21%
SM	Simona-Bippus complex, 0 to 5 percent slopes		0.07%
TN	Tonuco loamy fine sand, 0 to 3 percent slopes, eroded	Moderately dark colored noncalcareous soils that have been worked by wind	0.03%
TPE	Tortugas-Deama association, moderately steep	Shallow, well-drained soils that formed in residuum from limestone and calcareous sandstone.	0.70%
UG	Upton gravelly loam, 0 to 9 percent slopes	Calcareous, gravelly soils that developed in old alluvium derived from calcroues sedimentary rocks.	9.87%
Uo	Upton gravelly loam, 0 to 9 percent slopes		0.50%
Up	Upton soils, 0 to 1 percent slopes		0.02%
UR	Upton-Reagan complex, 0 to 9 percent slopes		2.32%
Ut	Upton soils, 1 to 3 percent slopes		0.04%
W	Water		0.08%

Table 34: Distribution vegetation in the study area. (Altschul et al. 2005, USGS 2012b).

Pump III Vegetation	Percent of Land Coverage	GAP Data for this project	Percent of Land Coverage
Chihuahuan Foothill-Piedmont Desert Grassland	68.65%	Chihuahuan Creosotebush, Mixed Desert and Thorn Scrub	46.70%
Chihuahuan Desert Scrub	15.32%	Apacherian-Chihuahuan Semi-Desert Grassland and Steppe	32.98%
Chihuahuan Desert Grassland	5.53%	Western Great Plains Shortgrass Prairie	10.48%
Rocky Mountain/ Great Basin Closed Conifer Woodland	3.05%	Apacherian-Chihuahuan Mesquite Upland Scrub	4.69%
Chihuahuan Lowland/Swale Desert Grassland	2.05%	Chihuahuan Mixed Salt Desert Scrub	1.14%
Short Grass Steppe	1.44%	Chihuahuan Stabilized Coppice Dune and Sand Flat Scrub	0.49%
Southwest and Plains Forested/Shrub Wetland	.92%	Coahuilan Chaparral	0.48%
Madrean Open Oak Woodland	.62%	North American Warm Desert Riparian Woodland and Shrubland	0.44%
Rocky Mountain Montane Scrub and Interior Chaparral	.6%	Developed, High Intensity	0.38%
Broadleaf Evergreen Interior Chaparral	.59%	Western Great Plains Sandhill Steppe	0.31%
Rocky Mountain/ Great Basin Closed Conifer Woodland	.49%	Chihuahuan Gypsophilous Grassland and Steppe	0.25%
Urban Vegetated	.3%	Cultivated Cropland	0.22%
Basin/Playa	.17%	Developed, Low Intensity	0.20%
Barren	.12%	North American Warm Desert Lower Montane Riparian Woodland and Shrubland	0.18%
Rock Outcrop	.08%	North American Warm Desert Bedrock Cliff and Outcrop	0.18%
Riverine/Lacustrine	.07%	Inter-Mountain Basins Semi-Desert Shrub Steppe	0.17%
Irrigated Agriculture	.06%	North American Warm Desert Wash	0.15%
		Open Water (Fresh)	0.11%
		Western Great Plains Cliff and Outcrop	0.11%
		Southern Rocky Mountain Juniper Woodland and Savanna	0.07%
		Mogollon Chaparral	0.04%
		North American Warm Desert Playa	0.04%
		Madrean Juniper Savanna	0.04%
		Chihuahuan Succulent Desert Scrub	0.03%

Pump III Vegetation	Percent of Land Coverage	GAP Data for this project	Percent of Land Coverage
		Chihuahuan Sandy Plains Semi-Desert Grassland	0.03%
		Madrean Encinal	0.03%
		North American Arid West Emergent Marsh	0.02%
		Madrean Pinyon-Juniper Woodland	0.01%
		North American Warm Desert Active and Stabilized Dune	0.002%
		Rocky Mountain Lower Montane Riparian Woodland and Shrubland	0.001%
		Colorado Plateau Mixed Low Sagebrush Shrubland	0.001%
		Madrean Pine-Oak Forest and Woodland	0.0003%

Table 35: Data from USGS water data gathering stations in and next to the subject area.

Test Point	Location	Drainage Area	Period of Record	Annual Runoff (ac-ft) 2010	Annual Runoff (ac-ft) 1964-2010	Days with flow 2010	Avg. Annual Runoff Against Avg. Rain Fall (assuming avg. 12 inches of rain)
08401900 Rocky Arroyo at Highway Bridge, Near Carlsbad, NM	Lat 32°30'21.89", long 104°22'29.96" referenced to North American Datum of 1983, Eddy County, NM, Hydrologic Unit 13060011, at downstream end of bridge pier nearest left bank on U.S. Highway 285, 2.1 mi upstream from mouth, and 10 mi northwest of Carlsbad. Mouth at Pecos River mile 475.2.	285 mi^2, approximately.	November 1963 to current year.	939	4,040	5	2.1%
08405050 Last Chance Canyon, Near Carlsbad Caverns, NM	Lat 32°17’31’’, long 104°36’25’’ referenced to North American Datum of 1983, Eddy County, NM, Hydrologic Unit 13060011, upstream from culvert on State Highway 137, 0.1 mi north of road to Sitting Bull Falls, and 12.5 mi northwest of Carlsbad Caverns.	2 mi^2.	Water years 1959 to 1996, 2005 to current year	Not listed	Not listed	Not Listed	
08405100 Mosley Canyon, Near White City, NM	Lat 32°15’27’’, long 104°22’43’’ referenced to North American Datum of 1983, Eddy County, NM, Hydrologic Unit 13060011, 600 ft downstream from dip on Dark Canyon Road, and 5.5 mi north of Whites City.	14.6 mi^2.	Water year 1959 to current year.	Not listed	Not listed	Not Listed	
08405105 Dark Canyon Draw Near White City, NM	Lat 32°17’25.55’’, long 104°20’57’’ referenced to North American Datum of 1983, Eddy County, NM, Hydrologic Unit 13060011, on left bank 0.25 mi upstream from mouth of canyon, and approximately 11.0 mi upstream from Dark Canyon Draw at Carlsbad.	327 mi^2, approximately.	February 2002 to current year.	2280	2560	7	1.2%
08405150 Dark Canyon Draw at Carlsbad, NM	Lat 32°24’12’’, long 104°13’46’’ referenced to North American Datum of 1983, Eddy County, NM, Hydrologic Unit 13060011, on right bank and upstream from San Jose Street, and 1.0 mi upstream from mouth. Mouth at Pecos River mile 459.2.	451 mi^2, approximately.	January 1973 to current year.	260	3290 (+ 2100 for irrigation) 5390	1	1.9%

Table 36: Conversion table for Runoff Curve Numbers from Antecedent Moisture Condition Class **II** to AMC Class **I** or Class **III** (after USDA-SCS (1972)).

AMC II	AMC I	AMC III	AMC II	AMC I	AMC III
100	100	100	58	38	76
98	94	99	56	36	75
96	89	99	54	34	73
94	85	98	52	32	71
92	81	97	50	31	70
90	78	96	48	29	68
88	75	95	46	27	66
86	72	94	44	25	64
84	68	93	42	24	62
82	66	92	40	22	60
80	63	91	38	21	58
78	60	90	36	19	56
76	58	89	34	18	54
74	55	88	32	16	52
72	53	86	30	15	50
70	51	85	25	12	43
68	48	84	20	9	37
66	46	82	15	6	30
64	44	81	10	4	22
62	42	79	5	2	13
60	40	78	0	0	0

Table 37: Soil water carrying capacity (% of unit of measurement).

Code	Water capacity	Project Area Coverage	Code	Water capacity	Project Area Coverage
Aa	.12	0.13%	Pe	.2	0.13%
At	.19	0.01%	PM	.2	0.70%
AH	.1-.09	0.01%	RA	.14	2.45%
Ah	.12	0.004%	Rc	.13	0.13%
Ao	.19	0.02%	Rd	.13	0.39%
DP	.2-.12	3.76%	RE	.14-.13	9.34%
DRG	.08	0.04%	RG	.13	0.55%
DYE	.15	0.02%	RM	.13	1.17%
EC	.09	24.60%	RPG	.07-0	0.56%
EE	.06	23.55%	RTE	.07-0	0.28%
ER	.14-.09	8.45%	SG	.11	0.21%
GA	N/A	0.14%	SM	.15-.11	0.07%
GC	.15	0.30%	TN	.09	0.03%
GP	N/A	0.004%	TPE	.08	0.70%
Ha	.15	0.01%	UG	.1	9.87%
Hk	.16	0.02%	Uo	.1	0.50%
Ku	.18	0.01%	Up	.07	0.02%
LN	.17	0.18%	UR	.14-.11	2.32%
LT	N/A	8.97%	Ut	.09	0.04%
MXC	.16	0.19%	W	N/A	0.08%
PD	.14-.13	0.06%			

Table 38: Soil absorption rates for the soils in project area.

Code	Soil absorption rates	Project Area Coverage	Code	Soil absorption rates	Project Area Coverage
Aa	2.00 to 6.00 in/hr	0.13%	Pe	0.20 to 0.60 in/hr	0.13%
At	0.00 to 0.06 in/hr	0.01%	PM	0.20 to 0.60 in/hr	0.70%
AH	0.00 to 0.06 in/hr	0.01%	RA	0.60 to 2.00 in/hr	2.45%
Ah	2.00 to 6.00 in/hr	0.004%	Rc	0.60 to 2.00 in/hr	0.13%
Ao	0.00 to 0.06 in/hr	0.02%	Rd	0.60 to 2.00 in/hr	0.39%
DP	2.00 to 6.00 in/hr	3.76%	RE	0.60 to 2.00 in/hr	9.34%
DRG	0.00 to 0.06 in/hr	0.04%	RG	0.00 to 0.06 in/hr	0.55%
DYE	0.00 to 0.06 in/hr	0.02%	RM	0.00 to 0.06 in/hr	1.17%
EC	0.06 to 2.00 in/hr	24.60%	RPG	0.00 to 0.06 in/hr	0.56%
EE	0.06 to 2.00 in/hr	23.55%	RTE	0.00 to 0.06 in/hr	0.28%
ER	0.06 to 2.00 in/hr	8.45%	SG	0.00 to 0.06 in/hr	0.21%
GA	0.20 to 2.00 in/hr	0.14%	SM	0.00 to 0.06 in/hr	0.07%
GC	0.20 to 2.00 in/hr	0.30%	TN	0.00 to 0.06 in/hr	0.03%
GP	N/A	0.004%	TPE	0.00 to 0.06 in/hr	0.70%
Ha	0.57 to 1.98 in/hr	0.01%	UG	0.01 to 0.60 in/hr	9.87%
Hk	0.60 to 2.00 in/hr	0.02%	Uo	0.01 to 0.60 in/hr	0.50%
Ku	0.20 to 0.60 in/hr	0.01%	Up	0.01 to 0.60 in/hr	0.02%
LN	0.20 to 0.60 in/hr	0.18%	UR	0.01 to 0.60 in/hr	2.32%
LT	bedrock	8.97%	Ut	0.01 to 0.60 in/hr	0.04%
MXC	0.20 to 0.60 in/hr	0.19%	W	N/A	0.08%
PD	2.00 to 6.00 in/hr	0.06%			

Table 39: Soil hydrology classification for soils in project area.

Code		Project Area Coverage	Code		Project Area Coverage	Total	
Aa	B	0.13%	Pe	B	0.13%	A	3.76%
At	C	0.01%	PM	B	0.70%	B	15.46%
AH	D	0.01%	RA	B	2.45%	C	13.22%
Ah	B	0.004%	Rc	B	0.13%	D	67.49%
Ao	C	0.02%	Rd	B	0.39%	N/A	0.08%
DP	A	3.76%	RE	B	9.34%		
DRG	D	0.04%	RG	B	0.55%		
DYE	D	0.02%	RM	B	1.17%		
EC	D	24.60%	RPG	D	0.56%		
EE	D	23.55%	RTE	D	0.28%		
ER	D	8.45%	SG	D	0.21%		
GA	C	0.14%	SM	D	0.07%		
GC	C	0.30%	TN	D	0.03%		
GP	A	0.004%	TPE	D	0.70%		
Ha	B	0.01%	UG	C	9.87%		
Hk	B	0.02%	Uo	C	0.50%		
Ku	B	0.01%	Up	C	0.02%		
LN	B	0.18%	UR	C	2.32%		
LT	D	8.97%	Ut	C	0.04%		
MXC	B	0.19%	W	N/A	0.08%		
PD	B	0.06%					

Table 40: Carlsbad historical climate data. For monthly and annual means, thresholds, and sums: Months with 5 or more missing days are not considered, Years with 1 or more missing months are not considered, Seasons are climatological not calendar seasons. (Winter = Dec., Jan., and Feb.; Spring = Mar., Apr., and May; Summer = Jun., Jul., and Aug.; Fall = Sep., Oct., and Nov.).

From Year=1900 To Year=2012														
	Precipitation											Total Snowfall		
	Mean (in.)	High (in.)	Year	Low (in.)	Year	1 Day Max. (in.)		>= 0.01 in.	>= 0.10 in.	>= 0.50 in.	>= 1.00 in.	Mean (in.)	High (in.)	Year
								# Days						
January	0.4	2.31	1949	0	1912	0.79	1980	3	1	0	0	1.2	17.2	1949
February	0.44	2.26	1997	0	1900	1.25	1997	2	1	0	0	1	17.8	1905
March	0.48	4.39	1919	0	1903	2.41	1919	2	1	0	0	0.3	6	1969
April	0.65	5.04	1915	0	1902	2.86	2004	2	1	0	0	0.2	13.5	1928
May	1.19	12.28	1941	0	1903	3.41	1959	4	2	1	0	0	0	1905
June	1.49	6.24	1948	0	1928	3.8	1972	4	3	1	0	0	0	1905
July	1.86	10.5	1902	0	1903	3.8	1902	5	3	1	0	0	0	1905
August	1.79	7.7	1984	0.01	1938	5.12	1916	5	3	1	0	0	0	1909
September	2.14	12.27	1980	0	1907	4.6	1980	5	3	1	1	0	0	1906
October	1.34	8.08	1907	0	1903	4.3	1945	4	2	1	0	0	0	1905
November	0.58	4.58	2004	0	1915	2	2000	3	1	0	0	0.6	10.3	1976
December	0.51	3.79	1991	0	1903	1.18	1986	3	1	0	0	1.2	16.2	2009
Annual	12.87	33.94	1941	2.95	1924	5.12	1916	42	25	8	3	4.5	20	2007
Winter	1.35	6.16	1992	0	1934	1.25	1997	7	4	1	0	3.4	22.2	2010
Spring	2.32	17.99	1941	0	2011	3.41	1959	8	5	1	1	0.4	6	1969
Summer	5.14	18.06	1902	0.74	1924	5.12	1916	15	9	3	1	0	0	1910
Fall	4.05	16.01	1974	0.27	1951	4.6	1980	12	7	3	1	0.6	10.3	1976

Table 41: Rain fall and peak mean flow for Carlsbad in recorded years.

Date	high mean flow ac-ft (USGS)	max 1-day rain (in) (WWCR)	Date	high mean flow ac-ft(USGS)	max 1-day rain (in) (WWCR)
Sep-74	2,170	4.62	Jun-95	83	0.95
Oct-74	4,260	1.64	Sep-95	82	1.32
Sep-78	3,480	2.62	Oct-97	31.	3
Nov-78	355	1.99	Jul-02	554	1.58
May-79	227	1.3	Oct-03	20	0.55
Jun-79	137	1.46	Apr-04	12,400	2.86
Jul-81	367	1.13	Jul-04	203	1.42
Aug-84	2,670	2.35	Aug-04	27	1.13
Jun-86	8,360	2.73	Sep-04	58	1.12
May-87	74	1.15	Sep-06	228	1.28
Aug-88	46	1.65	May-07	216	1.45
Sep-88	19	2.85	Sep-07	195	1.85
Sep-90	93	0.16	Jun-09	460	1.18
Sep-91	523	2.26	Sep-10	131	1.12
Aug-94	140	0.76			

Table 42: List of rain fall in inches of Carlsbad New Mexico from 2001-2011. a = 1 day missing from record, b = 2 days missing from record, c = 3 days, ..etc.., z = 26 or more days missing, A = Accumulations present.

Year(s)	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
2001	0.90	0.48	0.88	0.49	0.53	1.22	0.45	0.52	0.96e	0.11	1.25	0.20	7.99
2002	0.17	0.63	2.15	0.00	0.00	0.26	2.49	2.08	0.97	1.99a	0.56	0.50	11.80
2003	0.00	0.99	0.35	0.00	1.19	0.50a	0.56	0.31	0.00z	1.23	0.48	0.00	5.61
2004	0.25	0.95	1.71	4.17	0.00z	1.67	2.85	2.05	4.12	0.77	4.58	0.82	23.94
2005	0.45	1.66	0.67	0.18	1.51	0.12	0.57	2.40	0.26	1.05	0.00	0.00	8.87
2006	0.00	0.31	1.20	0.10	0.00z	1.80a	0.54b	0.00z	4.30	0.00z	0.09	0.33	8.67
2007	1.64	0.48	2.64	0.59	3.45	1.01	1.30	1.47	5.11	0.00	0.41	0.92	19.02
2008	0.01	0.04	0.23	0.00	0.86	0.62	2.52	1.88	2.33	0.73	0.04	0.13	9.39
2009	0.00	0.00z	0.14	0.01a	0.40	1.78	6.13	0.63	0.26	1.02	0.10	1.49	11.96
2010	0.98	1.34	0.41	0.61	0.86	1.24	7.36	1.00	3.41	0.10	0.00	0.01	17.32
2011	0.00	0.43	0.00	0.00	0.00	0.04	0.62	0.50	1.67	0.27	0.05	1.48	5.06

Table 43: Site Classifications in NMCRIS and categories used in this project.

Project Category	Other Classifications	Count
Apache	Apache 1539 AD to 1846AD, Apache 1850 AD to 1879AD, Apache 1867AD to 1875AD, Apache 1870AD 1890AD, Spanish Contact to Mexican 1539 AD to 1846 AD, Apache 1840AD to 1870AD	11
Archaic	Archaic 5500BC to 900AD, Archaic 5500BC to 200AD	2
Clovis to Late Paleoindian	Clovis to Late Paleoindian 9500 BC to 6600 BC	1
Early Archaic	Early Archaic 5500 BC to 3000 BC	2
Early Archaic to Late Archaic	Early Archaic to Late Archaic 4000BC to 1000AD, Early Archaic to Late Archaic 5500BC to 200AD	2
Early Archaic to Middle Archaic	Early Archaic to Middle Archaic 5500BC to 1800AD	1
Early Pithouse	Early Pithouse to Early Pithouse 200AD to 950AD	1
Early Pithouse to Early Pueblo	Early Pithouse to Early Pueblo 600AD to 1200AD	1
Early Pithouse to Late Pithouse	Early Pithouse to Late Pithouse 200AD to 750AD, Early Pithouse to Late Pithouse 500AD to 1000AD	2
Early Pithouse to Late Pueblo	Early Pithouse to Late Pueblo 1000 AD to 1400AD, Early Pithouse to Late Pueblo 1AD to 1375AD, Early Pithouse to Late Pueblo 2000AD to 1400AD, Early Pithouse to Late Pueblo 200AD to 1100AD, Early Pithouse to Late Pueblo 200AD to 1400AD, Early Pithouse to Late Pueblo 350AD to 1400AD, Early Pithouse to Late Pueblo 500AD to 1400AD, Early Pithouse to Late Pueblo 650AD to 1300AD, Early Pithouse to Late Pueblo 650AD to 1400AD, Early Pithouse to Late Pueblo 700AD to 1600AD	15

Project Category	Other Classifications	Count
Early Pueblo	Early Pueblo 1100AD to 1200AD, Early Pueblo 950AD to 1200AD	7
Early Pueblo to Late Pueblo	Early Pueblo to Late Pueblo 1075AD to 1375AD, Early Pueblo to Late Pueblo 1100 AD to 1400 AD, Early Pueblo to Late Pueblo 1100AD to 1350AD, Early Pueblo to Late Pueblo 1100AD to 1500AD, Early Pueblo to Late Pueblo 1150AD to 1250AD, Early Pueblo to Late Pueblo 950AD to 1300AD, Early Pueblo to Late Pueblo AD 1100-1400, Early Pueblo to Late Pueblo AD 1175-1400	28
Historic		92
Late Archaic	Late Archaic (1486 BC) to Late Archaic (200 AD), Late Archaic 0AD to 900AD, Late Archaic 1000BC to 200AD, Late Archaic 1000BC to 500AD, Late Archaic 1000BC to 750AD, Late Archaic 100BC to 900AD, Late Archaic 1065 BC to 145AD, Late Archaic 1500BC to 300AD, Late Archaic 1800 BC to 200AD, Late Archaic 1800BC to 700AD, Late Archaic 1800BC to 900 AD, Late Archaic 1800BC to 950AD, Late Archaic 1AD to 1000AD, Late Archaic 200AD to 1200 AD, Late Archaic 200AD to 900AD, Late Archaic 200BC to 600AD, Late Archaic 200BC to 900AD, Late Archaic 2300 BC to 1100AD, Late Archaic 500BC to 200AD, Late Archaic 750AD to 950AD, Late Archaic 800BC to 600AD, Late Archaic 820BC to 900AD	55
Late Archaic to Unspecific Archaic 800BC to 200AD	Late Archaic to Unspecific Archaic 800BC to 200AD	1
Late Paleoindian 8000BC to 6600BC	Late Paleoindian 8000BC to 6600BC	2
Late Pithouse	Late Pithouse 750AD to 1100AD, Late Pithouse 750AD to 1400AD, Late Pithouse 850AD to 1000AD, Late Pithouse 900AD to 1040AD, Late Pithouse 900AD to 1100AD, Late Pithouse 900AD to 1400AD, Late Pithouse to Late Pithouse (750 AD) to (1100 AD), Late Pithouse to Late Pithouse 750AD to 1100AD, Late Pithouse to Late Pithouse 850AD to 1000AD, Late Pithouse to Late Pithouse 900AD to 1100AD	16
Late Pithouse to Early Pueblo	Late Pithouse to Early Pueblo (750 AD) to (1200 AD), Late Pithouse to Early Pueblo 1000AD to 1200AD, Late Pithouse to Early Pueblo 750AD to 1100AD, Late Pithouse to Early Pueblo 750AD to 1175AD, Late Pithouse to Early Pueblo 750AD to 1200AD	7
Late Pithouse to Late Pueblo	Late Pithouse to Late Pueblo (1000 AD) to (1300 AD), Late Pithouse to Late Pueblo (700 AD) to (1500 AD), Late Pithouse to Late Pueblo (750 AD) to (1400 AD), Late Pithouse to Late Pueblo (900 AD) to (1350 AD), Late Pithouse to Late Pueblo (900 AD) to (1400 AD), Late Pithouse to Late Pueblo (900 AD) to (1600 AD), Late Pithouse to Late Pueblo 750 AD to 1400 AD, Late Pithouse to Late Pueblo 900 AD to 1400 AD, Late Pithouse to Late Pueblo 900 AD to 1450 AD, Late Pithouse to Late Pueblo 1000AD to 1200AD, Late Pithouse to Late Pueblo 1100AD to 1500AD, Late Pithouse to Late Pueblo 750AD to 1200AD, Late Pithouse to Late Pueblo 750AD to 1400AD, Late Pithouse to Late Pueblo 750AD to 1450AD, Late Pithouse to Late Pueblo 850AD to 1200AD, Late Pithouse to Late Pueblo 900AD to 1350AD, Late Pithouse to Late Pueblo 900AD to 1400AD, Late Pithouse to Late Pueblo 900AD to 1450AD, Late Pithouse to Late Pueblo 950AD to 1300AD, Late Pithouse to Late Pueblo 950AD to 1350AD	69
Late Pueblo	Late Pueblo 1175AD to 1400AD, Late Pueblo 1175AD to 1500AD, Late Pueblo 1175AD to 1545AD, Late Pueblo 1200AD to 1400AD, Late Pueblo 1275AD to 1400AD, Late Pueblo 1200AD to 1500AD	8

Project Category	Other Classifications	Count
Middle Archaic	Middle Archaic 2500BC to 800BC, Middle Archaic 3000BC to 1800BC	3
Middle Archaic to Late Archaic	Middle Archaic to Late Archaic 2500BC to 1000BC, Middle Archaic to Late Archaic 3000BC to 1000 BC, Middle Archaic to Late Archaic 3000BC to 200AD, Middle Archaic to Late Archaic 4850BC to 100AD, Middle to Late Archaic 2000BC to 200AD	9
Plains Woodland to Panhandle Aspect 250AD to 1400 AD	Plains Woodland to Panhandle Aspect 250AD to 1400 AD	1
Terminal Paleoindian 6600BC to 5500 BC	Terminal Paleoindian 6600BC to 5500 BC	1
Unknown/ Unspecific	Unknown, Unknown 25000 BC to 1994 AD, Unknown 25000 BC to 1995 AD, Unknown 25000 BC to 1996 AD, Unknown 25000 BC to 1997 AD, Unknown 25000 BC to 9999 AD, Unknown 9500 BC to 1993 AD, Unknown 9500 BC to 1997 AD, Unknown 9500 BC to 1999 AD, Unknown 9500 BC to 9999 AD, Unspecific 1300 AD to 1600AD, Unspecific 500 AD to 1750AD, Unspecific 750 AD to 1840AD, Unspecific 9500 BC to 1550AD, Unspecific 9500 BC to 1840AD, Unspecific 9500 BC to 1850AD, Unspecific 9500 BC to 1860AD, Unspecific 9500 BC to 1870AD, Unspecific 9500 BC to 1880AD, Unspecific 9500 BC to 1993AD, Unspecific 9500 BC to 1994AD, Unspecific 9500 BC to 1997AD, Unspecific 9500 BC to 1999AD, Unspecific 9500 BC to 2001AD, Unspecific 9500 BC to 9999 AD	541
Unspecific Archaic	Unspecific Archaic 4850 BC to 110AD, Unspecific Archaic 5500 BC to 200AD, Unspecific Archaic 5500BC to 900AD, Unspecific Archaic 5500BC-200AD	10
Unspecific Jornada	Unspecific Jornada 200AD to 1000AD, Unspecific Jornada 200AD to 1400AD, Unspecific Jornada 900AD to 1400AD	28

Table 44: Sites in the project area. LA- LA number, H- Hearth, RM- Ring Midden/roasting pit/ mescal pit, B- Burial, C-Cave/Rockshelter, S- shell midden, M- midden, Q- quarry, R-mortar, O- other, V- Number of visits.

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
1770	Late Pueblo 1175AD to 1400AD	0	0		1								1
1774	Early Pueblo to Late Pueblo AD 1100-1400											1	2
1775	Early Pueblo to Late Pueblo AD 1175-1400											4	1
8049	Unspecific Archaic 5500BC to 900AD	0	Unknown		1			1					2
8050	Unspecific 9500 BC to 1840AD	0	Unknown		7								1
8051	Unspecific 9500 BC to 1840AD	0	Unknown		1						1		3
9052	Archaic 5500BC to 900AD											1	1
14000	Historic	0	100 to 999		1								2

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
14179	Historic											4	1
14288	Early Pueblo 1100AD to 1200AD											4	2
14289	Late Pithouse to Late Pueblo 900 AD to 1400 AD	1 to 9	10 to 99	3	11							4	2
14304	Unspecific 9500 BC to 1840AD	Unknown	Unknown		2								1
15919	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	1
16462	Unspecific 9500 BC to 1840AD	0	1 to 9		1								1
17787	Unspecific 9500 BC to 1840AD	0	10 to 99		1								1
18152	Unspecific 9500 BC to 1840AD	0	0		1								1
20974	Unspecific 9500 BC to 1840AD	0	0		1								2
28684	Unspecific 9500 BC to 1840AD	0	10 to 99										1
28685	Unspecific 9500 BC to 1840AD	0	1000 to 9999	52	6				1			1	2
28752	Unspecific 9500 BC to 1840AD	0	Unknown		1								1
29504	Apache 1870AD 1890AD												1
29505	Unspecific 9500 BC to 1840AD	0	10 to 99	1	2								1
30625	Unspecific 9500 BC to 1840AD	0	10 to 99		3								2
33077	Late Pueblo 1200AD to 1400AD	0	10 to 99				1						1
33078	Unspecific 9500 BC to 1880AD	0	100 to 999	1	4								2
33970	Late Pithouse to Late Pueblo 750AD to 1400AD	0	1000 to 9999	4	8								3
33971	Unspecific 9500 BC to 1840AD	0	0		2								1
35557	Unspecific 9500 BC to 1840AD	0	0		1								2
35681	Unspecific 9500 BC to 1840AD	0	1000 to 9999										1
36609	Unspecific Jornada 200AD to 1400AD	Unknown	Unknown										2
36610	Unspecific Jornada 200AD to 1400AD	Unknown	Unknown										2
38193	Historic	1 to 9	100 to 999	4	4								3
38194	Unknown 9500 BC to 1993 AD	0	1 to 9		1								1
38195	Unknown 9500 BC to 1993 AD	0	100 to 999	8									1
38196	Unknown 9500 BC to 1993 AD	0	10 to 99										3
38197	Unknown 9500 BC to 1993 AD	0	10 to 99		1								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
38198	Unspecific 9500 BC to 1993AD	0	0				1						1
38199	Unspecific 9500 BC to 1880AD, Historic	1 to 9	10 to 99										1
38201	Unknown 9500 BC to 1993 AD	10 to 99	10 to 99										1
38202	Late Archaic 1800 BC to 200AD	1 to 9	100 to 999	1 4								1	3
38203	Unknown 9500 BC to 1993 AD	0	10 to 99										5
38204	Unknown 9500 BC to 1993 AD	10 to 99	100 to 999								2	7	6
38205	Unknown 9500 BC to 1993 AD	Unknown	Unknown		10	1						1	6
38206	Unknown 9500 BC to 1993 AD	0	Unknown		7								2
38207	Unknown 9500 BC to 1993 AD	0	0	1	5		2						1
38208	Unknown 9500 BC to 1993 AD	0	10 to 99				1			1			1
38209	Unknown 9500 BC to 1993 AD	0	1 to 9		5		1						1
38210	Unspecific 9500 BC to 1880AD	0	10 to 99		3								4
38211	Unknown 9500 BC to 1993 AD	0	1000 to 9999		14					1			2
38212	Unknown 9500 BC to 1993 AD	0	1000 to 9999	2	7								4
38213	Unknown 9500 BC to 1993 AD	0	100 to 999										1
38214	Unknown 9500 BC to 1993 AD	0	1 to 9		1								1
38215	Late Archaic 750AD to 950AD, Historic	1 to 9	1000 to 9999										5
38216	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	1
38217	Unknown 9500 BC to 1993 AD	0	10 to 99		4								1
38731	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	10 to 99	6	5								2
43426	Unspecific 9500 BC to 1840AD	0	10 to 99										1
43427	Early Pueblo to Late Pueblo 1100AD to 1400AD												1
43428	Late Pithouse to Late Pueblo 750AD to 1400AD	10 to 99	100 to 999	5	3								1
43429	Late Pithouse to Late Pueblo 750AD to 1400AD	Unknown	Unknown		3							1	1
43430	Early Pueblo 1100AD to 1200AD												2
43431	Late Pithouse to Early Pueblo 750AD to 1200AD	1 to 9	10 to 99		1								2
43432	Early Archaic to Late Archaic 4000BC to 1000AD, Late Pithouse to Late Pueblo (750 AD) to (1400 AD)	1 to 9	100 to 999		2								3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
43433	Late Pithouse to Late Pueblo 750AD to 1400AD	0	10 to 99										2
43434	Late Pithouse to Late Pueblo 750 AD to 1400 AD	0	1 to 9		1							1	3
43435	Unspecific 9500 BC to 1840AD	0	10 to 99		4								2
43436	Unspecific 9500 BC to 1840AD	Unknown	Unknown		1								2
43437	Late Pithouse to Late Pueblo 750AD to 1400AD	0	10 to 99		1								1
43438	Unspecific 9500 BC to 1880AD	1 to 9	10 to 99		7								2
43439	Early Pueblo to Late Pueblo 1150AD to 1250AD											2	1
43440	Late Archaic 200AD to 1200 AD, Spanish Contact to Mexican 1539 AD to 1846 AD	Unknown	Unknown		3								1
43441	Unspecific 9500 BC to 1840AD	0	10 to 99		3								3
43442	Unspecific 9500 BC to 1840AD	0	Unknown	2	5							1	3
43443	Unspecific 9500 BC to 1840AD	0	1 to 9		4								3
43444	Late Archaic 1000BC to 750AD, Late Pithouse to Late Pueblo (900 AD) to (1350 AD)	0	10 to 99		2								2
43445	Unspecific 9500 BC to 1840AD	0	100 to 999										4
43446	Early Pueblo to Late Pueblo AD 1100-1400												1
43447	Late Pueblo 1200AD to 1400AD, Spanish Contact to Mexican 1539 AD to 1846 AD	1 to 9	1 to 9	6	2							1	3
43448	Unspecific 9500 BC to 1840AD	0	10 to 99										2
43449	Early Pueblo to Late Pueblo 1100 AD to 1400 AD												1
43450	Unspecific 9500 BC to 1840AD	0	Unknown										1
43451	Early Pueblo to Late Pueblo 1100 AD to 1400 AD, Apache 1539AD to 1846AD	Unknown	Unknown		1							8	1
43460	Historic	0	10 to 99										1
43496	Late Pithouse to Late Pueblo (900 AD) to (1350 AD)	Unknown	Unknown				1		1				1
43507	Unspecific 9500 BC to 1840AD	0	100 to 999		2								2
43537	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	2
43671	Unspecific 9500 BC to 1840AD	0	1 to 9		1								1
43672	Apache 1539 AD to 1846AD						1						1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
43673	Late Archaic 1800BC to 200AD, Late Pithouse to Late Pueblo 900AD to 1400AD	1 to 9	1000 to 9999	1									2
43674	Unspecific Jornada 200AD to 1400AD	Unknown	Unknown		3		1		1		2		4
43675	Unspecific Archaic 5500 BC to 200AD, Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999	2	5						1	2	3
43676	Unspecific 9500 BC to 1840AD	Unknown	Unknown			1							1
43677	Late Pithouse to Late Pueblo 900AD to 1400AD	1 to 9	10 to 99		1								2
43679	Early Pueblo to Late Pueblo AD 1100-1400											5	1
43680	Late Pithouse to Late Pueblo 750AD to 1400AD	10 to 99	10 to 99		1								1
43682	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	1000 to 9999	10	6								3
43684	Early Pueblo 1100AD to 1200AD	0	0										2
43685	Late Pithouse to Early Pueblo 750AD to 1100AD, Historic	1 to 9	10 to 99										1
43704	Late Pithouse to Late Pueblo (750 AD) to (1400 AD), Apache AD 1539-1846	0	0									4	1
43732	Late Pithouse to Late Pueblo 750AD to 1200AD	10 to 99	100 to 999	25	2								1
43733	Early Pueblo to Late Pueblo AD 1100-1400											1	1
43734	Unspecific 9500 BC to 1840AD	0	Unknown	1									1
43735	Historic	0	1 to 9										1
43736	Unspecific 9500 BC to 1850AD	0	1 to 9		1								2
43737	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	1 to 9		16						5	2	2
43738	Late Pithouse to Late Pueblo 750AD to 1400AD	10 to 99	100 to 999	4	3								2
43741	Unknown 9500 BC to 1993 AD	0	10 to 99										1
45231	Unknown 9500 BC to 1993 AD	0	10 to 99										1
45861	Unspecific 9500 BC to 1880AD	0	100 to 999		1								3
45863	Late Archaic 820BC to 900AD, Late Pueblo 1200AD to 1500AD	1 to 9	1000 to 9999	1	4	7	3		2				6
45865	Unknown 9500 BC to 1993 AD	0	1 to 9	2	4								3
45866	Unknown 9500 BC to 1993 AD	0	1 to 9	3									1
45956	Unknown 9500 BC to 1993 AD	0	100 to 999	1	3								2

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
48735	Unspecific 9500 BC to 1993AD	0	0				1						1
48736	Early Pithouse to Late Pithouse 200AD to 750AD											4	2
48737	Unknown 9500 BC to 1993 AD	0	100 to 999		1								2
48738	Unspecific 750AD to 1840AD	0	0		1								2
48759	Unspecific 9500 BC to 1840AD	0	10 to 99										1
48772	Early Pueblo to Late Pueblo AD 1100-1400, Apache AD 1539-1846	0	0		1								1
48773	Late Archaic (1486 BC) to Late Archaic (200 AD), Unspecific 9500 BC to 1880AD	0	10 to 99										1
48774	Unspecific 9500 BC to 1840AD	Unknown	Unknown	3									1
49461	Late Archaic 1800BC to 900AD	1 to 9	100 to 999		10				1				1
49462	Late Archaic 1800BC to 900 AD	1 to 9	100 to 999		1							1	1
49463	Unspecific 9500 BC to 1840AD	0	Unknown	1									1
50484	Unspecific 9500 BC to 1880AD	0	100 to 999		2								2
51007	Early Pueblo to Late Pueblo 950AD to 1300AD, Historic	0	100 to 999	1									3
55986	Late Archaic to Unspecific Archaic 800BC to 200AD, Historic	Unknown	10 to 99		1								1
61245	Unspecific Archaic 5500BC to 900AD	0	1000 to 9999		1								1
61246	Unspecific 9500 BC to 1880AD	0	1 to 9		12								3
61348	Unspecific 9500 BC to 1840AD	Unknown	Unknown		4								1
61349	Unspecific 750 AD to 1840AD, Historic	1 to 9	0		1								1
64499	Unspecific 9500 BC to 1840AD	0	1 to 9		1								1
65400	Late Pithouse to Late Pueblo 900AD to 1400AD, Historic	1 to 9	10 to 99		1								1
67384	Unspecific 9500 BC to 1840AD	0	1000 to 9999	1	2								3
67513	Terminal Paleoindian 6600BC to 5500 BC, Archaic 5500BC to 200AD	0	0									1	1
67514	Historic	0	1000 to 9999		1							2	4
67515	Unspecific 9500 BC to 1880AD	0	10 to 99										2
67519	Early Archaic 5500 BC to 3000BC					1							1
67520	Unspecific 9500 BC to 1550AD	0	10 to 99		2							6	1
67529	Late Pithouse to Early Pueblo 750AD to 1175AD	10 to 99	10 to 99	1	5								2
67530	Unspecific 9500 BC to 1840AD	0	1 to 9	6	4								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
67531	Historic	0	100 to 999	14	5				1			1	1
67854	Late Archaic 1800BC to 900AD	1 to 9	10 to 99		1								2
67855	Late Pithouse 750AD to 1100AD	Unknown	Unknown		1								1
67856	Unspecific 9500 BC to 1840AD	0	1 to 9		1								1
67861	Unspecific 9500 BC to 1880AD	0	10 to 99									3	4
67862	Late Pithouse 900AD to 1100AD, Historic	0	10 to 99	1									1
67863	Late Pithouse to Late Pueblo 750AD to 1400AD	Unknown	Unknown		1	1	1						1
68252	Unknown 9500 BC to 1993 AD	0	100 to 999										1
68481	Unspecific 750AD to 1840AD	0	1 to 9	6	4								1
69032	Unspecific 9500 BC to 1840AD	Unknown	Unknown		4								1
69033	Historic	Unknown	1000 to 9999		28						50		6
69034	Unspecific 9500 BC to 1840AD	0	10 to 99										1
69035	Unspecific 9500 BC to 1840AD	0	Unknown		4								2
70107	Unspecific 9500 BC to 1840AD	0	10 to 99		1								1
71989	Apache 1539 AD to 1846AD, Historic	Unknown	Unknown			1	1						2
72256	Early Pithouse to Late Pueblo 700AD to 1600AD											4	1
72265	Unspecific 9500 BC to 1840AD	0	10 to 99		3								1
72390	Unspecific 9500 BC to 1840AD	0	10 to 99		2								3
73721	Late Pithouse 900AD to 1100AD	Unknown	10 to 99		2								2
73722	Unspecific 9500 BC to 1880AD	0	10 to 99										2
73723	Unspecific 9500 BC to 1880AD	Unknown	100 to 999	3									3
75564	Unspecific 9500 BC to 1840AD	0	10 to 99										1
76486	Unspecific 9500 BC to 1840AD	Unknown	Unknown										1
76487	Unspecific 9500 BC to 1840AD	0	100 to 999										
76488	Unspecific 9500 BC to 1840AD	0	100 to 999		1								2
77952	Unspecific 9500 BC to 1840AD	0	100 to 999										1
77953	Unspecific 9500 BC to 1840AD	0	Unknown		1								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
77954	Unspecific 9500 BC to 1840AD	0	10 to 99		1								2
77955	Unspecific 9500 BC to 1840AD	0	Unknown		2								1
77956	Unspecific 9500 BC to 1840AD	0	10 to 99										1
77957	Unspecific 9500 BC to 1840AD, Historic	1 to 9	100 to 999										1
77958	Historic	0	100 to 999	16	3								3
78425	Late Pithouse to Late Pueblo 750AD to 1450AD	0	10 to 99										
78426	Unspecific 9500 BC to 1840AD	0	0		1								2
78427	Unspecific 9500 BC to 1840AD	Unknown	0	5									1
78428	Unspecific 9500 BC to 1840AD	0	10 to 99	2									2
79813	Unspecific 9500 BC to 1840AD	0	0		1								1
81502	Unspecific 9500 BC to 1880AD	0	10 to 99	2	3								2
82638	Unspecific 9500 BC to 1840AD	0	10 to 99	2	7								4
82639	Unspecific 9500 BC to 1840AD	0	Unknown		3								1
82640	Unspecific 9500 BC to 1840AD	0	0		1								1
82641	Unspecific 9500 BC to 1840AD	0	1 to 9		5								2
82642	Unspecific 9500 BC to 1840AD	Unknown	Unknown		4								2
84548	Unspecific 9500 BC to 1840AD	0	0	1	7							9	3
84549	Unspecific 9500 BC to 1840AD, Historic	1 to 9	1 to 9		1	4	1		1				1
84550	Unspecific 9500 BC to 1840AD	0	Unknown		4								1
84551	Unspecific 9500 BC to 1840AD	0	10 to 99		2								1
84552	Unspecific 9500 BC to 1840AD	Unknown	Unknown	4									1
84553	Late Pithouse 900AD to 1100AD	0	0									2	2
84554	Unspecific 9500 BC to 9999 AD	0	10 to 99										1
84820	Late Archaic 2300 BC to 1100AD	Unknown	Unknown		3								2
84992	Late Pithouse to Late Pithouse 900AD to 1100AD	1 to 9	10 to 99		1								1
84993	Unspecific 9500 BC to 1840AD	0	0		1								3
84994	Unspecific 9500 BC to 1840AD	0	Unknown										1
84995	Unspecific 9500 BC to 1840AD	0	10 to 99		3								5

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
84996	Unspecific 9500 BC to 1840AD	0	0		1								2
86122	Unspecific 9500 BC to 1840AD	0	10 to 99										1
86197	Unspecific 9500 BC to 1840AD	Unknown	Unknown		2								1
87012	Unspecific 9500 BC to 1840AD	0	10 to 99										1
87013	Unspecific 9500 BC to 1840AD	0	1 to 9										1
87014	Historic	1 to 9	100 to 999	1	7								5
87015	Late Pueblo 1175AD to 1545AD	0	10 to 99				1						1
87038	Unspecific 9500 BC to 1840AD	0	1 to 9										1
88071	Late Archaic 1500BC to 300AD, Late Pithouse to Late Pueblo (900 AD) to (1600 AD)	10 to 99	100 to 999		1								1
88108	Historic											9	1
89375	Unspecific 9500 BC to 1994AD	0	100 to 999		1		1						1
89376	Late Pithouse to Late Pueblo 900AD to 1400AD	0	10 to 99	1	6								3
89377	Late Pithouse to Late Pueblo (900 AD) to (1400 AD)	1 to 9	1000 to 9999		28				1			3	2
89378	Unspecific 9500 BC to 1994AD	0	1 to 9	1	7								2
89379	Unspecific 9500 BC to 1994AD	0	10 to 99										1
89380	Unspecific 9500 BC to 1840AD, Historic	10 to 99	10 to 99	1	1								5
89381	Late Pithouse to Late Pueblo 750 AD to 1400 AD	1 to 9	100 to 999		38								5
89382	Late Pithouse to Late Pueblo 900AD to 1400AD	0	10 to 99										1
89383	Late Pithouse to Late Pueblo 900 AD to 1400 AD	0	10 to 99										1
89529	Unspecific 9500 BC to 1994AD	0	10 to 99										1
89530	Unspecific 9500 BC to 1994AD	0	10 to 99										1
89531	Unspecific 9500 BC to 1840AD	0	1 to 9										1
89532	Late Pithouse to Late Pueblo 900 AD to 1450 AD	10 to 99	100 to 999	3	1								2
89857	Unspecific 9500 BC to 1840AD	0	10 to 99										1
89859	Unspecific 9500 BC to 1840AD	0	10 to 99										1
89860	Unspecific 9500 BC to 1840AD	0	10 to 99										1
89861	Unspecific 9500 BC to 1840AD, Historic	Unknown	Unknown	4	9								2
89878	Unspecific 9500 BC to 1840AD	0	1 to 9		1								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
98320	Unspecific 9500 BC to 1840AD	0	100 to 999		6								1
98322	Late Archaic 1800BC to 200AD	Unknown	Unknown										2
98804	Unspecific 9500 BC to 1840AD	0	1 to 9										1
98805	Unspecific 9500 BC to 1840AD	0	1 to 9	5									3
99464	Late Pithouse to Late Pueblo 900AD to 1450AD	0	0		6								1
99932	Unspecific 9500 BC to 9999AD	0	100 to 999									2	3
101201	Middle Archaic 2500BC to 800BC	0	10 to 99										1
101202	Unknown 9500 BC to 1993 AD	0	1 to 9		12								4
101308	Unspecific 9500 BC to 1840AD	0	1 to 9	1	2								3
101494	Unspecific 9500 BC to 1880AD, Unspecific Jornada 200AD to 1400AD	Unknown	100 to 999	1	2							2	3
101495	Unspecific Jornada 200AD to 1400AD	Unknown	Unknown			1	1						2
103221	Unknown 25000 BC to 1994 AD	0	0			1	1						2
103222	Unknown 25000 BC to 1994 AD	Unknown	Unknown								8		1
103223	Unknown 25000 BC to 1994 AD	Unknown	Unknown	1									1
103224	Unknown 25000 BC to 1994 AD	Unknown	1 to 9	1									2
103225	Unknown 25000 BC to 1994 AD	Unknown	1 to 9	1	1								2
103226	Historic	10 to 99	100 to 999		6								1
103227	Unknown 25000 BC to 1994 AD	Unknown	Unknown	1									1
103228	Unknown 25000 BC to 1994 AD	Unknown	Unknown		2								1
104447	Unknown 25000 BC to 1995 AD, Historic	1 to 9	10 to 99		2								3
104544	Late Pithouse to Early Pueblo 1000AD to 1200AD	1 to 9	100 to 999		47								1
104545	Unspecific 9500 BC to 1550AD	0	100 to 999		4								1
104546	Late Pithouse to Late Pithouse 750AD to 1100AD	0	100 to 999		4								1
104948	Unspecific 9500 BC to 1550AD	0	1 to 9										1
104982	Unspecific 9500 BC to 1840AD	Unknown	1 to 9	1	1								2
104983	Unspecific 9500 BC to 1880AD	0	10 to 99										2
105253	Unspecific 9500 BC to 1550AD	0	100 to 999										3
105254	Unspecific 9500 BC to 1550AD	0	100 to 999										1
105505	Unspecific Archaic 5500 BC to 200AD	1 to 9	10 to 99										1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
106005	Unspecific 9500 BC to 1550AD	0	10000 to 99999		1							1	4
106049	Unspecific 9500 BC to 1880AD	0	100 to 999										2
107125	Unknown 25000 BC to 1994 AD	Unknown	Unknown		1								1
107436	Clovis to Late Paleoindian 9500 BC to 6600 BC											1	1
107437	Late Archaic 1800BC to 200AD, Apache 1840AD to 1870AD	0	0									9	2
107743	Unspecific 9500 BC to 1550AD	0	0		4								1
107902	Unspecific 9500 BC to 1550AD	0	0	7	6								2
107903	Unspecific 9500 BC to 1550AD	0	0		6								1
107937	Late Pithouse 900AD to 1040AD	Unknown	Unknown		1								1
107938	Unspecific 9500 BC to 1550AD	0	10 to 99		11								1
107939	Unspecific 9500 BC to 1880AD	0	100 to 999		1							1	4
108063	Unspecific 9500 BC to 1550AD	0	10 to 99		7								1
108064	Unspecific 9500 BC to 1550AD	0	0		8								1
108138	Early Pithouse to Late Pueblo 200AD to 1100AD											1	1
108348	Late Pithouse to Late Pueblo 900AD to 1400AD	1 to 9	10000 to 99999		400		3				3	16	1
108349	Middle Archaic to Late Archaic 3000BC to 1000 BC	0	10 to 99	1	1								
108350	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	100 to 999		1								1
108376	Early Pithouse to Early Pueblo 600AD to 1200AD											3	3
108377	Unspecific 9500 BC to 1550AD	0	0		2								1
108378	Unspecific 9500 BC to 1550AD	0	10 to 99										2
108556	Unspecific 9500 BC to 1550AD	0	1 to 9		33								1
108557	Unspecific 9500 BC to 1550AD	0	0		8								1
108626	Late Archaic 1800BC to 200AD	1 to 9	10 to 99	8	16								1
108627	Late Pithouse to Late Pueblo 900AD to 1400AD	0	1 to 9										
108628	Unspecific 9500 BC to 1550AD	0	0		3							1	1
108641	Unspecific 9500 BC to 1550AD	0	0		7								1
108642	Unspecific 9500 BC to 1550AD	0	0	1									3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
108643	Unspecific 9500 BC to 1550AD	0	1 to 9	3	9								3
108644	Unspecific 9500 BC to 1550AD	0	1000 to 9999										1
108645	Unspecific 9500 BC to 1550AD, Historic	1 to 9	1000 to 9999	65	70							9	2
108646	Unspecific 9500 BC to 1550AD	0	0		4							1	2
108647	Unspecific 9500 BC to 1550AD, Historic	0	100 to 999										1
109230	Late Archaic 1065 BC to 145AD	0	10 to 99		2								2
109323	Early Archaic to Late Archaic 5500BC to 200AD, Early Pueblo 1100AD to 1200AD	0	100 to 999										1
109335	Early Pithouse to Late Pueblo 200AD to 1400AD											1	1
109909	Unspecific 9500 BC to 1550AD, Unspecific 1550AD to 1996 AD	Unknown	1 to 9		29								2
109910	Unspecific 9500 BC to 1550AD, Historic	10 to 99	100 to 999		20							1	4
110143	Middle Archaic 3000BC to 1800BC, Early Pithouse to Late Pueblo 200AD to 1400AD	0	10 to 99	2	4								1
110147	Unspecific 9500 BC to 1880AD, Historic	0	10 to 99										1
110359	Early Pueblo 1100AD to 1200AD											1	3
110385	Unspecific 9500 BC to 1880AD	0	100 to 999		2								2
110387	Early Pithouse to Late Pueblo 500AD to 1400AD, Historic	10 to 99	1000 to 9999	14	19				1				3
110388	Early Pueblo to Late Pueblo 1100AD to 1400AD			1								2	2
110389	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	10 to 99	1	11								3
110513	Unspecific 9500 BC to 1550AD	0	1 to 9		2								1
110660	Unspecific 9500 BC to 1550AD, Historic	1 to 9	10 to 99		7								2
110909	Unspecific 9500 BC to 1550AD	0	0		1								1
111137	Unspecific 9500 BC to 1550AD	0	1 to 9		3								1
111138	Unspecific 9500 BC to 1880AD	0	10 to 99	3									1
111215	Late Archaic 1800BC to 200AD, Late Pithouse 750AD to 1400AD	1 to 9	Unknown	1									1
111216	Early Pithouse to Late Pueblo 650AD to 1300AD											2	1
112327	Early Pithouse to Late Pueblo 200AD to 1400AD			1								3	1
112347	Unspecific 9500 BC to 1550AD	0	0		1								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
112610	Unspecific 9500 BC to 1550AD	0	1 to 9		4								1
112611	Unspecific 9500 BC to 1550AD	0	1 to 9		53								1
112612	Unspecific 9500 BC to 1550AD	0	0		1								1
112613	Unspecific 9500 BC to 1550AD	0	0	1									1
112614	Unspecific 9500 BC to 1880AD	0	100 to 999		4								3
112615	Unspecific 9500 BC to 1550AD	0	0		2								1
112616	Unspecific 9500 BC to 1550AD	Unknown	Unknown										2
112617	Unspecific 9500 BC to 1550AD	0	10 to 99	1									1
112618	Unspecific 9500 BC to 1550AD	0	10 to 99										1
112619	Unspecific 9500 BC to 1550AD	0	10 to 99	1	3								1
112620	Unspecific 9500 BC to 1550AD	0	1 to 9		1								2
112621	Unspecific 9500 BC to 1550AD	0	100 to 999										3
112622	Unspecific 9500 BC to 1550AD	0	10 to 99										1
112623	Unspecific 9500 BC to 1550AD	0	10 to 99										1
112624	Unspecific 9500 BC to 1550AD	0	100 to 999										2
112625	Unspecific 9500 BC to 1550AD	0	10 to 99	2	6								1
112626	Unspecific 9500 BC to 1840AD	Unknown	Unknown	1	2								3
112627	Unspecific 9500 BC to 1550AD	0	10 to 99	1	3								1
112628	Unspecific 9500 BC to 1550AD	0	10 to 99		4						1	1	1
112629	Unspecific 9500 BC to 1550AD	0	10 to 99		4								1
112632	Unspecific 9500 BC to 1550AD	0	10 to 99										1
112633	Unspecific 9500 BC to 1550AD	0	10 to 99										4
112634	Unspecific 9500 BC to 1550AD	0	10 to 99										2
112635	Unspecific 9500 BC to 1550AD	0	100 to 999										1
112636	Unspecific 9500 BC to 1550AD	0	10 to 99		1								1
112637	Unspecific 9500 BC to 1550AD	0	10 to 99										1
112638	Unspecific 9500 BC to 1550AD	1 to 9	100 to 999										1
112639	Historic	0	1000 to 9999		1								2
112995	unknown	0	10 to 99	5									1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
113067	excavated											4	1
113503	Late Pithouse to Late Pueblo 900AD to 1400AD, Historic	1 to 9	10 to 99	1									2
113504	Unspecific 9500 BC to 1880AD	0	10 to 99		7								4
113505	Unspecific 9500 BC to 1550AD	0	100 to 999										1
113528	Unknown	0	10 to 99										1
113534	Late Pueblo 1200AD to 1400AD	0	1 to 9				1						1
113535	Early Pueblo to Late Pueblo 1075AD to 1375AD											2	1
113537	Late Pithouse to Late Pueblo 900AD to 1400AD	0	10 to 99		1								1
113538	Unspecific 9500 BC to 1550AD	0	100 to 999										2
113539	Unspecific 9500 BC to 1550AD	0	0		7	1							2
113563	Unspecific 9500 BC to 1550AD	0	10 to 99										2
113606	Unspecific 9500 BC to 1550AD	0	0	1	2								1
113995	Unspecific 9500 BC to 1550AD	0	10 to 99										1
114090	Unspecific 9500 BC to 1550AD	0	10 to 99										2
114091	Unspecific 9500 BC to 1880AD	0	10 to 99										2
114134	Unknown 25000 BC to 1997 AD	0	Unknown	3									1
114136	Unknown 25000 BC to 1997 AD	0	1 to 9	5									1
114137	Unknown 25000 BC to 1997 AD	Unknown	Unknown	1									1
114139	Unknown 25000 BC to 1997 AD, LA 111215 & 114139 both are part of much enlarged site. See LA 114139 site visit comments.	1 to 9	0		4								1
114140	Unknown 25000 BC to 1997 AD	Unknown	Unknown	2									1
114141	Unknown 25000 BC to 1997 AD	Unknown	Unknown	1									2
114145	Unknown 9500 BC to 1997 AD	0	1 to 9		5								2
114146	Historic	0	10 to 99										1
114147	Late Archaic 100BC to 900AD	0	10 to 99		6								2
114148	Unknown 25000 BC to 1997 AD	Unknown	Unknown		1								1
114149	Late Pithouse to Late Pueblo 1100AD to 1500AD	1 to 9	10 to 99	3									3
114184	Late Pithouse to Late Pueblo 950AD to 1350AD	0	10 to 99	2									1
114193	Unknown 25000 BC to 1997 AD	Unknown	1 to 9	8									1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
114194	Late Pithouse 900AD to 1400AD	1 to 9	100 to 999		1								3
114195	Unknown 25000 BC to 1997 AD	Unknown	Unknown		1								1
114196	Unknown 25000 BC to 1997 AD	Unknown	Unknown		8								4
114197	Late Pithouse to Late Pueblo 900AD to 1400AD	0	100 to 999	4	4								1
114198	Early Pueblo to Late Pueblo 1100AD to 1350AD												2
114286	Unspecific 9500 BC to 1550AD	0	10 to 99										1
114487	Unknown 25000 BC to 1996 AD	Unknown	Unknown		2								1
114533	Unspecific 9500 BC to 1550AD	0	10 to 99										1
114739	Unknown 25000 BC to 1997 AD	0	10 to 99	1	1								1
114741	Unknown 25000 BC to 1997 AD	Unknown	Unknown		7								1
114749	Unspecific 9500 BC to 1550AD	0	100 to 999		2								1
114983	Early Pithouse to Late Pueblo 350AD to 1400AD												1
115031	Unspecific 9500 BC to 1550AD	0	10 to 99										1
115242	Unspecific 9500 BC to 1550AD	0	1 to 9		2								1
115569	Early Pithouse to Late Pueblo 650AD to 1400AD												1
115570	Unspecific 500 AD to 1750AD	0	1 to 9		1								1
115706	Historic	10 to 99	100 to 999	3	7		2		1				3
115707	Historic	0	100 to 999		1								4
115708	Historic	1 to 9	10 to 99	1	2								1
115864	Unspecific 9500 BC to 1550AD	0	10 to 99										2
115865	Unspecific 9500 BC to 1550AD	0	1 to 9		7							1	1
115866	Unspecific 9500 BC to 1550AD	0	10 to 99										1
115908	Unspecific 9500 BC to 1997AD	0	100 to 999										1
116027	Unspecific 9500 BC to 1550AD	0	10 to 99										1
116373	Unspecific 9500 BC to 1880AD	0	10 to 99										2
116398	Unknown 25000 BC to 1997 AD	Unknown	Unknown	4	1								2
116399	Unspecific 9500 BC to 1550AD	0	0		1								1
116400	Unknown 25000 BC to 1997 AD	0	Unknown		2								1
116403	Historic	0	100 to 999		8								3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
116404	Unknown 25000 BC to 1997 AD	0	Unknown	1									1
116405	Unknown 25000 BC to 1997 AD	0	Unknown		2								1
116470	Unspecific 9500 BC to 1870AD	0	100 to 999		24								3
116471	Late Archaic 500BC to 200AD, Early Pueblo to Late Pueblo AD 1100-1400	0	1000 to 9999	5									2
116472	Historic	1 to 9	100 to 999		11								3
116480	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	1
116523	Unspecific 9500 BC to 1880AD	0	1 to 9		1								1
116524	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	2
116525	Unspecific 9500 BC to 1550AD	0	1 to 9										1
117294	Late Pithouse to Late Pueblo 950AD to 1300AD, Historic		1 to 9		2								2
117366	Unspecific 9500 BC to 1880AD	0	10 to 99										2
117367	Unspecific 9500 BC to 1999AD	0	10 to 99										1
117368	Unspecific 9500 BC to 1999AD	0	100 to 999										1
117369	Unspecific 9500 BC to 1999AD	0	10 to 99		1								1
117370	Early Pithouse to Late Pueblo 2000AD to 1400AD											1	1
117371	Unspecific 9500 BC to 1550AD	0	10 to 99										1
117372	Unspecific 9500 BC to 1550AD	0	10 to 99										1
117373	Unspecific 9500 BC to 1550AD	0	100 to 999										1
117374	Late Archaic 1800BC to 200AD	1 to 9	1000 to 9999		14								1
117375	Unspecific 9500 BC to 1550AD	0	1000 to 9999										1
117376	Unspecific 9500 BC to 1550AD	0	100 to 999										1
117377	Unspecific 9500 BC to 1550AD	0	100 to 999									2	1
117445	Unspecific 9500 BC to 1550AD	0	100 to 999		1								1
117590	Unspecific 9500 BC to 1550AD	0	100 to 999										1
117591	Unspecific 9500 BC to 1550AD	0	100 to 999										2
117693	Unspecific 9500 BC to 1550AD	0	0		1								1
118059	Unspecific 9500 BC to 1550AD	0	100 to 999		1								1
118237	Late Pithouse to Late Pueblo 900AD to 1450AD	0	0		1								1
118565	Unspecific 9500 BC to 1840AD	0	10 to 99	1	2								3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
119364	Early Pueblo to Late Pueblo AD 1100-1400											8	1
119406	Historic	0	10 to 99										1
119407	Unspecific Archaic 5500BC-200AD	0	10 to 99		4								3
119408	Historic	1 to 9	1000 to 9999		2								1
119410	Late Pithouse to Late Pithouse (750 AD) to (1100 AD)	10 to 99	100 to 999	5	8				1				1
119411	Unspecific Jornada 200AD to 1400AD	0	10 to 99		2								2
119791	Unspecific 9500 BC to 1550AD	0	0		2								1
119792	Unspecific 9500 BC to 1550AD	0	0		1								1
119793	Early Pueblo to Late Pueblo 1100AD to 1400AD											2	1
119794	Unspecific Archaic 5500BC to 200AD, Unspecific Jornada 200AD to 1400AD	Unknown	Unknown	1	1								1
119795	Unspecific Archaic 5500BC to 200AD	1 to 9	100 to 999		1								2
119796	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	10 to 99		1								2
119797	Unspecific Jornada 200AD to 1400AD	0	100 to 999										1
119798	Unspecific 9500 BC to 1999AD	0	10 to 99										1
119799	Unspecific Jornada 900AD to 1400AD	0	100 to 999		2								1
119800	Unspecific Jornada 900AD to 1400AD	1 to 9	10 to 99		2								1
119802	Historic	0	10 to 99										2
119803	Unspecific 9500 BC to 1999AD	0	10 to 99										1
119804	Unspecific 9500 BC to 1999AD	0	10 to 99		3								1
119805	Unspecific 9500 BC to 1550AD	0	0		2								1
119806	Historic	0	10 to 99										1
119808	Unspecific Jornada 200AD to 1400AD	0	10 to 99		3								1
119809	Unspecific 9500 BC to 1550AD	0	100 to 999										1
119810	Unspecific 9500 BC to 1550AD	0	10 to 99										2
119961	Unspecific 9500 BC to 1550AD	0	100 to 999	5	5							3	1
120233	Unspecific 9500 BC to 1550AD	0	10 to 99										1
120234	Unspecific 9500 BC to 1550AD	0	100 to 999	1	1								2
120235	Unspecific 9500 BC to 1550AD	0	10 to 99	2	1								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
120236	Unspecific 9500 BC to 1550AD	0	10 to 99	2	1								1
120237	Unspecific 9500 BC to 1550AD	1 to 9	1000 to 9999										1
120238	Late Pithouse 850AD to 1000AD	Unknown	Unknown		2						1		3
120239	Late Archaic 1800BC to 200AD, Late Pithouse to Late Pueblo 900AD to 1400AD	Unknown	1 to 9		2								3
120240	Unspecific Jornada 200AD to 1400AD	0	1000 to 9999		11								1
120241	Late Pithouse to Late Pithouse 850AD to 1000AD	100 to 999	100 to 999	2	27							2	2
120242	Unspecific Jornada 200AD to 1400AD	0	10 to 99		2								1
120243	Unspecific 9500 BC to 1880AD	0	10 to 99										1
120244	Unspecific 9500 BC to 1880AD	0	1 to 9										2
120245	Unspecific Jornada 200AD to 1400AD	0	1 to 9		1								1
120246	Unspecific 9500 BC to 1550AD	0	10 to 99										1
120247	Unspecific 9500 BC to 1550AD	0	100 to 999										1
120248	Unspecific 9500 BC to 1550AD	0	10 to 99										1
120249	Unspecific 9500 BC to 1550AD	0	100 to 999										1
120250	Unspecific 9500 BC to 1550AD	0	10 to 99										1
120251	Historic	0	100 to 999										2
120252	Late Archaic 1000BC to 200AD, Late Pithouse to Late Pueblo 900AD to 1400AD	10 to 99	100 to 999		1								1
120302	Late Archaic 1800BC to 200AD	Unknown	Unknown		2								1
120303	Unspecific 9500 BC to 1999AD	0	0		1								1
120304	Late Pithouse to Early Pueblo 1000AD to 1200AD	1 to 9	1000 to 9999	3	6						1	5	5
120305	Unknown 9500 BC to 1999 AD	0	10 to 99										1
120306	Unspecific 9500 BC to 1550AD	0	10 to 99										1
120307	Unspecific 9500 BC to 1550AD	0	1 to 9	1	1				1				1
120308	Unspecific 9500 BC to 1550AD	0	1 to 9										1
120309	Late Archaic 1800BC to 200AD, Late Pithouse 850AD to 1000AD	Unknown	Unknown		2								1
120310	Historic	0	10 to 99										1
120311	Unspecific Jornada 200AD to 1400AD	10 to 99	10 to 99									2	1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
120312	Unspecific 9500 BC to 1999AD	0	100 to 999										1
120313	Unspecific 9500 BC to 1999AD	0	1 to 9		1								1
120314	Middle to Late Archaic 3000BC to 200AD	Unknown	Unknown		1								1
120315	Early Pithouse to Late Pueblo 1000 AD to 1400AD											1	1
120316	Middle to Late Archaic 2000BC to 200AD	Unknown	Unknown	1									1
120317	Unspecific 9500 BC to 1999AD	0	10 to 99		1								1
120318	Unspecific Jornada 200AD to 1400AD	10 to 99	100 to 999		10								2
120319	Unspecific Archaic 5500BC to 200AD, Unspecific Jornada 200AD to 1400AD	Unknown	Unknown										2
120320	Unspecific 9500 BC to 1550AD	0	10 to 99		1								2
120321	Unspecific 9500 BC to 1999AD	0	10 to 99										1
120377	Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999		2								2
120378	Historic	0	10 to 99		1								1
120379	Unspecific 9500 BC to 1550AD	0	10 to 99	7	1								1
120380	Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999		2								1
120381	Unspecific Jornada 200AD to 1400AD	1 to 9	10 to 99		1								1
120382	Unspecific 9500 BC to 1550AD	0	1 to 9		1								1
120383	Unspecific Jornada 200AD-1400AD	10 to 99	100 to 999		1								1
120384	Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999		2								1
120385	Unspecific 9500 BC to 1550AD	0	Unknown	1	1				1				2
120386	Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999		4								1
120387	Unspecific Jornada 200AD to 1000AD	0	10 to 99										2
120388	Unspecific Jornada 200AD to 1400AD	10 to 99	100 to 999		1								1
120389	Late Pithouse to Late Pueblo (1000 AD) to (1300 AD)	1 to 9	100 to 999		41								5
120390	Unspecific 9500 BC to 1999AD	0	0									1	1
120391	Unspecific 9500 BC to 1999AD	0	10 to 99										2
120392	Unspecific 9500 BC to 1550AD	0	1 to 9	2									2
120393	Late Archaic 1800BC to 200AD	1 to 9	10 to 99		2							6	1
120394	Unspecific Jornada 200AD to 1400AD	1 to 9	100 to 999		2								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
120395	Late Archaic 1800BC to 200AD	0	100 to 999										2
120396	Early Pueblo to Late Pueblo 1100AD to 1400AD											1	1
120498	Late Paleoindian 8000BC to 6600BC	Unknown	Unknown										1
120512	Unspecific 9500 BC to 1999AD	0	10 to 99										1
120513	Late Pithouse to Late Pueblo 1000AD to 1200AD	0	100 to 999	2	1								6
120514	Unknown 9500 BC to 1999 AD	0	10 to 99										4
120645	Early Pithouse to Late Pueblo 200AD to 1400AD											8	2
120646	Historic	0	1000 to 9999		5								2
120716	Late Archaic 1000BC to 200AD	0	100 to 999		1								2
120717	Unspecific 9500 BC to 1550AD	0	1 to 9						1				1
120719	Late Pithouse to Late Pueblo 900AD to 1400AD												
120720	Early Pithouse to Late Pueblo 1AD to 1375AD											1	1
120943	Unspecific 9500 BC to 1880AD	0	10 to 99										1
121074	Late Pithouse to Late Pueblo 900AD to 1450AD	0	1 to 9	1	14								1
121075	Unspecific 9500 BC to 1999AD	0	10 to 99	5									1
121118	Unspecific 9500 BC to 1550AD	0	10 to 99										1
121119	Unspecific 9500 BC to 1550AD	0	1 to 9		2								2
121138	Unspecific 9500 BC to 1880AD	0	10 to 99										1
121206	Late Pithouse to Late Pueblo 900AD to 1400AD	0	10000 to 99999							1			3
121207	Unspecific 9500 BC to 1880AD	0	10 to 99										1
121289	Middle Archaic to Late Archaic 3000BC to 200AD	0	10 to 99										1
121290	Late Pithouse to Late Pueblo 750AD to 1400AD	Unknown	Unknown		8								1
121291	Late Pithouse to Late Pueblo 900AD to 1400AD	0	100 to 999										2
121292	Late Pithouse to Late Pueblo 900AD to 1450AD	0	0		3								3
121294	Unspecific 9500 BC to 1880AD	0	100 to 999										2
121499	Unspecific 9500 BC to 1880AD	0	100 to 999		1								1
121500	Unspecific 9500 BC to 1880AD	0	10 to 99										2
121703	Unspecific 9500 BC to 9999 AD	1 to 9	1000 to 9999		13								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
121704	Unspecific 9500 BC to 1840AD	0	1 to 9	16	2								2
121708	Early Pithouse to Late Pithouse 500AD to 1000AD											1	1
121709	Late Archaic 1AD to 1000AD	10 to 99	10 to 99	1	8		3						2
121856	Unspecific 9500 BC to 1880AD	0	10 to 99										2
121857	Early Pueblo to Late Pueblo 1100AD to 1350AD												
121969	Unspecific 9500 BC to 1880AD	0	0	1									1
122138	Unspecific 9500 BC to 1880AD	0	1 to 9		5								5
125151	Apache 1850 AD to 1879AD									2			1
125281	Unspecific 9500 BC to 1880AD	0	10 to 99	1									1
125721	Apache 1867AD to 1875AD									1		14	3
125722	US Territorial 1880AD to 1899AD	0	0									1	1
126113	Early Pueblo to Late Pueblo 1100AD to 1500AD											1	1
126114	Unspecific 9500 BC to 1880AD	0	1 to 9		3								2
126115	Unspecific 9500 BC to 1880AD	0	0		1								2
126160	Unspecific 9500 BC to 1880AD	0	10 to 99										1
126161	Unspecific 1300 AD to 1600AD	0	0		1								1
126257	Unspecific 9500 BC to 9999 AD	1000 to 9999	1000 to 9999		7							2	3
126562	Unknown 9500 BC to 9999 AD	0	10 to 99										1
126563	Unknown 25000 BC to 9999 AD	Unknown	Unknown										1
126587	Unspecific 9500 BC to 1880AD	0	1 to 9	6									2
126588	Unknown 25000 BC to 9999 AD	0	0	2	2								1
126611	Unspecific 9500 BC to 1880AD	0	10 to 99		8								1
126612	Unspecific 9500 BC to 1880AD	0	1 to 9	1									2
126613	Unspecific 9500 BC to 1880AD	0	10 to 99										1
126614	Late Archaic 200BC to 600AD	100 to 999	100 to 999		1								1
126637	Unspecific 9500 BC to 1880AD	0	100 to 999		3								1
126639	Middle Archaic to Late Archaic 3000BC to 200AD	0	10 to 99										3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
127013	Late Pithouse to Late Pueblo 900AD to 1400AD	Unknown	Unknown										
127014	Unspecific 9500 BC to 1550AD, A.D. 900 to A.D. 1400.	Unknown	Unknown				1		1				3
127015	Late Pithouse to Late Pueblo 900AD to 1400AD	0	0		1								2
127182	Unspecific 9500 BC to 1550AD	Unknown	Unknown	1									2
127183	Unspecific 9500 BC to 1550AD	0	1 to 9		1								1
127785	Middle Archaic 3000BC to 1800BC	0	100 to 999	10	13								1
128124	Unspecific 9500 BC to 1880AD	0	1 to 9		1								1
128356	Late Pithouse to Late Pueblo 750AD to 1400AD	0	10 to 99	1	1								1
128358	Unspecific 9500 BC to 1880AD	0	1 to 9		1								1
128361	Unspecific 9500 BC to 1880AD	0	100 to 999		1								1
128438	Unspecific 9500 BC to 1850AD, Historic	10 to 99	100 to 999		2								3
128491	Unspecific 9500 BC to 1550AD	0	Unknown				2		1				1
128496	Unknown 9500 BC to 9999 AD	0	10 to 99										
128502	Unknown 9500 BC to 9999 AD	0	100 to 999		2								2
128656	Unspecific 9500 BC to 1850AD	0	1000 to 9999		106							1	2
128683	Unspecific 9500 BC to 1880AD, Historic	10 to 99	100 to 999	1	1								2
128825	Unspecific 9500 BC to 1550AD	Unknown	Unknown		1		1						1
129361	Unspecific 9500 BC to 1550AD	0	Unknown				1						2
129362	Late Archaic 0AD to 900AD	0	10000 to 99999							1			1
129363	Unspecific 9500 BC to 1550AD	Unknown	Unknown			1	1						1
129364	Late Archaic 1800BC to 200AD, Early Pithouse to Late Pueblo 200AD to 1200AD	Unknown	Unknown			1	1						1
129365	Unspecific 9500 BC to 1550AD	0	Unknown				1						2
129366	Unspecific 9500 BC to 1550AD	Unknown	Unknown		1		1				1		1
129367	Unspecific 9500 BC to 1550AD	1 to 9	10 to 99		3		1		1				3
129396	Early Pueblo to Late Pueblo AD 1100-1400											4	1
129397												3	3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
129398	Historic	0	10 to 99	3	1							1	2
129405	Unspecific 9500 BC to 1840AD	0	0		1								1
129440	Unspecific 9500 BC to 1840AD	0	0		1								1
129441	Unspecific 9500 BC to 1840AD	0	1 to 9		5								1
129464	Historic	0	1 to 9		1								1
129465	Unspecific 9500 BC to 1850AD	1 to 9	100 to 999	5	52		1				4		2
129466	Late Paleoindian 8000BC to 6600BC	10 to 99	10 to 99	2	4							1	4
129478	Unspecific 9500 BC to 1880AD	0	10 to 99										1
129583	Unspecific 9500 BC to 1880AD	0	1 to 9		1								2
129584	Unspecific 9500 BC to 1880AD	0	1000 to 9999										1
129599	Late Pithouse to Late Pueblo 900AD to 1350AD, Historic	0	10 to 99	1	1								4
129788	Unspecific 9500 BC to 1850AD	0	10 to 99	2	3								2
130082	Unspecific 9500 BC to 1840AD	0	10 to 99										1
130083	Unspecific 9500 BC to 1840AD	0	100 to 999		13								4
130133	Late Pithouse 850AD to 1000AD	Unknown	Unknown		7								1
130402	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
130415	Unspecific 9500 BC to 1840AD	0	10 to 99	1									4
130416	Unspecific 9500 BC to 1840AD	Unknown	Unknown		1								2
130417	Late Archaic 1000BC to 500AD, Late Pithouse to Late Pueblo 750AD to 1400AD	100 to 999	100 to 999										1
130460	Late Archaic 1800 BC to 200AD	1 to 9	1000 to 9999		37								3
130469	Historic	0	10 to 99	1									1
130485	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	1000 to 9999		1								1
130591	Late Pithouse 900AD to 1100AD	1 to 9	10 to 99	3									1
130684	Late Pithouse to Late Pueblo 900AD to 1350AD	0	10 to 99										4
130685	Historic	0	10 to 99										2
130686	Unspecific 9500 BC to 1880AD	0	10 to 99										1
130687	Unspecific 9500 BC to 1880AD	0	10 to 99										2
130688	Historic	0	10 to 99										1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
130689	Early Archaic 5500 BC to 3000 BC											1	2
130690	Unspecific 9500 BC to 1880AD	0	10 to 99		2								1
130691	Late Pueblo 1275AD to 1400AD	0	10 to 99				1					1	1
130692	Unspecific 9500 BC to 1880AD	0	1000 to 9999		6								1
130693	Late Archaic 1800BC to 700AD	10 to 99	1000 to 9999		2							1	2
130694	Unspecific 9500 BC to 1880AD	0	100 to 999		5								1
130696	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
130733	Historic	0	10 to 99										
130734	Unspecific 9500 BC to 1880AD	0	10 to 99	1									1
130744	Historic	0	1000 to 9999		22								1
130745												1	1
130746	Unspecific 9500 BC to 1840AD	0	10 to 99										1
130747	Late Archaic 1000BC to 200AD	1 to 9	10 to 99	1	1								2
130748	Unspecific 9500 BC to 1840AD	0	100 to 999		8								1
130803	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
130804	Middle Archaic to Late Archaic 2500BC to 1000BC	0	1 to 9	1									1
130805	Unspecific 9500 BC to 1880AD	0	10 to 99		2								1
130806	Unspecific 9500 BC to 1880AD	0	1 to 9		2							1	1
130807	Unspecific 9500 BC to 1880AD	0	10 to 99	1	3								1
130808	Unspecific 9500 BC to 1880AD	0	10 to 99									2	2
130846	Unspecific 9500 BC to 1860AD	0	100 to 999	1 1									2
130848	Historic	0	100 to 999		1								1
130849	Unspecific 9500 BC to 1880AD	0	10 to 99	1	7								3
130850	Unspecific 9500 BC to 1880AD	0	1 to 9		1								1
130851	Unspecific 9500 BC to 1880AD	0	10 to 99	3								3	2
130852	Unspecific 9500 BC to 1880AD	0	1 to 9		2								4
130853	Unspecific 9500 BC to 1880AD	0	10 to 99		4								2
130854	Unspecific 9500 BC to 1880AD	0	10 to 99		6								2

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
130855	Historic	0	Unknown		5								1
130856	Unspecific 9500 BC to 1880AD, Historic	Unknown	Unknown		1		1				1		1
130857	Late Archaic 1800BC to 200AD	1 to 9	100 to 999										2
130858	Unspecific 9500 BC to 1950AD	0	0				1						2
130859	Unspecific 9500 BC to 1850AD	Unknown	Unknown		10							1	5
130949	Late Pithouse to Late Pueblo 900AD to 1350AD	0	100 to 999	15	4								1
130972	Late Archaic 1800BC to 200AD	1 to 9	10000 to 99999		24						49	5	5
131110	Unspecific Archaic 5500 BC to 200AD	0	10 to 99										1
131111	Unspecific Jornada 200AD to 1400AD	0	10 to 99		1								1
131112	Historic	Unknown	1 to 9		4								3
131215	Unspecific 9500 BC to 1850AD	0	1 to 9		4							4	2
131222	Unspecific 9500 BC to 1880AD	0	1 to 9		2								2
131263	Plains Woodland to Panhandle Aspect 250AD to 1400 AD	0	10 to 99		3								1
131359	Historic		10 to 99		7							1	4
131361	Early Pueblo to Late Pueblo AD 1100-1400											3	1
131362	Unspecific 9500 BC to 1880AD	0	10 to 99										2
131366	Unspecific 9500 BC to 1840AD	0	1 to 9		2								1
131367	Unspecific 9500 BC to 1840AD	0	10 to 99		1								1
131686	Middle Archaic to Late Archaic 3000BC to 200AD	0	100 to 999										2
131687	Unspecific 9500 BC to 1880AD	0	10 to 99	3									1
131763	Unspecific 9500 BC to 1840AD	0	10 to 99	1									1
131764	Unspecific 9500 BC to 1840AD	0	1 to 9		4								1
131765	Unspecific 9500 BC to 1880AD	0	100 to 999	1	5								3
131766	Unspecific 9500 BC to 1840AD	0	100 to 999		1								1
132037	Unspecific 9500 BC to 1880AD	0	1 to 9		2								1
132127	Unspecific 9500 BC to 2001AD	Unknown	100 to 999										1
132233	Early Pueblo to Late Pueblo 1100AD to 1400AD												1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
132271	Late Pithouse to Late Pueblo 900AD to 1400AD	0	0		1								1
132354	Late Archaic 800BC to 600AD	Unknown	Unknown		1	7							2
132516	Historic	0	10 to 99										1
132517	Unspecific 9500 BC to 1880AD	0	100 to 999	2	1								4
132607	Historic	0	10 to 99		2								1
132608	Unspecific 9500 BC to 1880AD	0	10 to 99										1
132609	Unspecific 9500 BC to 1880AD	0	10 to 99		10								3
132610	Unspecific 9500 BC to 1880AD	0	100 to 999										1
132800	Unspecific 9500 BC to 1840AD	0	10 to 99	1	1								2
132801	Unspecific 9500 BC to 1840AD	0	10 to 99	1									1
132802	Unspecific 9500 BC to 1840AD	0	10 to 99		7								2
132803	Unspecific 9500 BC to 1840AD	0	10 to 99										1
133778	Unspecific 9500 BC to 1840AD	0	10 to 99										1
133905	Unspecific 9500 BC to 1850AD	0	10 to 99									7	4
133967	Historic	0	10 to 99		3								2
133968	Late Pueblo 1175AD to 1500AD	0	10 to 99										1
134105	Unspecific 9500 BC to 1880AD	0	100 to 999										4
134106	Unspecific 9500 BC to 1880AD	0	100 to 999										2
134107	Unspecific 9500 BC to 1880AD	0	1 to 9	2	1								1
134216	Unspecific 9500 BC to 1880AD	0	0		2								1
134274	Unspecific 9500 BC to 1840AD	0	10 to 99		1								1
134275	Unspecific 9500 BC to 1840AD	0	10 to 99		1								1
134276	Unspecific 9500 BC to 1840AD	0	100 to 999		1								2
134277	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
134451	Unspecific 9500 BC to 1840AD	0	1000 to 9999							1			1
134510	Historic	0	10 to 99		11								1
134638	Middle Archaic to Late Archaic 3000BC to 200AD	Unknown	Unknown	1									1
134690	Early Pueblo 1100AD to 1200AD											1	1
134712	Unspecific 9500 BC to 1550AD	0	1 to 9				1				3		1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
134713	Unspecific 9500 BC to 1550AD	0	Unknown		3		1						1
134714	Unspecific 9500 BC to 1550AD	0	10 to 99				1						1
134715	Unspecific 9500 BC to 1550AD	0	Unknown							1			1
134716	Unspecific 9500 BC to 1550AD	0	Unknown	1	1			1					2
134717	Unspecific 9500 BC to 1550AD	0	Unknown	1	1								1
134718	Unspecific 9500 BC to 1550AD	0	0		1								1
134719	Unspecific 9500 BC to 1550AD	0	Unknown		1								3
134810	Late Pithouse to Early Pueblo 750AD to 1200AD	1 to 9	10 to 99		1								1
134920	Unspecific 9500 BC to 1880AD	0	10 to 99	4									2
134930	Unspecific 9500 BC to 1880AD	0	10 to 99	1	2								1
134931	Unspecific 9500 BC to 1880AD	0	10 to 99		8								1
134946	Late Pithouse to Late Pueblo 850AD to 1200AD	0	1000 to 9999		4								1
135222	Unspecific 9500 BC to 1880AD	0	10 to 99										1
135223	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
135279	Unspecific 9500 BC to 1840AD	Unknown	10 to 99										2
135280	Unspecific 9500 BC to 1880AD	0	100 to 999	1	9								1
135281	Unspecific 9500 BC to 1880AD	0	10 to 99	2									1
135282	Unspecific 9500 BC to 1840AD	0	100 to 999		3								1
135283	Late Archaic 1800BC to 200AD	1 to 9	10 to 99										2
135284	Unspecific 9500 BC to 1840AD	0	0		1								2
135286	Unspecific 9500 BC to 1840AD	0	1 to 9		4								1
135287	Unspecific 9500 BC to 1840AD	0	0		1								1
135288	Unspecific 9500 BC to 1840AD, Historic	10 to 99	10 to 99		8								3
135312	Unspecific 9500 BC to 1550AD	0	10 to 99	2	13				1				5
135313	Late Archaic 1800BC to 200AD	Unknown	Unknown		2	1	1		1				1
135320	Unspecific 9500 BC to 1880AD	0	10 to 99	5									1
135416	Unspecific 9500 BC to 1840AD	0	10 to 99										1
135417	Unspecific 9500 BC to 1880AD	0	0									1	1
135418	Late Archaic 1800BC to 200AD	0	100 to 999										3

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
135594	Unspecific 9500 BC to 1880AD	0	10 to 99		4								1
135596	Unspecific 9500 BC to 1880AD	0	10 to 99		2								1
135597	Historic	0	10 to 99		15								1
135600	Unspecific 9500 BC to 1550AD	0	1 to 9		1								1
135790	Unspecific 9500 BC to 1880AD	0	1 to 9		2								1
135813	Unspecific 9500 BC to 1550AD	0	10 to 99		5								2
136257	Unspecific 9500 BC to 1840AD	0	1 to 9		1								2
137409	Late Archaic 1800BC to 900AD	1 to 9	100 to 999		9		1					4	1
137410	Unspecific 9500 BC to 1840AD	0	1 to 9		36							1	4
137426	Unspecific 9500 BC to 1840AD	Unknown	1 to 9		2								1
137427	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
137428	Unspecific 9500 BC to 1840AD	0	100 to 999										1
137661	Unspecific 9500 BC to 1840AD	0	10 to 99								12		3
137662	Unspecific 9500 BC to 1840AD	Unknown	Unknown		1	6	3						1
137698	Unspecific 9500 BC to 1840AD	0	0				1				1	1	2
137778	Unspecific 9500 BC to 1880AD	0	10 to 99		4								1
138753	Late Archaic 1800BC to 200AD	0	10 to 99	1	3								2
138771	Late Archaic 1000BC to 500AD	1 to 9	100 to 999		1								1
138786	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
139072	Unspecific 9500 BC to 1880AD	0	100 to 999		12							4	1
139073	Late Archaic 1800BC to 950AD	Unknown	Unknown				1		1				1
139107	Early Pithouse to Late Pueblo 200AD to 1400AD												2
139186	Early Archaic to Middle Archaic 5500BC to 1800AD											3	3
139891	Unspecific 9500 BC to 1550AD	0	Unknown										1
139892	Late Archaic 1800BC to 200AD	1 to 9	10 to 99	4	11								5
139951	Late Pithouse to Late Pueblo 750AD to 1400AD	0	10 to 99	2	5								4
139973	Unspecific 9500 BC to 1880AD	0	10 to 99		7								1
140785	Historic	0	100 to 999		22								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
140826	Late Pithouse to Late Pueblo (700 AD) to (1500 AD)	Unknown	Unknown				1		1				1
140828	Unspecific 9500 BC to 1880AD	0	1 to 9		1								1
140829	Unspecific 9500 BC to 1880AD	0	10 to 99		1							2	1
140830	Unspecific 9500 BC to 1880AD	0	10 to 99		8								1
140831	Unspecific 9500 BC to 1880AD	0	1 to 9		16								1
140832	Unspecific 9500 BC to 1880AD	0	1 to 9		13								1
140833	Unspecific 9500 BC to 1880AD	0	10 to 99		30							1	2
140834	Unspecific 9500 BC to 1880AD, Historic past NM Statehood	Unknown	Unknown				2						2
140835	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140836	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
140837	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140838	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
140839	Unspecific 9500 BC to 1880AD	0	1 to 9		6							1	1
140840	Unspecific 9500 BC to 1880AD	0	100 to 999	4									1
140841	Unspecific 9500 BC to 1880AD	0	10 to 99	8									1
140842	Unspecific 9500 BC to 1880AD	0	100 to 999										1
140843	Unspecific 9500 BC to 1880AD	0	10 to 99		4								1
140844	Unspecific 9500 BC to 1880AD	0	1 to 9		5								1
140845	Unspecific 9500 BC to 1880AD	0	1 to 9		16								1
140846	Unspecific 9500 BC to 1880AD, Historic	0	100 to 999	4	9							1	2
140847	Unspecific 9500 BC to 1880AD	0	1 to 9		8								1
140848	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
140849	Late Pithouse to Late Pueblo 900AD to 1400AD	0	0		4								1
140850	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1
140851	Unspecific 9500 BC to 1880AD	0	10 to 99		9								1
140852	Historic	0	10 to 99										2
140853	Unspecific 9500 BC to 1880AD	0	1 to 9										1
140856	Unspecific 9500 BC to 1880AD	0	10 to 99										1
140857	Unspecific 9500 BC to 1880AD	0	10 to 99		3								1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
140858	Unspecific 9500 BC to 1880AD	0	10 to 99		3								2
140859	Unspecific 9500 BC to 1880AD	0	10 to 99		2								2
140863	Unspecific 9500 BC to 1880AD	0	10 to 99		1							1	1
140864	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
140865	Unspecific 9500 BC to 1880AD	0	10 to 99		2								1
140866	Late Archaic 1800BC to 200AD	0	10 to 99		1								2
140867	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140868	Unspecific 9500 BC to 1880AD	0	100 to 999		4								1
140869	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140870	Late Archaic 1800BC to 900AD	10 to 99	1000 to 9999	1	19							1	2
140872	Late Archaic 2300BC to 1100AD	Unknown	Unknown		1		1						2
140873	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
140874	Unspecific 9500 BC to 1880AD	0	100 to 999		4								1
140875	Unspecific 9500 BC to 1880AD	0	10 to 99		5							1	1
140876	Late Archaic 1800BC to 900AD	Unknown	Unknown	1			1		1				2
140877	Late Archaic 200AD to 900AD	Unknown	100 to 999				2		1		1		1
140878	Unspecific 9500 BC to 1880AD	0	1 to 9		2								1
140879	Late Archaic 1800BC to 900AD	Unknown	Unknown				1						2
140880	Unspecific 9500 BC to 1880AD	0	100 to 999		23								1
140881	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140882	Late Pithouse to Late Pueblo (900 AD) to (1350 AD)	1 to 9	Unknown				1						2
140883	Unspecific 9500 BC to 1880AD	0	100 to 999		5								1
140884	Unspecific 9500 BC to 1880AD	0	10 to 99		8								1
140885	Late Archaic 1800BC to 900AD, Historic	0	100 to 999										1
140889	Late Archaic 200BC to 900AD	10 to 99	1000 to 9999	7	2								3
140890	Unspecific 9500 BC to 1880AD	0	1 to 9										
140891	Unspecific 9500 BC to 1880AD	0	10000 to 99999		370							4	1
140892	Historic	0	10 to 99										1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
140893	Historic	10 to 99	10 to 99		1	10	1				1		2
140894	Unspecific Archaic 4850 BC to 110AD, Late Pithouse to Late Pueblo 900AD to 1400AD	1 to 9	10000 to 99999		150						10	7	11
140895	Unspecific 9500 BC to 1880AD	0	10 to 99										2
140896	Historic	0	10 to 99		2								1
140897	Unspecific 9500 BC to 1880AD, Historic	1 to 9	10 to 99										1
140898	Unspecific 9500 BC to 1880AD	0	10 to 99		3				1				1
140899	Late Archaic 200AD to 900AD	Unknown	Unknown			1	1		1				1
140900	Unspecific 9500 BC to 1880AD, Historic	100 to 999	1 to 9										1
140901	Unspecific 9500 BC to 1880AD	0	1 to 9	6	2								2
140902	Unspecific 9500 BC to 1880AD	0	10 to 99	13	1								1
140903	Unspecific 9500 BC to 1880AD	0	10 to 99	2									1
140904	Historic	0	100 to 999		2						1	1	1
140905	Unspecific 9500 BC to 1880AD	0	1 to 9	3								1	1
140906	Late Pithouse 900AD to 1100AD	1 to 9	100 to 999		5								2
140919	Unspecific 9500 BC to 1880AD	0	101 to 999		9								1
140920	Unspecific 9500 BC to 1880AD	0	100 to 999		9							2	1
140921	Unspecific 9500 BC to 1880AD	0	1 to 9		2								1
140922	Unspecific 9500 BC to 1880AD	0	100 to 999		12								1
140942	Middle Archaic to Late Archaic 4850BC to 100AD, Late Pithouse to Late Pueblo (900 AD) to (1400 AD)	1 to 9	10 to 99										3
140943	Unspecific 9500 BC to 1880AD	0	10 to 99		1								1
140944	Unspecific 9500 BC to 1880AD	0	10 to 99		5								1
140945	Unspecific 9500 BC to 1880AD	0	0		1								1
140946	Unspecific 9500 BC to 1880AD	0	100 to 999		4								1
141421	Unspecific 9500 BC to 1880AD	0	0		1								1
141685	Early Pithouse to Early Pithouse 200AD to 950AD	0	0									1	1
141858	Historic	0	10 to 99										1

LA	Time Period	Ceramics	Lithics	H	RM	B	C	S	M	Q	R	O	V
141859	Early Pueblo 950AD to 1200AD			2									1
142044	Historic	0	100 to 999	2	19						1		1
142340	Historic	0	100 to 999		58				1				4
142341	Historic	0	1000 to 9999		15							1	1
142530	Late Pithouse to Late Pueblo 750AD to 1400AD	1 to 9	100 to 999										3
142531	Unspecific 9500 BC to 1880AD	0	0		1								2
142532	Unspecific 9500 BC to 1880AD	0	0									1	1
142533	Late Pithouse to Early Pueblo (750 AD) to (1200 AD)	1 to 9	100 to 999		2								1
142534	Unspecific 9500 BC to 1880AD	Unknown	Unknown									1	1
142814	Unspecific 9500 BC to 1880AD	Unknown	Unknown	1									1
143256	Unspecific 9500 BC to 1950AD	0	0				1						1
146499	Historic	Unknown	100 to 999			1 2				2	3	1	3
148614	Unspecific 9500 BC to 1850AD	0	100 to 999	2	2								2
148619	Historic	Unknown	Unknown	5	1							1	5
148988													
153611	Historic	0	10 to 99		1								1
153612	Historic	1 to 9	100 to 999		21								7
153613	Historic	Unknown	1000 to 9999		50								3
155665	Unspecific 9500 BC to 1850AD	0	10 to 99										3
159757	reserved	0	0		8								1
162057	Unspecific 9500 BC to 9999AD	0	10 to 99										1

Appendix C- NetLogo code from project models.

Water Model Code:

```
extensions [ gis ]
globals [ elevation slope aspect soil mi ma]
patches-own [ slope-self aspect-self sg CN Q hr-1 day-1 day-2 day-3 day-5 day-7 hiI lowI storage
stored peak-flow]
turtles-own [water]
to setup
    clear-all
    gis:load-coordinate-system ("sa.prj")
    set elevation gis:load-dataset"sa.asc"
    set soil gis:load-dataset "sasoil.asc"
    ;;gis:set-world-envelope gis:envelope-of elevation
    gis:set-world-envelope (gis:envelope-union-of (gis:envelope-of elevation)(gis:envelope-of
soil))
     ;;let horizontal-gradient gis:convolve elevation 3 3 [ 1 1 1 0 0 0 -1 -1 -1 ] 1 1
     ;;let vertical-gradient gis:convolve elevation 3 3 [ 1 0 -1 1 0 -1 1 0 -1 ] 1 1
    let horizontal-gradient gis:convolve elevation 3 3 [ 1 0 -1 2 0 -2 1 0 -1 ] 1 1
     let vertical-gradient gis:convolve elevation 3 3 [ 1 2 1 0 0 0 -1 -2 -1 ] 1 1
    set slope gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
     set aspect gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
     let x 0
    repeat (gis:width-of slope)
    [ let y 0
     repeat (gis:height-of slope)
    [ let gx gis:raster-value horizontal-gradient x y
    let gy gis:raster-value vertical-gradient x y
    if ((gx <= 0) or (gx >= 0)) and ((gy <= 0) or (gy >= 0))
    [ let s sqrt ((gx * gx) + (gy * gy))
    gis:set-raster-value slope x y s
    ifelse (gx != 0) or (gy != 0)
        [ gis:set-raster-value aspect x y atan gy gx ]
```

```
[ gis:set-raster-value aspect x y 0 ] ]
set y y + 1 ]
set x x + 1 ]
gis:set-sampling-method aspect
"NEAREST_NEIGHBOR"
set mi gis:minimum-of elevation
set ma gis:maximum-of elevation
ask patches [
 set pcolor scale-color green (gis:raster-sample elevation self) mi ma
 set slope-self gis:raster-sample slope self
set aspect-self gis:raster-sample aspect self
set sg gis:raster-sample soil self]
reset-ticks
end
to reset
reset-ticks
clear-turtles
end
to test
 ask patches
 [NG-run-off]
end
to go
rain
loop[
tick
ask turtles [
if xcor >= (max-pxcor - 1)
[die ]
if xcor <= (min-pxcor + 1)
[ die ]
if ycor >= (max-pycor - 1)
[ die ]
if ycor <= (min-pycor + 1)
```

```
[ die ]
if dd = "elevation"
[let target min-one-of neighbors [ gis:raster-sample elevation self + (sum [water] of turtles-
here)]
face target
f (gis:raster-sample elevation target + ((sum [water] of turtles-here) / 39.3701))
< (gis:raster-sample elevation self + ((sum [water] of turtles-here) / 39.3701))
[ let r ((water / 39.3701) * 55.56) / (55.56 + (2 * (water / 39.3701))) ^ (2 / 3)
let s ((([slope-self] of patch-here + .00001) / 100) ^ (1 / 2))
forward (((r * s / .035) * 300) / 55.56) ]]  ;; manning formula = radius ^ 2/3 * slope ^ 1/2
divdied by Manning's roughness coefficient
if dd = "aspect"
[set heading [aspect-self] of patch-here
if (gis:raster-sample elevation patch-ahead 1 + ((sum [water] of turtles-here) / 39.3701))
 < (gis:raster-sample elevation self + ((sum [water] of turtles-here) / 39.3701))
[ let r ((water / 39.3701) * 55.56) / (55.56 + (2 * (water / 39.3701))) ^ (2 / 3)
let s ((([slope-self] of patch-here + .00001) / 100) ^ (1 / 2))
forward (((r * s / .035) * 300) / 55.56) ] ]
if dd = "aspect-1"
[set heading [aspect-self] of patch-here
let r ((water / 39.3701) * 55.56) / (55.56 + (2 * (water / 39.3701))) ^ (2 / 3)
 let s ((([slope-self] of patch-here + .00001) / 100) ^ (1 / 2))
forward (((r * s / .035) * 300) / 55.56) ]
 set water (water - ( EV * .7 ))
if ticks > 12
[if [storage] of patch-here - [stored] of patch-here > 0
[set water (water - ([lowI] of patch-here) / 12)
 ask patch-here[set stored (stored + (lowI / 12) )] ] ]
 if [peak-flow] of patch-here < sum [water] of turtles-here
[set peak-flow sum [water] of turtles-here ]
if water < 0
[ die ]
 set color 99 - round(sum [water] of turtles-here)
if count turtles-here >1
```

```
            [ask patch-here
            [set Q (sum [water] of turtles-here)
            ask turtles-here [die]
            sprout 1
            [set color 99
            set shape "circle"
            set water Q ]
            ]  ]
            ]
            set-days
            if not any? Turtles
            [ stop ]]
end
to no-NG-runoff
            ask patches
            [set Q down-pour]
            end
            to rain
            ask patches
            [if [Q] of self > .001
            [sprout 1
            [set color 99
             set shape "circle"
             set water Q ]
             ] ]
end
to set-days
            if ticks = 12
            [ask patches[set hr-1 (sum [water] of turtles-here)] ]
            if ticks = 288
            [ask patches[set day-1 (sum [water] of turtles-here)] ]
            if ticks = 576
            [ask patches[set day-2 (sum [water] of turtles-here)] ]
            if ticks = 864
```

```
        [ask patches[set day-3 (sum [water] of turtles-here)] ]
        if ticks = 1440
        [ask patches[set day-5 (sum [water] of turtles-here)] ]
        if ticks = 2016
        [ask patches[set day-7 (sum [water] of turtles-here)] ]
end
to show-days
        ask patches [set pcolor scale-color green (gis:raster-sample elevation self) mi ma]
        if Days = "1 Hr"
        [ask patches [ if hr-1 > 1 [set pcolor scale-color red hr-1 0 10  ]]]
        if Days = "Day 1"
        [ask patches [ if day-1 > 1 [set pcolor scale-color red day-1 0 10  ]]]
        if Days = "Day 2"
        [ask patches [ if day-2 > 1 [set pcolor scale-color red day-2 0 10  ]]]
        if Days = "Day 3"
        [ask patches [ if day-3 > 1 [set pcolor scale-color red day-3 0 10  ]]]
        if Days = "Day 5"
        [ask patches [ if day-5 > 1 [set pcolor scale-color red day-5 0 10  ]]]
        if Days = "Day 7"
        [ask patches [ if day-7 > 1 [set pcolor scale-color red day-7 0 10  ]]]
end
to find-dips
        ask patches[
         if [gis:raster-sample elevation self] of min-one-of neighbors [ gis:raster-sample elevation self
] > gis:raster-sample elevation self
        [set pcolor red]
        ]
end
to show-peak-flow
        ask patches [set pcolor scale-color blue peak-flow 0 20]
end
to NG-run-off
        let S 0
        if [sg] of self = 1 and A-veg = "poor"
```

```
[set CN 63]
if [sg] of self = 1 and A-veg = "fair"
[set CN 55]
if [sg] of self = 1 and A-veg = "good"
[set CN 49]
if [sg] of self = 2 and B-veg = "poor"
[set CN 77]
if [sg] of self = 2 and B-veg = "fair"
[set CN 72]
if [sg] of self = 2 and B-veg = "good"
[set CN 68]
if [sg] of self = 3 and C-veg = "poor"
[set CN 85]
if [sg] of self = 3 and C-veg = "fair"
[set CN 82]
if [sg] of self = 3 and C-veg = "good"
[set CN 79]
if [sg] of self = 4 and D-veg = "poor"
[set CN 88]
if [sg] of self = 4 and D-veg = "fair"
[set CN 86]
if [sg] of self = 4 and D-veg = "good"
[set CN 84]
if [sg] of self = 5
[set CN 100]
if AMC = "AMCI"
[ let AI ((4.2 * CN) / (10 - (0.58 * CN)))
  set CN AI]
if AMC = "AMCIII"
[ let AIII ((23 * CN) / (10 - (0.13 * CN)))
set CN AIII]
if CN > 0
[set S ((1000 / CN) - 10)
 set Q ((down-pour - (.2 * S)) ^ 2 / (down-pour + (.8 * S)))]
```

```
set stored (down-pour - Q)
;;if Q > 0.1
;;[set pcolor green]
end
to set-sg
ask patches
[ if sg >= 8 and sg < 41
[set sg 1 set lowI 0 set hiI 0 set storage 0]
if sg >= 41 and sg < 72
[set sg 5 set lowI 0 set hiI 0 set storage 0]
if sg >= 72 and sg < 74
[set sg 2 set lowI .2 set hiI .6 set storage 9.6]
if sg >= 74 and sg < 76
[set sg 4 set lowI 0 set hiI .06 set storage 1.1]
if sg >= 76 and sg < 84
[set sg 4 set lowI 0 set hiI .06 set storage 2.6]
if sg >= 84 and sg < 88
[set sg 4 set lowI 0 set hiI .06 set storage 1]
if sg >= 88 and sg < 90
[set sg 4 set lowI 0 set hiI .06 set storage .8]
if sg >= 90 and sg < 101
[set sg 4 set lowI 0 set hiI .06 set storage .7]
if sg >= 101 and sg < 102
[set sg 3 set lowI .01 set hiI .06 set storage 1.9]
if sg >= 103 and sg < 104
[set sg 3 set lowI 0 set hiI .06 set storage 6.4]
if sg >= 104 and sg < 105
[set sg 3 set lowI .01 set hiI .06 set storage 1.3]
if sg >= 105 and sg < 106
[set sg 3 set lowI .01 set hiI .06 set storage 1.4]
if sg >= 106 and sg < 107
[set sg 2 set lowI .2 set hiI 6 set storage 8.2]
if sg >= 107 and sg < 108
[set sg 2 set lowI .2 set hiI .06 set storage 11.9]
```

```
if sg >= 108 and sg < 111
[set sg 3 set lowI 0 set hiI .06 set storage 6.4]
if sg >= 111 and sg < 113
[set sg 2 set lowI .6 set hiI 2 set storage 8.2]
if sg >= 113 and sg < 115
 [set sg 2 set lowI .2 set hiI .6 set storage 10.5]
if sg >= 115 and sg < 117
[set sg 4 set lowI .06 set hiI 2 set storage .5]
if sg >= 117 and sg < 119
[set sg 2 set lowI .57 set hiI 1.98 set storage 9]
if sg >= 119 and sg < 120
[set sg 2 set lowI .06 set hiI 2 set storage 8.2]
if sg >= 120 and sg < 121
[set sg 3 set lowI .01 set hiI .6 set storage 1.4]
if sg >= 121 and sg < 122
[set sg 2 set lowI 2 set hiI 6 set storage 7.2]
if sg >= 122 and sg < 123
[set sg 3 set lowI .01 set hiI .6 set storage 1.4]
if sg >= 123 and sg < 124
[set sg 2 set lowI .06 set hiI 2 set storage 8.2]
if sg >= 124 and sg < 125
[set sg 4 set lowI 0 set hiI .06 set storage 2.1]
if sg >= 125 and sg < 126
[set sg 4 set lowI 0 set hiI .06 set storage 1.3]
if sg >= 126 and sg < 130
[set sg 2 set lowI 2 set hiI 6 set storage 8.4]
if sg >= 130 and sg < 131
[set sg 4 set lowI .06 set hiI 2 set storage .4]
if sg >= 131 and sg < 134
[set sg 2 set lowI .06 set hiI 2 set storage 9.6]
if sg >= 134 and sg < 137
[set sg 2 set lowI .2 set hiI .6 set storage 11.9]
if sg >= 137 and sg < 140
[set sg 4 set lowI 0 set hiI .06 set storage 6]
```

```
        if sg >= 140 and sg < 142
        [set sg 4 set lowI 0 set hiI .06 set storage 2.1]
        if sg >= 142 and sg < 143
        [set sg 2 set lowI .2 set hiI .06 set storage 10]
        if sg >= 143 and sg < 146
        [set sg 2 set lowI 0 set hiI .06 set storage 4.3]
        if sg >= 146 and sg < 150
        [set sg 3 set lowI .2 set hiI 2 set storage 1.2]
        if sg >= 150 and sg < 151
        [set sg 4 set lowI .06 set hiI 2 set storage .5]
        if sg >= 151 and sg < 152
        [set sg 1 set lowI 2 set hiI 6 set storage 4.3]
        if sg >= 152 and sg < 156
        [set sg 3 set lowI .2 set hiI 2 set storage 1.2]
        if sg >= 156 and sg < 168
        [set sg 2 set lowI 0 set hiI .06 set storage 4.3]
        if sg >= 168 and sg < 173
        [set sg 2 set lowI 2 set hiI 6 set storage 7.2]
        if sg >= 173 and sg < 200
        [set sg 4 set lowI 0 set hiI 0 set storage 0]
        if sg = 1
        [set pcolor scale-color red (gis:raster-sample elevation self) mi ma ]
        if sg = 2
        [set pcolor scale-color blue (gis:raster-sample elevation self) mi ma ]
        if sg = 3
        [set pcolor scale-color black (gis:raster-sample elevation self) mi ma ]
        if sg = 4
        [set pcolor scale-color yellow (gis:raster-sample elevation self) mi ma ]
        if sg = 5
        [set pcolor scale-color green (gis:raster-sample elevation self) mi ma ]
        ]
end
```

Least cost code

```
extensions [ gis ]
globals [ elevation slope aspect mi ma ]
```

```
patches-own [ pcost slope-self aspect-self pmarked fx fy gozone xhome yhome visited]
turtles-own [tcost]
breed [walkers walker]
breed [marks mark]
breed [RMs RM]
to setup
	clear-all
	set elevation gis:load-dataset "sa.asc"
	gis:set-world-envelope gis:envelope-of elevation
	 let horizontal-gradient gis:convolve elevation 3 3 [ 1 1 1 0 0 0 -1 -1 -1 ] 1 1
	let vertical-gradient gis:convolve elevation 3 3 [ 1 0 -1 1 0 -1 1 0 -1 ] 1 1
	set slope gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
	set aspect gis:create-raster gis:width-of elevation gis:height-of elevation gis:envelope-of
elevation
	let x 0
	 repeat (gis:width-of slope)
	[ let y 0
	repeat (gis:height-of slope)
	[ let gx gis:raster-value horizontal-gradient x y
	let gy gis:raster-value vertical-gradient x y
	if ((gx <= 0) or (gx >= 0)) and ((gy <= 0) or (gy >= 0))
	[ let s sqrt ((gx * gx) + (gy * gy))
	gis:set-raster-value slope x y s
	ifelse (gx != 0) or (gy != 0)
	[ gis:set-raster-value aspect x y atan gy gx ]
	[ gis:set-raster-value aspect x y 0 ] ]
	set y y + 1 ]
	set x x + 1 ]
	gis:set-sampling-method aspect "bilinear"
	set mi gis:minimum-of elevation
	set ma gis:maximum-of elevation
	 ask patches [
	 set pcolor scale-color green (gis:raster-sample elevation self) mi ma
```

```
    set slope-self gis:raster-sample slope self
    set aspect-self gis:raster-sample aspect self
    set pcost 10000
    ]
    reset-ticks
end
to mark-out
    random-seed 584965489
    let exy 10
    let wxy 10
    repeat 86 [
    set wxy 10
    ask patch max-pxcor exy
    [repeat 86 [
     sprout-walkers 1
    [set color red
    set pcost 0]
    ask walkers
    [ while [pxcor > min-pxcor]
    [facexy min-pxcor wxy
    fd 1
    ask patch-here
    [;;set pcolor blue
     if patch-at-heading-and-distance 0 100 != nobody
     [ask patch-at-heading-and-distance 0 100
    [set gozone 1
    ;;set pcolor red
    ]]
    if patch-at-heading-and-distance 180 100 != nobody
    [ask patch-at-heading-and-distance 180 100
    [set gozone 1
     ;;set pcolor yellow
    ]]]]
    die]
```

```
ask patches with [gozone = 1]
[ ask neighbors
[set gozone 1] ]
set pcost 0
set pmarked 1
replicate
while [[pmarked] of patch min-pxcor wxy < 2 ]
[ask marks
[ask patch-here
[replicate]
 die
]
]
ask patches with [gozone = 1]
[set gozone 0]
ask patch min-pxcor wxy
[ sprout-RMS 1
[set color red]]
ask RMS
[while [[pcost] of patch-here != 0]
[
 setxy [xhome] of patch-here  [yhome] of patch-here
 ask patch-here
[set pcolor red
set visited visited + 1] ]
die]
ask patches with [pmarked > 0]
[set pcost 10000
set pmarked 0]
set wxy wxy + 5
]
]
set exy exy + 5]
stop
```

```
end
to east
    random-seed 584965489
    let exy 20
    let wxy 20
    repeat 40 [
    set wxy 20
    ask patch min-pxcor exy
    [repeat 40 [
    sprout-walkers 1
    [set color red
    set pcost 0]
    ask
    walkers
    [ while [pxcor < max-pxcor]
    [facexy max-pxcor wxy
    fd 1
    ask patch-here
    [   ;;set pcolor blue
    if patch-at-heading-and-distance 0 100 != nobody
    [ask patch-at-heading-and-distance 0 100
    [set gozone 1
    ;;set pcolor red
    ]]
    if patch-at-heading-and-distance 180 100 != nobody
    [ask patch-at-heading-and-distance 180 100
    [set gozone 1
    ;;set pcolor yellow
    ]]]]
    die]
    ask patches with [gozone = 1]
    [ ask neighbors
    [set gozone 1] ]
    set pcost 0
```

```
        set pmarked 1
        replicate
        while [[pmarked] of patch max-pxcor wxy < 2 ]
        [ask marks
        [ask patch-here
        [replicate]
         die
         ]
        ]
        ask patches with [gozone = 1]
        [set gozone 0]
        ask patch max-pxcor wxy
        [ sprout-RMS 1
        [set color red]]
        ask RMS
        [while [[pcost] of patch-here != 0]
        [
        setxy [xhome] of patch-here  [yhome] of patch-here
        ask patch-here
        [set pcolor blue
        set visited visited + 1] ]
        die]
        ask patches with [pmarked > 0]
        [set pcost 10000
        set pmarked 0]
        set wxy wxy + 10
        ]
        ]set exy exy + 10]
        stop
end
to replicate
        let pheading 0
        let sheading 0
        let pc 0
```

```
let sc 0
let hy [pycor] of self
let hx [pxcor] of self
let scost [pcost] of self
let angle 0
let p patch-at-heading-and-distance angle 1
repeat 4 [
set p patch-at-heading-and-distance angle 1
if p != nobody and gozone = 0 and [pmarked] of p != 2[
set pheading (subtract-headings [aspect-self] of p angle)
ifelse pheading >= 90
[set pheading pheading - 90]
[set pheading (abs (pheading - 90))]
set pc (6 * (exp (-3.5 * ( abs ( (pheading / 90) * ((gis:raster-sample slope p / 100) + .05)))  )))
set sheading (subtract-headings aspect-self angle)
ifelse sheading >= 90
[set sheading sheading - 90]
[set sheading (abs (sheading - 90))]
set sc (6 * (exp (-3.5 * ( abs ( (sheading / 90) * ((gis:raster-sample slope self / 100) + .05)))  )))
if [pcost] of p > (scost + (.5 / pc) + (.5 / sc))
[ask p
[set pcost (scost + (.5 / pc) + (.5 / sc))
set xhome hx
set yhome hy
set pmarked 1
if count turtles-here = 0
[sprout-marks 1
[set color red]] ]]] set angle angle + 90 ]
set angle 45
repeat 4 [
set p patch-at-heading-and-distance angle 1
if p != nobody and gozone = 0 and [pmarked] of p != 2[
set pheading (subtract-headings [aspect-self] of p angle)
ifelse pheading >= 90
```

```
        [set pheading pheading - 90]
        [set pheading (abs (pheading - 90))]
        set pc (6 * (exp (-3.5 * ( abs ( (pheading / 90) * ((gis:raster-sample slope p / 100) + .05)))  )))
        set sheading (subtract-headings aspect-self angle)
        ifelse sheading >= 90
        [set sheading sheading - 90]
        [set sheading (abs (sheading - 90))]
        set sc (6 * (exp (-3.5 * ( abs ( (sheading / 90) * ((gis:raster-sample slope self / 100) + .05)))  )))
        if [pcost] of p > (scost + (sqrt .5 / pc) + (sqrt .5 / sc))
        [ask p
        [set pcost (scost + (sqrt .5 / pc) + (sqrt .5 / sc))
        set xhome hx
        set yhome hy
        set pmarked 1
        if count turtles-here = 0
        [sprout-marks 1
        [set color red]]  ]]] set angle angle + 90 ]
        set pmarked 2
end
to f
        ask patches [set pcolor scale-color 0 (gis:raster-sample elevation self) mi ma]
        ask patches with [visited > 100]
        [set pcolor scale-color red visited 1 1000 ]
end
to re-set
        clear-turtles
        clear-patches
        ask patches [set pcolor scale-color green (gis:raster-sample elevation self) mi ma
        set slope-self gis:raster-sample slope self
        set aspect-self gis:raster-sample aspect self
        set gozone 0
        set pcost 10000]
        reset-ticks
end
```

```
to c-cost
	clear-turtles
	ask patches
	[set pcolor scale-color green ([pcost] of self) 0 10000]
	end
	to pause
	stop
end
```

References

Adriaensen, F., J. P. Chardon, G. De Blust, H. Gulinck, E. Matthysen, E. Swinnen and S. Villalba (2003). "The application of 'least-cost' modelling as a functional landscape model." Landscape and Urban Planning **64**: 233-247.
Aitchison, K., E. Alphas, V. Ameels, M. Bentz, C. Borş, E. Cella, K. Cleary, C. Costa, P. Damian, M. Diniz, C. Duarte, J. Frolík, C. Grilo, N. Kangert, R. Karl, A. K. Andersen, V. Kobrusepp, T. Kompare, E. Krekovič, M. L. da Silva, A. Lawler, I. Lazar, K. Liibert, A. Lima, G. MacGregor, N. McCullagh, M. Mácalová, A. Mäesalu, M. Malińska, A. Marciniak, M. Mintaurs, K. Möller, U. Odgaard, E. Parga-Dans, D. Pavlov, V. P. Kocuvan, D. Rocks-Macqueen, J. Rostock, J. P. Tereso, A. Pintucci, E. S. Prokopiou, J. Raposo, K. Scharinghausen, T. Schenck, M. Schlaman, J. Skaarup, A. Šnē, D. Staššíková-Štukovská, I. Ulst, M. van den Dries, H. van Londen, R. Varela-Pousa, C. Viegas, A. Vijups, N. Vossen, T. Wachter and L. Wachowicz (2014). Discovering the Archaeologists of Europe 2012–14: Transnational Report.
Aldenderfer, M. and H. D. G. Maschner, Eds. (1996). Anthropology, Space and Geographic Information Systems. New York, Oxford University Press.
Allen, K. M. S. (2000). Considering of Scale in Modeling Settlement Patterns Using GIS: An Iroquois Example. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. K. L. W. a. R. J. Brandon. New York, Taylor & Francis 101-110.
Allen, K. M. S., S. W. Green and E. B. W. Zubrow, Eds. (1990). Interpreting space: GIS and archaeology. London, Taylor and Francis
Allen, R. J. (2010). Survey of Agent Based Modelling and Simulation Tools. Warrington, Science and Technology Facilities Council.
Altschul, J., W. E. Hayden, C. M. Rohe, L. Sebastian and S. Thompson (2005). Adaptive Management and Planning Models for Cultural Resources in Oil and Gas Fields in New Mexico, Pump III Project. Carson City, Nevada, Gnomon, Inc.
Altschul, J. H. (1988). Models and Modeling Process. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 61-96.
Altschul, J. H. (1990). Red flag models: the use of modelling in management contexts. Interpreting Space: GIS and archaeology. K. M. Allen, Green, S.W., and Zubrow, E.B.W. London, Taylor Francis**:** 226-238.
Altschul, J. H., K. Heidelberg and L. Sebastian (2004). Predictive Modeling in the Military Similar Goals, Divergent Paths. Rio Rancho, New Mexico, SRI Foundation.
Altschul, J. H., T. H. Klein and L. Sebastian (2005). A Workshop on Predictive Modeling & Cultural Resource Management on Military Installations. Rio Rancho, NM, SRI Foundation.
Altschul, J. H. and C. L. Nagel (1988). Collecting New Data For the Purpose of model Development. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 257-299.
Altschul, J. H. and M. R. Rose (1998). An Overview of Statistical method and Theory for Quantitative Model Building. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 173-255.
Anderson, G. W., J. C. Chugg, L. H. Jones and D. L. King (1971). Soil Survey. Eddy Area, New Mexico, United States Department of Agriculture.

Anderson, O. J., Jones.G.E. and G. N. Green (1997). Geologic map of New Mexico: U.S. Geological Survey Open-File Report 97-52.
Anderson, W. (2004). "Lechuguilla (Agave lechuguilla)."   Retrieved 6/6/12, 2012, from http://museum2.utep.edu/chih/gardens/plants/agavlech.htm.
Andrefsky Jr., W. (1994). "Raw-material availability and the organization of technology." American Antiquity **59**: 21-34.
Applegarth, J. S. (1979). Environmental Implications of Herpetofaunal Remains from Archaeological Sites West of Carlsbad, New Mexico. Biological investigations in the Guadalupe Mountains National Park, Texas : proceedings of a symposium held at Texas Tech University, Lubbock, Texas, April 4-5, 1975 H. H. Genoways and R. J. Baker. Washington National Park Service.
Applegarth, S. M. (1976). Prehistoric Utilization of the Environment of the Eastern Slopes of the Guadalupe Mountains, Southeastern New Mexico. PhD, University of Wisconsin.
Artz, J. A. and K. Reid (1983). Modeling the Location of Buried Archaic Sites in Northeastern Oklahoma: A geomorphic Perspective. 48th Annual Meeting of the Society of American Archaeology. Pittsburgh.
Asch, D. L. and R. E. Warren (2000). A Predictive Model of Archaeological Site Location in the Eastern Prairie Peninsula. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. L. W. a. B. Konnie, R. Joe. New York, Taylor & Francis**:** 5-25.
Askenazi, M., R. Burkhart, C. Langton and N. Minar (1996). The Swarm simulation system: A toolkit for building multi-agent simulations. Santa Fe, Santa Fe Institute.
Axelrod, R. (1997). The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. Princeton, NJ, Princeton Univ. Press.
Axtell, R. L. and J. M. Epstein (1996). Growing Artificial Societies: Social Science from the Bottom Up. Cambridge, MA, MIT Press.
Baden, W. W. and C. S. Beekman, Eds. (2005). Nonlinear models for archaeology and anthropology: Continuing the revolution. Aldershot, Ashgate.
Bailey, K., T. Grossardt, P. B. Mink and J. Ripy (2009). Predictive Archaeological Modeling using GIS-Based Fuzzy Set Estimation. ESRI International Users Conference.
Bamforth, D. B. (2006). "The Windy Ridge quartzite quarry: hunter–gatherer mining and hunter–gatherer land use on the North American Continental Divide." World Archaeology **38**: 511-527.
Banning, E. B. (2002). Archaeological Survey New York, Kluwer Academic.
Barber, R. and M. E. Roberts (1979). Archaeology and Paleontology. A Summary and Analysis of Cultural Resource Information on the Continental Shelf from the Bay of Fundy to Cape Hatteras, Final Report. Conservation Archaeological Report 88. Cambridge, Peabody Museum, Harvard University. **II**.
Barlow, K. R. and D. Metcalfe (1992). "A model for exploring the optimal trade-off between field processing and transport." American Anthropologist **94**: 340-356.
Beckman, K. A. and R. B. Duncan (2000). The Application of GIS Site Location Models within Pennsylvania and West Virginia. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. L. W. a. B. Konnie, R. Joe. New York, Taylor & Francis 33-56.
Bell, J. A. (1994). Reconstructing prehistory: Scientific method in archaeology. Philadelphia, Temple University Press.
Bell, T. and G. Lock (2000). Topographic and cultural influences on walking the Ridgeway in later prehistoric times. Beyond the Map. Archaeology and Spatial Technologies. G. Lock. Oxford, IOS Press.

Bell, T., A. Wilson and A. Wickham (2002). "Tracking the Samnites: Landscape and Communications Routes in the Sangro Valley, Italy." American Journal of Archaeology **106**: 169-186.
Bell, W. H., E. Castetter and A. Grove (1938). "The Early Utilization and the Distribution of Agave in the American Southwest." The University of New Mexico Bulletin **5**(4): 1-93.
Bennett, D. A. and J. B. Hageman (2000). Construction of Digital Elevation Models for Archaeological Applications. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. K. L. W. a. R. J. Brandon. New York, Taylor & Francis 113-124.
Bentley, R. A. and H. D. G. Maschner, Eds. (2003). Complex systems and archaeology. Empirical and theoretical applications. Salt Lake, University of Utah.
Berger, J.-F. and P. Verhagen (2007). The hidden reserve. Predictive modelling of buried archaeological sites in the Tricastin-Valdaine region. Case Studies in Archaeological Predictive Modelling. . P. Verhagen, Leiden University Press**:** 41-70.
Berman, M. J., E. L. Camilli and J. I. Ebert (1996). GIS in the Analysis of Disributional Archaeological Data. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press**:** 25-37.
Berry, M. S. (1984). " Sampling and Predictive Modeling on Federal Lands in the West " American Antiquity **49**(4): 842-853.
Bertoncello, F., T. Brughmans, L. Nuninger and P. Verhagen (2012). The Long and Winding Road: Combining Least Cost Paths and Network Analysis Techniques for Settlement Location Analysis and Predictive Modelling. Computer Applications and Quantitative Methods in Archaeology. Southampton.
Bettinger, R. L. (1980). Explanatory/predictive models of hunter-gatherer adaptation. Advances in Archaeological Method and Theory. M. B. Schiffer. New York, Academic Press. **3:** 189-255.
Bettinger, R. L. and D. H. Thomas (1976). "Prehistoric Pinon Ecotone Settlements of the Upper Reese River Valley, Central Nevada." Anthropological Papers of the American Museum of Natural History **53**.
Bevan, A. (2012). Travel and Interaction in the Greek and Roman World. A Review of Some Computational Modelling Approaches. London, UCL.
Bithell, M. and J. Brasington (2009). "Coupling agent-based models of subsistence farming with individual-based forest models and dynamic models of water distribution." Environmental Modelling & Software **24**: 173-190.
Blaney, H. F. and W. D. Criddle (1950). Determining water requirements in irrigated areas from climatological and irrigation data. Washington DC, USDA Soil Conservation Service.
Bohensky, E., J. Butler, G. Carlin, M. Hodgen, A. House, F. Pambudhi and A. Smajgl (2009). Design Document for agent-based model SimPaSI East Kalimantan. Townsville, CSIRO.
Bonabeau, E. (2002). "Agent-based modeling: Methods and techniques for simulating human systems." Proceedings of the National Academy of Science **99**(3): 7280-7287.
Brandon, R. J., G. R. Burgett and T. Church (2000). GIS Applications in Archaeology: Method in Search of Theory. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. K. L. W. a. R. J. Brandon. New York**:** 135-150.
Brandon, R. J. and K. L. Wescott, Eds. (2000). Practical applications of GIS for archaeologists: A predictive modeling kit. London, Taylor and Francis.

Brandt, R. W., B. J. Groenewoudt and K. L. Kvamme (1992). "An experiment in archaeological site location: Modelling in the Netherlands using GIS techniques." World Archaeology (24): 268-282.
Breitenecker, F. and N. Popper (2011). "Simulation. From Theory to Application.".
Brewster, A. and S. Reddy (1999). "Applying GIS to Archaeological Site Prediction on Camp Pendleton, Southern California." Pacific Coast Archaeological Society Quarterly **35**(1): 7-18.
Brooke, M., H. Dawson, K. E. Kolm and P. K. M. Van der Heijde (1997). Determining Water Infiltration Routes from Structures Located Above Carlsbad Cavern, Carlsbad Caverns National Park, Carlsbad, New Mexico. Fort Collins, National park Service.
Brujes, K., W. Ballard, M. Humphrey, F. Harwell, N. McIntyre, P. Krausman and M. Wallace (2006). "Habitat use by sympatric mule and white-tailed deer in Texas." The Journal of Wildlife Management **70**: 1351-1359.
Burgess, R. J., K. L. Kvamme, P. R. Nickens, A. D. Reed and G. C. J. Tucker (1980). Class II Cultural Resource Inventory of the Glenwood Springs resource Area, Grand Junction District, Colorado. Grand Junction, Nickens and Associates. Submitted to the Bureau of Land management.
Burgess, T. L. (1979). Agave- Complex of the Gudalupe Mountains National Park: Putative Hybridization Between Members of Different Subgenera. Biological investigations in the Guadalupe Mountains National Park, Texas : proceedings of a symposium held at Texas Tech University, Lubbock, Texas, April 4-5, 1975. H. H. Genoways and R. J. Baker. Washington, National Park Service.
Burns, I. and T. Whitley (2012). A Paleoeconomic Approach to Predictive Modeling in the Lower Mississippi River Region (Southern Arkansas, Northern Louisiana, and Western Mississippi, USA). CAA 2012. Southampton.
Buskirk, W. (1986). The Western Apache: Living with the Land before 1950. Norman, University of Oklahoma Press.
Cambridge Systematics, I. (2009). Effective Practices for Considering Historic Preservation in Transportation Planning and Early Project Development, American Association of State Highway and Transportation Officials (AASHTO) Standing Committee on the Environment.
Canning, S. (2005). "'BELIEF' in the past: Dempster-Shafer theory, GIS and archaeological predictive modelling." Australian Archaeology **60**: 6-15.
Caraher, W. R., D. Nakassis and D. K. Pettigrew (2006). "Siteless Survey and Intensive Data Collection in an Artifact-rich Environment: Case Studies from the Eastern Corinthia, Greece." Journal of Mediterranean Archaeology **19**(1).
Carmichael, D. L. (1982). Fresnal Shelter, New Mexico: Preliminary Dating and Evidence for Early Cultigens. 47th Annual Meeting of the Society for American Archaeology. Minneapolis, Minnesota.
Carmichael, D. L. (1990). GIS predictive modelling of prehistoric site distribution in central Montana. Interpreting Space: GIS and archaeology. K. M. Allen, Green, S.W., and Zubrow, E.B.W. London, Taylor Francis: 216-225.
Castetter, E. and M. E. Opler (1936). "The Ethnobiology of the Chiricahua and Mescalero Apache." The University of New Mexico Bulletin **4**(6).
Castle, C. and A. Crooks (2006). Principles and Concepts of Agent-Based Modelling for Developing Geospatial Simulations. London, University College London.
Cegielski, W. H. and D. Rogers (2016). "Rethinking the role of Agent-Based Modeling in archaeolog." Journal of Anthropological Archaeology **41**: 283-298.
Center, W. R. C. (2012). "Western Regional Climate Center " Retrieved 6/6/2012, 2012.

Chadwick, A. J. (1978). A computer simulation of Mycenaean settlement. Simulation studies in archaeology. I. Hodder. Cambridge, Cambridge University Press: 47-57.
Chapman, J. A., J. P. Dunn and R. E. Marsh (1982). Jackrabbits: Lepus californicus and allies. Wild mammals of North America: biology, management and economics. J. A. Chapman and G. A. Feldhamer. Baltimore, MD, The Johns Hopkins University Press.
Chisholm, M. (1962). Rural settlement and land use: an essay in location. London, Hutchinson University Library.
Church, R. L. and A. J. Ruggles (1996a). Spatial Allocation in Archaeology: An Opportunity fro Reevaluation. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press: 147-173.
Church, R. L. and A. J. Ruggles (1996b). An Analysis of Late-Horizon Settlement Patterns in the Teotihuacan-Temascalapa Basins: A Location-Allocation and GIS-Based Approach. Anthropology, Space, and Geographic Information Systems. M. Aldenderfer, and Maschner, Herbert D.G. London, Oxford University Press: 155-175.
Church, T. (1996). "Lithic Resources of the Bearlodge Mountains, Wyoming: Description, Distribution, and Implications." The Plains Anthropologist **41**(156): 135-164.
City of Carlsbad, N. M. (2012). "About Carlsbad, New Mexico." Retrieved 6/6/2012, 2012, from http://www.cityofcarlsbadnm.com/AboutCarlsbad.cfm.
Clarke, K. C., A. Ford and G. Raines (2009). "Modeling Settlement Patterns of the Late Classic Maya Civilization with Bayesian Methods and Geographic Information Systems " Annals of the Association of American Geographers **99**(3): 496-520.
Clement, C., R. W. Kloot and D. S. Ohm (2001). Using GIS to Model and Predict Likely Archaeological Sites. ESRI User’s Conference, San Diego, CA.
Cole, M., S. Gould, S. Madry, B. Resnick, S. Seibel and M. Wilkerson (2006). A GIS-Based Archaeological Predictive Model and Decision Support System for the North Carolina Department of Transportation. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York: 317-334.
Cole, M., S. Madry, M. L. Scott and S. Seibel (2006). Archaeological Predictive Modeling: Method and Theory, Environmental Services, Inc.
Collins, M. B. (2002). "The Gault Site, Texas, and Clovis Research." Athena Review **3**(2).
Conolly, J. and M. Lake (2006). Geographical Information Systems in Archaeology. Cambridge, Cambridge University Press.
Cordell, L. S. (1979). Cultural Resources Overview of the Middle Rio Grande Valley, New Mexico. Washington D.C., U.S. Government Printing Office.
Costopoulos, A. and M. Lake, Eds. (2010). Simulating Change: Archaeology Into the Twenty-First Century. Salt Lake City, University of Utah Press.
Cox, E. R. (1967). Geology and Hydrology Between Lake McMillian and Carlsbad Springs Eddy County, New Mexico. Washington DC, US Department of Interior.
Crawford, O. G. S. (1912). "Distribution of Early Bronze Age Settlements in Britain." Geographical Journal **40**: 184.
Crawford, O. G. S. (1921). Man and his Past. London, Oxford University Press.
Crawford, T. and S. Vaughn (2009). "A predictive model of archaeological potential: An example from northwestern Belize " Applied Geography **29**(4): 542-555.
Crosby, D. G. (2004). The Poisoned Weed: Plants Toxic to Skin. New York, Oxford University Press.

Cushman, D. W. and L. Sebastian (2008). Integrating Archaeological Models: Management and Compliance on Military Installations Legacy Resource Management Program, Project 06-167. Rio Rancho, New Mexico, SRI Foundation.
Custer, J. F., T. Eveleigh, V. Klemas and I. Wells (1986). "Application of LANDSAT Data and Synoptic Remote Sensing to Predictive Models for Prehistoric Archaeological Sites: An Example from the Delaware Coastal Plain " American Antiquity **51**(3): 572-588.
Dalla Bona, L. (2000). Protecting cultural resources through forest management planning in Ontario using archaeological predictive modeling. In Practice applications of GIS for archaeologists: A predictive modeling kit. K. L. W. a. R. J. Brandon. London, Taylor and Francis 73-100
Dalla Bona, L. (1994). Predictive Modelling Methodology. Thunder Bay, Ontario, Lakehead University: Center for Archaeological Resource Prediction.
Dalla Bona, L. and L. Larcombe (1996). Modelling prehistoric land use in northern Ontario. New Methods, old problems: Geographic information systems in modern archaeological research. H. D. G. Maschner. Carbondale, Southern Illinois University**:** 252-271.
Dean, J. S., G. J. Gumerman, J. M. Epstain, R. L. Axtell, A. C. Swedlund, M. T. Parker and S. Mccarroll (2000). Unterstanding Anasazi Culture Change Through Agent-Based Modelling. Dynamics in Human and Primate Societies. Agent-Based Modeling of social and spatial processes. . T. A. Kohler and G. J. Gumerman. Oxford, Oxford University Press**:** 179-207.
Deeben, J., D. Hallewas, M. v. Leusen, H. Kamermans, P. Verhagen and P. Zoetbrood (2007). First thought on the incorporation of cultural variables into predictive modelling. Case Studies in Archaeological Predictive Modelling. . P. Verhagen, Leiden University Press**:** 203-210.
DeGloria, S. D., R. Elliot and S. J. Riley (1999). "A Terrain Ruggedness Index That Quantifies Topographic Heterogeneity." Intermountain Journal of Sciences **5**: 23-27.
Delpech, J. and D. K. Grayson (1998). "Changing diet breadth in the early upper paleolithic of Southwestern France." Journal of Archaeological Science **25**: 1119-1130.
Dering, P. (1999). "Earth-Oven Plant Processing in Archaic Period Economies: An Example from a Semi-Arid Savannah in South-Central North America." American Antiquity **64**(4): 659-674.
Derr, P. S. (1981). Soil Survey of Otero Area, New Mexico. Parts of Otero, Eddy, and Chaves Counties, National Cooperative Soil Survey.
DesJeans, T. P., C. Feiss, T. A. Kohler and D. F. Thompson (1980). An Archaeological survey of Selected Areas of the Fort Benning Military Reservation, Alabama and Georgia. Atlanta, Untied States Army.
Devitt, M., J. B. Hill and M. Sergeyeva (2009). Predictive Modeling and Cultural Resource Preservation in Santa Cruz County, Arizona Center for Desert Archaeology.
Dijkstra, E. W. (1959). "A note on two problems in connexion with graphs." Numerische Mathematik **1**: 269-271.
Dincauze, D. F. and J. W. Meyer (1976). Prehistoric resources of East-Central New England: A Preliminary Predictive Study University of Massachusetts, Amherst. Submitted to the National Park Service.
Dobres, M. A. and J. E. Robb (2000). Agency in archaeology: Paradigm or platitude? Agency in archaeology. M. A. Dobres and J. E. Robb. London, Routledge**:** 3-17.

Doneus, M., R. Gietl and M. Fera (2008). Cost Distance Analysis in an Alpine Environment. Comparison of Different Cost-surface Modules. Conference on Computer Applications and Quantitative Methods in Archaeology (CAA). A. Posluschny, K. Lambers and I. Herzog. Berlin.
Doran, J. (1999). "Prospects for Agent-Based Modelling in Archaeology. ." Achaeologia e Calcolatori **10**: 33-44.
Dore, C. D. and L. Wandsnider (1999). "Creating cultural resource data layers: experiences from the Nebraska Cultural resources GIS Project, Part II: combining GPS and GIS data. ." Bulletin of the Society of American Archaeology **13**(5): 16-18.
Dore, C. D. and L. Wandsnider (2006). Modeling for Management in a Compliance World. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 75-96.
Ducke, B., A. R. Millard and M. van Leusen (2009). Dealing with uncertainty in archaeological prediction. Archaeological Prediction and Risk Management. H. Kamermans, Leusen, Martijn van, and Verhagen, Philip Leiden University Press**:** 123-160.
Duke, C. and J. Steele (2009). "Geology and lithic procurement in Upper Palaeolithic Europe: a weights-of-evidence based GIS model of lithic resource potential." Journal of Archaeological Science **30**: 1-12.
Dunne, T. (1978). Field Studies of Hillslope Flow Processes. Hillslope Hydrology. M. J. Kirkby, John Wiley and Sons**:** 227-293.
Dunnell, R. C. and W. S. Dancey (1983). The siteless survey: A regional scale data collection Strategy. Advances in Archaeological Method and Theory. M. B. Schiffer. New York, Academic Press**:** 267-287.
Duran, M. S., P. Eidenbach, D. T. Kirkpatrick and K. W. Laumbach (2001). Basin and Range Archaeology an Overview of Prehistory in South-Central New Mexico. The Archaeological Record of Southern New Mexico: Sites and Sequences in Prehistory. S. R. Katz and P. Katz. Santa Fe, New Mexico Historic Preservation Division.
Ebert, D. (2004). "Applications of Archaeological GIS." Canadian Journal of Archaeology/Journal Canadien d’Archéologie **28**: 319-341.
Ebert, D. and M. Singer (2004). "GIS, predictive modelling, erosion, site monitoring." Assemblage(8).
Ebert, J. I. and T. A. Kohler (1988). The Theoretical Basis of Archaeological Predictive Modeling and a Consideration of Appropriate Data-collection Methods. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 97-171.
Ebert, J. I., S. Larralde and L. Wandsnider (1984). Theoretical, Methodological and Economic Aspects of Nonsite Surface Survey, Nonsite Sampling and Predictive Modeling. 49th Annual Meeting of the Society for American Archaeology. Protland.
Ebert, J. J. (2000). The State of the Art in “Inductive” Predictive Modeling: Seven Big Mistakes (and Lots of Smaller Ones). Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. K. L. W. a. R. J. Brandon. New York, Taylor & Francis 129-133.
Eighmy, J. L. and J. M. LaBelle (1996). "Radiocarbon Dating of Twenty-Seven Plains Complexes and Phases." Plains Anthropologist **41**(155): 53-69.
Emery, J. (2012). "20th Annual Mescal Roast celebrates Apache History and Culture in New Mexico." from http://www.lasplash.com/publish/Domestic_150/20th_Annual_Mescal_Roast_celebrates_Apache_History_and_Culture_in_New_Mexico.php.

Emlen, J. M. (1966). "The role of time and energy in food preference." The American Naturalist **100**: 611-617.
Engineer, S. (2004). In the Matter of the Application by Seven Rivers, Inc for Permit to Temporarily Change Location of Well and Place or Purpose of Use of Gorund Water in the Roswell Basin of New Mexico. S. E. o. N. Mexico. Santa Fe.
Epstein, J. M. (2008). "Why Model?" ournal of Artificial Societies and Social Simulation, **11**(4).
Ericson, J. E. and R. GoldStenin (1980). "Space: A New Approach to the Analysis of Energy Expenditure within Site Catchments." Anthropology **1**: 21-30.
Esparza-Carlos, J. P., J. W. Laundre and V. J. Sosa (2011). "Precipitation Impacts on Mule Deer Habitat Use in the Chihuahuan Desert of Mexico." Journal of Arid Environments **75**: 1008-1015.
Fábrega Álvarez, P. and C. Parcero Oubiña (2007). "Proposals for an archaeological analysis of pathways and movement." Archeologia e Calcolatori **18**: 121-140.
Faunce, K. V. (2000). The Perception of Landscape in the Use and Settlement of the Tularosa Basin, New Mexico. PHD, University of Idaho.
Ferretti, G., A. E. Minetti, C. Moia, G. S. Roi and D. Susta (2002). "Energy cost for walking and running at extreme uphill and downhill slopes." Journal of Applied Physiology **93**: 1039-1046.
Finkea, P. A., E. Meylemansb and J. Van de Wauwa (2008). "Mapping the possible occurrence of archaeological sites by Bayesian inference." Journal of Archaeological Science **35**(10): 2786-2796.
Fleure, H. J. (1916). "Valleyward Movement." Archaeologia Cambrensis **71**: 101-140.
Fox, C. (1923). The Archaeology of the Cambridge Region. Cambridge, Cambridge University Press.
Fuller, S., L. M. Gregonis and A. E. Rogge (1976). Orme Alternatives: The Archaeological Resources of Roosevelt Lake and Horseshoe Reservoir. Tucson, Arizona State Museum. **Archaeological Series**.
Gaffney, V., Z. Stancic and H. Watson (1996). Moving from Catchments to Cognition: Tentative Steps Towards a Larger Archaeological Context for GIS. Anthropology, Space, and Geographic Information Systems. M. Aldenderfer, and Maschner, Herbert D.G. London, Oxford University Press**:** 132-154.
Gaffney, V. and M. van Leusen (1995). GIS, environmental determinism and archaeology : a parallel text. Archaeology and Geographical Information Systems: A European Perspective. G. a. S. Lock, Z. London, Taylor & Francis**:** 367-382.
Garcia Sanjuan, L. and D. Wheatley (1999). "The State of the Arc: Differential Rates of Adoption of GIS for European Heritage Management." European Journal of Archaeology **2**(2): 201-228.
Garen, D. C. and D. S. Moore (2005). "Curve Number Hydrology in Water Quality Modeling: Uses, Abuses, and Future Directions." Journal of the American Water Resources Association: 377-388.
Gartner. (2008). "Understanding Gartner's Hype Cycles, 2008 " Retrieved 5/5/2012, 2012, from http://www.gartner.com/DisplayDocument?doc_cd=158921&ref=g_noreg.
Gazenbeek , M. and P. Verhagen (2006). The Use of predictive Modeling for Guiding the Archaeological Survey of Roman Pottery Kilns in the Argonne Region. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 447-461.

Gazenbeek, M. and P. Verhagen (2007). The Use of Predictive Modelling for Guiding the Archaeological Survey of Roman Pottery Kilns in the Argonne Region. Case Studies in Archaeological Predictive Modelling. . P. Verhagen, Leiden University Press.**:** 29-40.
Gehlbach, F. R. (1967). "Vegetation of the Guadalupe Escarpment, New Mexico-Texas." Ecology **48**(3): 404-419.
Gilbert, G. N. (2008). Agent-based models. No. 153, Sage.
Gilbert, G. N. and K. G. Troitzsch (2009). Simulation for the social scientist. Berkshire., Open University.
Gili, S., R. Mico, R. Risch and P. Verhagen (2007). Modelling Prehistoric Land Use Distribution in the Rio Aguas Valley. Case Studies in Archaeological Predictive Modelling. P. Verhagen, Leiden University Press.
Gillings, M. and D. Wheatley (2002). Spatial Technology and Archaeology The Archaeological Application of GIS London, Taylor and Francis
Goel, G., D. Jackson, I. Moore and T. G. Whitley (2009). Beyond the Marsh: Settlement Choice, Perception, and Spatial Decision-Making on the Georgia Coastal Plain. CAA 2009 Williamsburg: Making History Interactive. Williamsburg, Virginia.
Goings, C. A. (2003). "A Predictive Model for Lithic Resources in Iowa." The Plains Anthropologist **48**(184): 53-67.
Goodbar, J. and S. Rice-Snow (2012). " Terrain factors in Capitan aquifer recharge,northeastern Guadalupe escarpment, New Mexico." New Mexico Geology **34**(1): 15-22.
Goudswaard, B., R. Isarin and P. Verhagen (2009). Archaeology as a risk in spatial planning: manoeuvring between objectivity and subjectivity. Archaeological Prediction and Risk Management. H. Kamermans, Leusen, Martijn van, and Verhagen, Philip Leiden University Press**:** 41-48.
Grady, M. A., M. McCarth and T. R. Scott (1978). Archaeological Survey in Cherokee, Smith and Rusk Counties, Texas: A lesson in Survey Methods. Archaeological Research Report 116. Dallas, Department of Anthropology, Southern Methodist University.
Graham, S. (2009). "Behaviour Space: Simulating Roman Social Life and Civil Violence. ." Digital Studies/Le champ numérique, **1**(2).
Graves, D. (2010). "The use of predictive modelling to target Neolithic settlement and occupation activity in mainland Scotland." Journal of Archaeological Science **30**(1): 1-24.
Gumerman, G. J. and T. A. Kohler, Eds. (2000). Dynamics in human and primate societies: Agent-based modeling of social and spatial processes. New York, Oxford University Press.
Hackenberger, S. (1978). Reconnaissance Sampling of Fort Sill: Methods and Results. An Archaeological Reconnaissance of Fort Sill, Oklahoma. C. R. Ferring. Lawton, Museum of the Great Plains**:** 118-161.
Haise, H. R. and M. E. Jensen (1963). "Estimating evapotranspiration from solar radiation." Journal of Irrigation and Drainage **96**: 25-28.
Hale, W. E. (1945). Ground-Water Condition in the Vicinity of Carlsbad, New Mexico. Santa Fe, State Engineer of New Mexico.
Hall, S. A. (1982). "Late Holocene Paleoecology of the Southern Plains." Quaternary Research (17): 391-407
Hames, R. and W. Vickers (1982). "Optimal diet breadth theory as a model to explain variability in Amazonian hunting." American Ethnologist **9**: 358-378.

Harris, T. and G. Lock (2006). Enhancing Predictive Archaeological Modeling: Integrating Location, Landscape, and Culture. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 41-62.
Harris, T. M. and G. Lock (1995). Toward an evaluation of GIS in European archaeology: the past, present and future of theory and applications. Archaeology and Geographical Information Systems: A European Perspective. G. a. S. Lock, Z. London, Taylor & Francis**:** 349-365.
Hasenstab, R. J. (1996). Settlement As Adaptation: Variability in the Iroquois Village Site Selection As Inferred Through GIS. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press**:** 223-241.
Hasenstab, R. J. and B. Resnick (1990). GIS in historical predictive modelling; a Fort Drum Project. Interpreting Space: GIS and archaeology. K. M. Allen, Green, S.W., and Zubrow, E.B.W. Taylor Francis, London**:** 284-306.
Hatzinikolaou, E. G. (2006). Quantitative methods in Archaeological Predictions: From Binary to Fuzzy Logic. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 437-446.
Hawkes, K. and J. F. O'Connell (1984). "Food choice and foraging sites among the Alyawara." Journal of Anthropological Research **40**: 435-504.
Hawkins, R. H. and V. M. Ponce (1996). "Runoff Curve Number: Has It Reached Maturity?" Journal of Hydrologic Engineering: 11-19.
Hawley, J. W. (1986). Phsiographic Provinces. New Mexico in Maps. J. L. Williams. Albuquerque, University of New Mexico Press**:** 23-27.
Hawthorne, L. S. (1994). I Never Left a Place That I Didn't Clean Up: The Legacy of Historic Settlement on Lands Administered by Holloman Air Force Base. Cultural Resources Publication. Holloman Air Force Base, U.S. Department of Defence, U.S. Air Force, Air Combat Command, Holloman Air Force Base, New Mexico. **1**.
Herzog, I. (2010). Theory and Practice of Cost Functions. Computer Applications and Quantitative Methods in Archaeology. J. Melero, P. Cano and J. Revelles.
Herzog, I. (2012). The Potential and Limits of Optimal Path Analysis.
Herzog, I. and Posluschny (2011). Tilt – Slope-Dependent Least Cost Path Calculations Revisited. Computer Applications and Quantitative Methods in Archaeology. E. Jerem and R. Szeverenyi. Budapest.
Higgs, E. S. and C. Vita-Finzi (1972). Prehistoric economies. A territorial approach. Papers in Economic Prehistory. E. S. Higgs. Cambridge Cambridge University Press**:** 27-36.
Hill, J. N. and F. T. Plog (1971). Explaining Variability in the Distribution of Sites. The Distribution of Prehistoric Population Aggregates. G. J. Gumerman. Prescott, Prescott College Press**:** 7-36.
Hodgson, W. (2001). Food Plants of the Sonoran Desert. Tucson, University of Arizona Press.
Hogan, P. H. (2006). Southeastern New Mexico Regional Research Design and Cultural Resource Management Strategy. Albuquerque, Office of Contract Archeology, University of New Mexico.
Holland, J. H. and J. H. Miller (1991). "Artificial Adaptive Agents in Economic Theory." American Economic Review **81**: 365-371.
Holmer, R. N. (1979). Split Mountain Cultural Study Tract. Vernal, Utah, Archaeological Center, Department of Utah. Submitted to the Bureau of Land Management.

Hong, N.-M., C.-J. Lin, Y.-P. Lin and P.-J. Wu (2007). "Modeling and assessing land-use and hydrological processes to future land-use and climate change scenarios in watershed land-use planning." Environmental Geography **53**: 623-634.
Howey, M. C. (2007). "Using multi-criteria cost surface analysis to explore past regional landscapes: A case study of ritual activity and social interaction in Michigan, A.D. 1200–1600." Journal of Archaeological Science(34): 1830–1846.
Hunter, A. J. S. and S. Steiniger (2010). Free and Open Source GIS Software for Building a Spatial Data Infrastructure. OGRS 2009, Springer.
Husdal, J. (2000). How to make a straight line square. Network Analysis in Raster GIS with time-dependent cost variables. MSc, Leicester.
Hussain, A. and M. Niazi (2011). "Agent-based computing from multi-agent systems to agent-based models: a visual survey." Scientometrics: 1-21.
Ingold, T. (1993). "The temporality of the landscape." World Archaeology **25**: 152-174.
Ingold, T. (2000). The perception of the environment. Essays on livelihood, dwelling and skill. London, Routledge.
Irwin-Williams, C. (1979). Post-Pleistocene Archeology, 7000–2000 BC. Southwest. A. Ortiz. Washington DC, Smithsonian Institution. **9:** 31-42.
Isard, W. (1956). "Location and space economy: a general theory relating to industrial location, market, land use, trade and urban structure." The Regional Science Studies Series **1**.
Jelinek, A. J. (1967). A Prehistoric Sequence in the Middle Pecos Valley, New Mexico. Ann Arbor, University of Michigan, Museum of Anthropology.
Jeneson, K. and P. Verhagen (2012). A Roman Puzzle. Trying to Find the Via Belgica with GIS. Thinking Beyond the Tool. A. Chrysanthi, P. Murrieta Flores and C. Papadopoulus. Oxford, BAR International Series **2344:** 123-130.
Jochim, M. A. (1976). Hunter-Gatherer Subsistence and Settlement: A Predictive Model. New York, Academic Press.
Johnson, E. and V. T. Holliday (1989). "Environmental Change on the Southern High Plains, USA." Journal of Quaternary Sciences **4**(2): 145-165.
Johnson, E. and V. T. Holliday (1995). "Late Quaternary Environments and Archaeology of the Southern High Plains." Bulletin of the Texas Archeology Society(66): 519-540.
Johnson, J. K. (1996). Delta Digitizing: GIS and Remote Sensing in Northwest Mississippi. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press**:** 242-251.
Jones, K. T. and D. B. Madsen (1989). "Calculating the cost of resource transportation." Current Anthropology **30**: 529-534.
Joolen, E. v. (2003). Archaeological land evaluation: a reconstruction of the suitability of ancient landscapes for various land uses in Italy focused on the first millennium BC. Groningen, Rijksuniversiteit Groningen.
Judge, W. J. (2007). Early Man: Plains and Southwest. An Interpretive Summary of the PaleoIndian Occupation of the Plains and Southwest. Handbook of North American Indians. D. Ubelaker, W. C. Sturtevant and D. Stanford. Washington D.C., Smithsonian Institution. **3**.
Judge, W. J. and L. Sebastian, Eds. (1988). Quantifying the present and predicting the past: theory, method and application of archaeological predictive modelling. Denver, US Department of Interior, Bureau of Land Management.
Kahneman, D. and A. Tversky (1979). "Prospect theory: An analysis of decision under risk." Econometrica **47**: 263-291.

Kahneman, D. and A. Tversky (1992). "Advances in prospect theory: Cumulative representation of uncertainty." Journal of Risk and Uncertainty **5**(4): 297-323.
Kamermans, H. (2000). Land evaluation as predictive modelling: a deductive approach. Beyond the Map: archaeology and spatial technologies. G. Lock. Amsterdam, IOS Press**:** 124-146.
Kamermans, H. (2006). Problems in Paleolithic Land Evaluation: A Cautionary Tale. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 97-122.
Kamermans, H. (2007). "Smashing the Crystal Ball. A Critical Evaluation of the Dutch National Archaeological Predictive Model." International Journal of Humanities and Arts Computing **1**(1): 71–84.
Kamermans, H., M. van Leusen and P. Verhagen, Eds. (2009a). Archaeological Prediction and Risk Management, Leiden University Press.
Kamermans, H., M. van Leusen and P. Verhagen (2009b). The future of archaeological predictive modelling. Archaeological Prediction and Risk Management. H. Kamermans, Leusen, Martijn van, and Verhagen, Philip Leiden University Press**:** 19-25.
Kamermans, H. and M. v. van Leusen, Eds. (2005). Predictive Modelling for Archaeological Heritage Management: A research agenda. Nederlandse Archeologische Rapporten. Amersfoort.
Kammerer, J. C. (2005). "Largest Rivers in the United States."   Retrieved 27/7/2012, 2012, from http://pubs.usgs.gov/of/1987/ofr87-242/.
Kantner, J. and R. Hobgood (2003). Digital Technologies and Prehistoric Landscapes in the American Southwest. The Reconstruction of Archaeological Landscapes through Digital Technologies. M. Forte and P. R. Williams, BAR International Series. **1151:** 117-124.
Katz, S. R. and P. Katz (1985). The Prehistory of the Carlsbad Basin, Southeastern New Mexico. San Antonio, Texas, Bureau of Reclamation, Southwest Regional Office.
Katz, S. R. and P. Katz (2001). The Southeast. The Archaeological Record of Southern New Mexico: Sites and Sequences in
Prehistory. S. R. Katz and P. Katz. Santa Fe, New Mexico Historic Preservation Division.
Kaufmann, K. E. (2006). Appropriateness and Applicability of GIS and Predictive Models with Regard to Regulatory and Nonregulatory Archaeology. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 265-278.
Kay, S. J. and R. E. Witcher (2009). "Predictive modelling of Roman settlement in the Middle Tiber Valley." Archeologia e Calcolatori(20): 277-290.
Keene, A. S. (1983). Biology, behaviour, and borrowing: a critical examination of optimal foraging theory in archaeology. Archaeological Hammers and Theories. J. A. Moore and A. S. Keene. New York, Academic Press**:** 137-155.
Kellogg, D. C. (1987). "Statistical relevance and site location data." American Antiquity **52**: 143-150.
Kelly, C. J. (1937). "The Route of Antonio De Espejo Down the Pecos River and Across the Texas Trans-Pecos Region in 1583; Its Relation to West Texas Archeology." West Texas Historical and Scientific Society **18**(4).
Kelly, R. L. (1995). The foraging spectrum: Diversity in hunter–gatherer lifeways. Washington DC, Smithsonian Institution.
Kennett, D. and B. Winterhalder (2006). Behavioral ecology and the transition from hunting and gathering to agriculture. Behavioral ecology and the transition to agriculture. B. Winterhalder and D. Kennett. Berkeley, University of California Press**:** 1-21.

Kohler, T. A. (1988). Predictive Locational Modeling: History and Current Practice. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 19-59.
Kohler, T. A. and K. Kvamme (1988). Geographic Information Systems: Technical Aids for Data Collection, Analysis, and Display. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. a. S. Judge, Lynn. Denver, US Department of the Interior, Bureau of Land Management Service Center**:** 493-547.
Kohler, T. A. and S. Parker (1986). Predictive models for archaeological resource location In Advances. Archaeological Methods and Theory. M. B. Schiffer. Oxford, Oxford University press**:** 397-452
Kohler, T. A. and S. van der Leeuw, Eds. (2007). Model-based archaeology of socionatural systems. Santa Fe, School of Advanced Research.
Kohler, T. A. and C. Van West (1996). A Time to Rend, A Time to Sew: New Perspectives on Northern Anasazi Sociopolitical Development in Late Prehistory. Anthropology, Space, and Geographic Information Systems. M. Aldenderfer and H. D. G. Maschner. New York, Oxford University Press.
Krebs, J. R. and D. W. Stephens (1986). Foraging theory. Princeton Princeton University Press.
Kuiper, J. A., J. Leveson and B. Verhaaren (2006). Relating Cultural Resources to Their Natural Environment Using the IEDROK GIS: A cultural Resources Management for the Republic of Korea. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 241-264.
Kuiper, J. A. and K. L. Wescott (2000). Using a GIS to Model Prehistoric Site Distribution in the Upper Chesapeake Bay. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. R. J. Konnie L. Wescott and Brandon. New York, Taylor & Francis 59-67.
Kvamme, K. (1989). Geographic information systems in regional archaeological research and data management. Archaeological Method and Theory. M. B. Schiffer. Tucson, University of Arizona Press. **1:** 139-203.
Kvamme, K. L. (1983a). "Computer processing techniques for regional modeling of archaeological site locations." Advances in Computer Archaeology(1): 26-52.
Kvamme, K. L. (1983b). A manual for predictive models: Examples from the Grand Junction District, Colorado. Denver, Bureau of Land Management.
Kvamme, K. L. (1985). Determining empirical relationships between the natural environment and prehistoric site locations: a hunter-gatherer example. For concordance in archaeological analysis: Bridging data structure, quantitative technique and theory. C. Carr. Kansas City Missouri, Westport**:** 208-238
Kvamme, K. L. (1988a). Using existing Archaeological Survey Data for Model Building. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 301-323.
Kvamme, K. L. (1988b). Development and Testing of Quantitative Models. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. W. J. Judge, and Sebastian, Lynne. Denver, Bureau of Land Management**:** 325-428.

Kvamme, K. L. (1990). The fundamental principles and practice of predictive archaeological modelling. Mathematics and information science in archaeology: a flexible framework. A. Voorips. Bonn, Holos-Verlag: 257-295.
Kvamme, K. L. (1999). "Recent Directions and Developments in Geographical Information Systems." Journal of Archaeological Research **7**(2): 153-201.
Kvamme, K. L. (2006). There and Back Again: Revisiting Archaeological Locational Modeling. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 3-40.
Ladefoged, T. N. and M. D. McCoy (2009). "New Developments in the Use of Spatial Technology in Archaeology " Journal of Archaeological Research **17**(3): 263-295.
Lake, M. (2010). The uncertain future of simulating the past. Simulating Change. Archaeology into the twenty-first century. Foundations of Archaeological Inquiry. A. Costopoulos and M. Lake. Salt Lake Citiy, The University of Utah Press: 12-20.
Lake, M. (2014). "Trends in Archaeological Simulation." Journal of Archaeological Method and Theory **21**(2): 258-287.
Langmuir, E. (1995). Mountaincraft and Leadership. Edinburgh, MLTUK.
Larralde, S. and L. Sebastian (1989). Living on the Land: 11,000 Years of Human Adaptation in Southeastern New Mexico. Roswell, Bureau of Land Management, Roswell District, Roswell, New Mexico.
Lee, E. and M. Newman (2012). The IMSIG Survey results. IfA Conference. Oxford, UK.
Lee, J., M. J. Munro-Stasiuk and C. Yu (2003). "Extensions to least-cost path algorithms for roadway planning." Int. Journal Geographical Information Science **17**(4): 361-376.
Legg, R. J. and D. Taylor (2005). "Modeling environmental influences on the locations of Irish early medieval ringforts." Geoarchaeology **21**(3): 201-220.
Leslie, R. H. (1979). The Eastern Jornada Mogollon: Extreme Southeastern New Mexico (A Summary). Jornada Mogollon Archaeology: Proceedings of the First Jornada Conference. P. H. Beckett and R. N. Wiseman. Santa Fe, New Mexico Historic Preservation Division: 179-199.
Levin, S. A. (1995). The problem of pattern and scale in ecology. Ecological Time Series. T. M. Powell and J. H. Steele. New York, Chapman & Hall: 277-326.
Linacre, E. T. (1994). "Estimating U.S. Class-A pan evaporation from few climate data." Water International **19**: 5-14.
Llobera, M. (1996). "Exploring the topography of mind: GIS, social space and archaeology." Antiquity **70**: 612-622.
Llobera, M. (2000). Understanding movement: a pilot model towards the sociology of movement. Beyond the Map: Archaeology and Spatial Technologies. G. Lock. Amsterdam, IOS Press: 65-84.
Llobera, M. (2001). "Building past landscape perception with GIS: Understanding topographic prominence." Journal of Archaeological Science **28**: 1005-1014.
Llobera, M. (2003). "Extending GIS-based visual analysis: The concept of visualscapes." International Journal of Geographical Information Science **17**: 25-48.
Llobera, M. and T. J. Sluckin (2007). "Zigzagging: Theoretical insights on climbing strategies." Journal of Theoretical Biology **239**: 206-217.
Lock, G. R. and Z. Stançic, Eds. (1995). Archaeology and Geographical Information Systems: a European Perspective. London, Taylor and Francis.
Macal, C. M. and M. J. North (2006). Introduction to Agent-based Modeling and Simulation MCS LANS Informal Seminar. Argonne National Laboratory, Argonne, IL

MacArthur, R. H. and E. R. Pianka (1966). "On the optimal use of a patchy environment." The American Naturalist **100**: 603-609.
Machielse, I. A., D. R. M. Timmermans and P. P. Wakker (2003). The effects of statistical information on insurance decisions and risk attitudes. Amsterdam, Department of Economics, University of Amsterdam.
Madey, G. and C. Nikolai (2009). "Tools of the Trade: A Survey of Various Agent Based Modeling Platforms " Journal of Artificial Societies and Social Simulation **12**(22).
Marcy, R. B. (1850). The report of the route from Fort Smith to Santa Fe. Reconnaissance in New Mexico and Texas, etc. Reports of the Secretary of War. Washington.
Marozas, B. A. and J. A. Zack (1990). GIS and archaeological site location. Interpreting Space: GIS and archaeology. K. M. Allen, Green, S.W., and Zubrow, E.B.W. London, Taylor Francis**:** 165-172.
Maschner, H. D. G. (1996). The politics of Settlement Choice on the Northwest Coast: Cognition, GIS, and Coastal Landscapes. Anthropology, Space, and Geographic Information Systems. M. Aldenderfer, and Maschner, Herbert D.G. London, Oxford University Press**:** 175-189.
McClaran, M. P. and T. R. Van Devender, Eds. (1995). The Desert Grassland. Tucson, The University of Arizona Press.
McDonnell, J. J. (2003). "Where does water go when it rains? Moving beyond the variable source area concept of rainfall-runoff response." Hydrological Processes **17**: 1869-1875.
McHugh, T. (1958). "Social Behavior of the American Buffalo " Zoologica **43**: 1-40.
MDOT. (2009). "Mn/Model Statewide Archaeological Predictive Model." Retrieved 11/11/2009, 2009, from http://www.mnmodel.dot.state.mn.us/.
Mehren, L. L. (1969). A History of the Mescalero Apache Reservation, 1869–1881. Masters, University of Arizona.
Mehrer, M. W. and K. L. Wescott, Eds. (2006). GIS and Archaeological Site Location Modeling. New York, Taylor & Francis.
Mera, H. P. (1938). Reconnaissance and excavation in southeastern New Mexico, American Anthropology Association.
Mesterházy, G., G. Padányi-Gulyás and M. Stibrányi (2012). Familiar road, unfamiliar ground: Archaeological Predictive Modelling in Hungary. CAA 2012. Southampton.
METI and NASA (2012). ASTER GDEM 2
Millard, A. R. (2005). What can Bayesian statistics do for archaeological predictive modelling? Predictive modelling for archaeological heritage management: a research agenda. P. M. van Leusen and H. A. Kamermans, Rijkdienst voor het Oudheidkundig Bodemonderzoek.
Mink, P. B., D. Pollack and B. J. Stokes (2006). Points vs. Polygons: A Test Case Using a Statewide GIS. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 219-240.
Monteith, J. L. (1965). Evaporation and the Environment. The state and movement of water in living organisms. G. E. Fogg. Cambridge, Cambridge University Press**:** 205-234.
Motts, W. S. (1968). "The Control of Ground–Water Occurrence by Lithofacies in the Guadalupian Reef Complex near Carlsbad, New Mexico." Geological Society of America Bulletin **79**(3): 283-298.
Mount, J. and K. R. Schwarz (2006). Integrating Spatial Statistics into Archaeological Data Modeling. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 163-190.

Murrieta-Flores, P. (2012). "Understanding human movement through spatial technologies. The role of natural areas of transit in the Late Prehistory of South-western Iberia." Trabajos De Prehistoria **69**(1): 103-122.
National Geodetic Survey, U. S. (2012). El Capitan NGS data sheet.
Naunapper, L. S. (2006). Archaeological GIS in Environmental Impact Assessment and Planning. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 279-290.
Neustupny, E. (1995). Beyond GIS. Archaeology and Geographical Information Systems: A European Perspective. G. a. S. Lock, Z. London, Taylor & Francis**:** 134-139.
NMAES, N. M. A. E. S. (1971). Soil Survey Eddy Area, New Mexico. Wahsington DC, USDA.
Office of the State Engineer, N. M. (2005). "Pecos River Basin." Retrieved 27/6/2012, 2012, from http://www.ose.state.nm.us/isc_pecos.html.
Opler, M. E. and C. H. Opler (1950). "Mescalero Apache History in the Southwest." New Mexico Historical Review **25**(1): 1-36.
Orians, G. F. and N. E. Pearson (1979). On the theory of central place foraging. Analysis of ecological systems. D. J. Horn, R. D. Mitchell and C. R. Stairs. Columbus, Ohio State University Press**:** 154-177.
Orton, C. (2003). "The Fourth Umpire: Risk in archaeology (Inaugural Lecture)."
Parmenter, R. R. and T. R. Van Devender (1995). Diversity, Spatial Variability, and Functional Roles of Vertebrates in the Desert Grassland. The Desert Grassland. M. P. McClaran and T. R. Van Devender. Tucson, The University of Arizona Press.
Podobnikar, T., K. Oštir, T. Veljanovski and Z. Stanèiè (2001). Archaeological Predictive Modelling in Cultural Resource Management. Brono.
Podobnikar, T. and I. Šprajc (2007). Spatial analyses and Maya cultural landscape.
Premo, L. S. (2008). Exploring behavioral terra incognita with archaeological agent-based models. Beyond Illustration: 2D and 3D Technologies as Tools of Discovery in Archaeology. . B. Frischer and A. Dakouri-Hild. Oxford, ArchaeoPress, .
Premo, L. S. (2010). Equifinality and explanation: The role of agent-modelling in post-positivist archaeology. Simulating Change. Archaeology into the twenty-first century. Foundations of Archaeological Inquiry. A. Costopoulos and M. Lake. Salt Lake City, University of Utah Press**:** 28-37.
Rapp, G. (2009). Archaeomineralogy. New York, Springer.
Reaney, S. M. (2008). "The use of agent based modelling techniques in hydrology: determining the spatial and temporal origin of channel flow in semi-arid catchments." Earth Surface Processes and Landforms **33**: 317-327.
Reeves, C. C. (1972). Tertiary-Quaternary Stratigraphy and Geomorphology of West Texas and Southeastern New Mexico. Guidebook of East-Central New Mexico. V. C. T. Kelley, F. D. Socorro, New Mexico Geological Society 23rd Field Conference. New Mexico Bureau of Mines and Minerals**:** 108-117.
Renfrew, C. and E. B. Zubrow (1994). The ancient mind. Elements of cognitive archaeology. Cambridge Cambridge University Press.
Resources, N. M. B. o. G. M. (2012). "Maps." Retrieved 6/6/2012, 2012, from http://geoinfo.nmt.edu/publications/maps/home.html.
Ridges, M. (2006). Regional Dynamics of Hunting and Gathering: An Australian Case Study Using Archaeological Predictive Modeling. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 123-146.

Ritzema, H. P., Ed. (1994). Drainage Principles and Applications. Wageningen, International Institute for Land Reclamation and Improvement.
Robb, J. and D. Van Hove (2003). "Gardening, foraging and herding: Neolithic land use and social territories in Southern Italy." Antiquity **77**(241-254).
Robertson, B. P. and L. B. Robertson (1978). The Generation of Locational Models in an Inductive Framework. Conservation Archaeology in the Northeast: Toward a Research Orientation. A. E. Spiess. Cambridge, Peabody Museum, Harvard University**:** 27-36.
Rocks-Macqueen, D. (2010). Investigation of Data Errors and Cultural Attributes on Site Predictive Modelling. Masters, University of Newcastle.
Rocks-Macqueen, D. (2016). "When You are more Likely to Die of Cancer than Become an Academic: What is the Role of PhD students?" Papers from the Institute of Archaeology **26**(1).
Roney, J. R. (1985). Prehistory of the Guadalupe Mountains. MA, Eastern New Mexico University.
Rothman (1998). Promise Beheld and the Limits of Place: A Historic Resource Study of Carlsbad Caverns and Guadalupe Mountains National Parks and the Surrounding Areas. Washington, DC, Department of the Interior
National Park Service.
Savage, S. H. (1990). GIS in archaeological research. Interpreting space. K. M. Allen, S. W. Green and E. B. W. Zubrow. London, Taylor & Francis**:** 22-33.
Schelling, T. C. (1971). "Dynamic Models of Segregation." Journal of Mathematical Sociology **1**: 143-186.
Schneider-Hector, D. (1993). White Sands: The History of a National Monument. Albuquerque, University of New Mexico Press.
Schriever, B. A., M. S. Shackley and M. S. Taliaferro (2010). "Obsidian procurement, least cost path analysis, and social interaction in the Mimbres area of southwestern New Mexico." Journal of Archaeological Science **37**: 536-548.
Senour, K., Jr. (1980). Field/Technical Reports Supporting the Bull Creek Environmental Impact Statement, White County, Arkansas. Oklahoma City, Mid-Arkansas Regional Water Distribution District.
Shanks, M. and C. Tilley (1987). Re-constructing archaeology: theory and practice. New York, Routledge.
Shannon, R., and James D. Johannes. (1977). "Systems Simulation: The Art and Science." Systems, Man and Cybernetics, IEEE Transactions **10**: 723-724.
Shermer, S. J. and J. A. Tiffany (1985). "Environmental variables as factors in site location: an example from the upper Midwest." Midcontinental Journal of Archaeology **10**: 215-240.
Simms, S. R. (1987). Behavorial ecology and hunter–gatherer foraging: An example from the Great Basin. Oxford.
Smith, E. A. (1991). Inujuamiut foraging strategies: Evolutionary ecology of an arctic hunting economy. New York, Aldine.
Society, A. M. (2012). "Rain."
Stafford, T. J. (1981). "Alluvial Geology and Archaeological Potential of the Texas Southern High Plains." American Antiquity **46**: 548-565.
Stancic, Z. and T. Veljanovski (2000). Understanding Roman settlement patterns through multivariate statistics and predictive modeling. Beyond the Map: Archaeology and Spatial Technologies. G. Lock. Amsterdam, IOS Press**:** 147-156.

Stancic, Z. and T. Veljanovski (2006). Predictive Modeling In Archaeological Location Analysis and Archaeological Resource Management: Principles and Applications. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 393-412.
Thomas, A. V. (1988). A Survey of Predictive Locational Models: Examples from the Late 1970s and Early 1980s. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological
Predictive Modeling. W. J. Judge and L. Sebastian. Denver, USDI Bureau of Land Management**:** 581-645.
Tobler, W. (1993). Three Presentations on Geographical Analysis and Modeling. Santa Barbara, National Center for Geographic Information and Analysis.
Tol, A. and P. Verhagen (2007). Establishing optimal core sampling strategies: theory, simulation and practical implications. Case Studies in Archaeological Predictive Modelling. P. Verhagen. Leiden, Leiden University Press.
Trifković, V. (2005). Construction of space in Early Holocene iron gates. Oxford, The Institute of Archaeology, University of Oxford.
Trigger, B. (2006). A History of Archaeological Thought. Cambridge, Cambridge University Press.
Truett, J. (1996). "Bison and Elk in the American Southwest: In Search of the Pristine." Environmental Management **20**(2): 195-206.
Turc, L. (1954). "Le bilan d'eau des sols. Relations entre les précipitations, l'évaporation et l'écoulement. Ann." Agron **6**: 5-131.
USDA-SCS, U. S. D. o. A.-S. C. S. (1972). Estimation of Direct Runoff From Storm Rainfall. S. C. S. U.S. Department of Agriculture. Washington DC**:** 10.11-10.24.
USDA, U. S. D. o. A. (1971). Soil Survey Eddy Area, New Mexico. Washington DC, USDA.
USDA, U. S. D. o. A. (1981). Soil Survey of Otero Area, New Mexico. Parts of Otero, Eddy, and Chaves Counties. Washington DC.
USDA, U. S. D. o. A. (1986). Urban Hydrology for Small Watersheds. U. S. D. o. Agriculture. Washington DC, U. S. Department of Agriculture.
USDA, U. S. D. o. A. (2012). "Web Soil Survey." Retrieved 6/6/2012, 2012, from http://websoilsurvey.nrcs.usda.gov/app/HomePage.htm.
USGS. (2012a). "History of USGS Topographic Mapping." Retrieved 6/6/2012, 2012, from http://gallery.usgs.gov/videos/261#.T9TDBVJntbw.
USGS. (2012b). "USGS GAP Analysis Program." 2012, from http://gapanalysis.usgs.gov/.
USGS, U. S. G. S. (2012c). "New Mexico Geology." Retrieved 6/6/2012, 2012, from http://mrdata.usgs.gov/sgmc/nm.html.
Van Devender, T. R. (1980). "Holocene Plant Remains from Rocky Arroyo and Last Chance Canyon, Eddy County, New Mexico." The Southwestern Naturalist **25**(3): 361-372.
Van Devender, T. R. (1995). Desert Grassland History: Changing Climates, Evolution, Biogeography, and Community Dynamics. The Desert Grassland. M. P. McClaran and T. R. Van Devender. Tucson, The University of Arizona Press.
Van Devender, T. R., W. G. Spaulding and M. R. Phillips (1979). Late Pleistocene Plant Communities in the Guadalupe Mountains. Biological Investigations in the Guadalupe Mountains National Park, Texas. H. H. Genoways and R. J. Baker. Washington, D.C. **4:** 13-30.
van Leusen, M. (1996). GIS and Locational Modeling in Dutch Archaeology: A review of Current Approaches. New Methods, Old Problems: Geographic Information Systems in

Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press: 177-197.
van Leusen, M. (2002). Pattern to Process. Methodological investigations into the formation and interpretation of spatial patterns in archaeological landscapes, Rijksuniversiteit Groningen.
Verbruggen (2009). On costs and benefits in archaeological prospection. Archaeological Prediction and Risk Management. H. Kamermans, Leusen, Martijn van, and Verhagen, Philip Leiden University Press: 27-32.
Verhagen, P. (2006). Quantifying the Qualified: The Use of Multicriteria Methods and Bayesian Statistics for the Development of Archaeological Predictive Models. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 191-218.
Verhagen, P., Ed. (2007a). Case Studies in Archaeological Predictive Modelling, Leiden University Press.
Verhagen, P. (2007b). Prospection strategie and archaeological predictive modelling. Case Studies in Archaeological Predictive Modelling. . P. Verhagen, Leiden University Press: 101-114.
Verhagen, P. (2007c). Predictive models put to the test. Case Studies in Archaeological Predictive Modelling. P. Verhagen, Leiden University Press: 115-168.
Verhagen, P. (2007d). A Condensed History of Predictive Modelling in Archaeology. Case Studies in Archaeological Predictive Modelling. P. Verhagen. Leiden, Leiden University Press.
Verhagen, P. (2007e). Some considerations on the use of archaeological land evaluation. Case Studies in Archaeological Predictive Modelling. P. Verhagen. Leiden, Leiden University Press: 193-200.
Verhagen, P. (2007f). Quantifying the Qualified: the Use of Multicriteria Methods and Bayesian Statistics for the Development of Archaeological Predictive Models. Case Studies in Archaeological Predictive Modelling. P. Verhagen. Leiden, Leiden University Press: 71-92.
Verhagen, P. (2009). Testing archaeological predictive models: a rough guide. Archaeological Prediction and Risk Management. H. Kamermans, Leusen, Martijn van, and Verhagen, Philip Leiden University Press: 63-70.
Verhagen, P. and T. G. Whitley (2011). "Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene." Journal of Archaeological Method and Theory 19(1) 49–100.
Verhart, L. and M. Wansleeben (1997). "Geographical Information Systems Methodological progress and theoretical decline." Archaeological Dialogues **4**(1): 53-64.
Vermeulen, F. (2006). Understanding Lines in the Roman Landscape: A Study of Ancient Roads and Field Systems Based on GIS Technology. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York: 291-316.
von Neumann, J. (1951). The general and logical theory of automata. Cerebral Mechanisms in Behavior – The Hixon Symposium. L. A. Jeffress. New York, John Wiley & Sons: 1-31.
Waguespack, N. M. (2005). "The Organization of Male and Female Labor in Foraging Societies: Implications for Early Paleoindian Archaeology." American Anthropologist **107**(4): 666-676.
Warren, R. E. (1990a). Predictive modeling of archaeological site location: a case study in the Midwest. Interpreting Space: GIS and Archaeology. K. Allen, Green, S., and Zubrow, E. London, Taylor & Francis: 201-215.

Warren, R. E. (1990b). Predictive Modelling in archaeology: a Primer. Interpreting Space: GIS and archaeology. K. M. Allen, Green, S.W., and Zubrow, E.B.W. London, Taylor Francis**:** 90-111.
Weber, D. J. (1994). The Spanish Frontier in North America, Yale University Press.
Wescott , K. L. (2000). Introduction. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. L. W. Konnie, and Brandon, R. Joe. London, Taylor & Francis**:** 1-5.
Wescott, K. L. (2006). One Step Beyond: Adaptive Sampling and Analysis Techniques to Increase the Value of Predictive Models. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 63-74.
Whalen, M. E. (1994). "Moving out of the Archaic on the Edge of the Southwest." American Antiquity **59**(4): 622-638.
Wheatley, D. (1993). Going over old ground: GIS, archaeological theory and the act of perception. Computing the Past: Computer Applications and Quantitative Methods in Archaeology. J. Andresen, Madsen, T., and Scollar, I. Aarhus, Aarhus University press**:** 133-138.
Wheatley, D. (1995). Curlative viewshed analysis: a GIS-based method for investigating intervisibility, and its archaeological application. Archaeology and Geographical Information Systems: A European Perspective. G. a. S. Lock, Z. London, Taylor & Francis**:** 171-185.
Wheatley, D. (1996). The Use of GIS to Understand Regional Variation in Earlier Neolithic Wessex. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. H. D. G. Maschner. Carbondale, Southern Illinois University Press**:** 75-103.
Wheatley, D. (2004). "Making space for an archaeology of place." Internet Archaeology **15**.
Whitley, T. G. (2000). Dynamical Systems Modeling in Archaeology: A GIS Approach to Site Selection Processes in the Greater Yellowstone Region. PhD, University of Pittsburgh.
Whitley, T. G. (2003). Causality and Cross-Purposes in Archaeological Predictive Modeling. Computer Applications in Archaeology Conference. Vienna, Austria.
Whitley, T. G. (2004a). "Spatial Variables as Proxies for Modelling Cognition and Decision-Making in Archaeological Settings: A Theoretical Perspective." Internet Archaeology(16).
Whitley, T. G. (2004b). Re-thinking Accuracy and Precision in Predictive Modeling. Computer Applications in Archaeology Conference. Prato, Italy.
Whitley, T. G. (2006). Predictive Modeling in a Homogrneous Environment: An Example from the Charleston naval weapons Station, South Carolina. GIS and Archaeological Site Location Modeling. M. W. M. a. K. L. Wescott. New York, Taylor & Francis 355-392.
Wilcox, W. (2009). Archaeological Predictive Modelling in East Anglia and Norflok. Computer Applications to Archaeology 2009. Williamsburg, Virginia, USA.
Willey, G. (1953). Prehistoric Settlement Patterns in the Viru Valley, Peru. Bureau of American Ethnology Bulletin. Washington D.C., Government Printing Office.
Willey, G. R., Ed. (1956). Prehistoric settlement pattern in the New World. New York, Viking Fund Publications in Anthropology.
Winterhalder, B. (1987). The analysis of hunter–gatherer diets: Stalking and optimal foraging model. Food and evolution. M. Harris and E. Ross. Philadelphia, Temple University Press**:** 311-319.
Wiseman, R. N. (1996). The Land in Between: Archaic and Formative Occupations along the Upper Rio Hondo of Southeastern New Mexico. Archaeology Notes. Santa Fe. **125**.

Witcher, R. E. (1999). GIS and landscapes of perception. Geographical information systems and landscape archaeology. M. Gillings, D. Mattingly and J. van Dalen. Oxford, Oxbow: 13-22.
WRCC, W. R. C. C. (2012). "Western Regional Climate Center." 2012, from http://www.wrcc.dri.edu/.
Zeder, M. A. (1997). The American Archaeologists: A Profile. Los Angeles, Altamira.
Zubrow, E. B. W. (1990). The fantasies of GIS software. Interpreting space: GIS and archaeology. K. M. S. Allen, S. W. Green and E. B. W. Zubrow. London, Taylor & Francis: 184-194.
Zubrow, E. B. W. (1994). Knowledge representation and archaeology: a cognitive example using GIS. The Ancient Mind: Elements of Cognitive Archaeology. E. B. W. a. R. Zubrow, C. Cambridge, Cambridge University Press: 107-118.