

**ARCHAEOLOGICAL PREDICTIVE MODELING FOR BLACK MESA
RANGER DISTRICT, APACHE-SITGREAVES NATIONAL FOREST, AZ**

Matthew H. Beavers. B.S.

A directed research report submitted to the
Department of Geography and Environmental Studies at
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Applied Geography
with a specialization in Geographic Information Science and Technology

December 2022

Committee Members:

Dr. Benjamin Zhan, Chair

Dr. Samantha Krause

COPYRIGHT

by

Matthew H. Beavers

2022

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Matthew H. Beavers, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my advisors for their open and willing assistance on short notice, and their encouragement to see the work completed. I am also deeply appreciative of the Heritage staff at the Apache-Sitgreaves National Forest who helped supply data and answer questions as they arose. Working with them was always a joy and inspired the present research.

And of course, to my family and friends for their unending support.

TABLE OF CONTENTS

ACKNOWLEDGMENTS.....	iv
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
CHAPTER	
1. INTRODUCTION.....	1
2. STUDY AREA.....	3
3. DATA.....	6
1. Archaeological Sites.....	6
2. Elevation.....	8
3. Slope.....	8
4. Aspect.....	8
5. Sun Exposure.....	9
6. Distance to Water.....	9
7. Precipitation Runoff.....	10
8. Local Relief.....	11
9. Land Texture.....	11

10. Land Shelter.....	12
11. Topographic Positional Index.....	12
4. METHODS.....	15
1. Binary Logistic Modeling.....	15
2. Frequency Ratios.....	19
3. Statistical Index Model (W_i).....	20
5. RESULTS.....	23
6. DISCUSSION.....	27
7. CONCLUSION.....	30
REFERENCES.....	32
Data.....	36

LIST OF TABLES

Table	Page
1. North's (North et al. 2003) Proposed Revisions to Haury's Original Chronology and Mills and Herr's (1999:279) Revised Chronology for the Mogollon Rim Region as provided by Dokter et al. 2020.....	4
2. Classification Scheme for Combined Models	17
3. BLR Model Methods and JMP Pro Results.....	18
4. BLR Model Equations.....	18
5. Pixel Counts for all Factors and Site Components.....	21, 22
6. Results of Kvamme's Gain.....	24
7. Verification Data Classification.....	25

LIST OF FIGURES

Figure	Page
1. Apache-Sitgreaves National Forest.....	3
2. Black Mesa Ranger District.....	7
3. Factor Maps.....	13
4. Factor Maps cont'd.....	14
5. Results for the Binary Logistic Regression Model.....	26
6. Results for the Statistical Index Model (Wi).....	26
7. Results for the Frequency Ratio Model.....	27
8. Bailey Ruin highlighted in blue.....	28
9. Site 03010201892.....	29

1. Introduction

While conducting an archaeological survey for the purpose of recording new data, companies and federal organizations alike often employ the use of an Archeological Predictive Model (APM) to aid in choosing their survey methods. Such methods primarily consist of sample survey, which covers a <100% portion of the project area (Plog 1981). This is done by spacing surveyors out in a row (~5-30m, depending on regional practice or local conditions) and walking over places they think are likely to hold cultural remains. APMs help them decide what those likely areas are by defining where sites have a higher probability of occurring (Verhagen and Whitley 2012). This creates the expectation that for n sites that exist, most will be found in High probability, less sites will be found in Low probability, and some number in between will be found in Moderate probability regions. The goal is to include as many sites as possible in higher categories of probability to reduce the chances of overlooking any resources.

APMs account for spatial characteristics associated with the location of archaeological resources in a geographical region and produce discreet areas of site likelihood. Typical factors include elevation, slope, aspect, distance to water, soil type, ground cover, and many others. Often these factors are considered in some way to reflect favorable living conditions, such as flat surface area for convenient travel, easy access to natural resources and good farmland. This research recorded statistical data for archaeological site characteristics on the Black Mesa Ranger District (BMRD) of the Apache-Sitgreaves National Forest (ASNF; the Forest) and used three methods for creating APMs: a Binary Logistic Regression (BLR) Model, Frequency Ratios (FR) and a Statistical Index Model (W_i). The rationale for an updated APM comes from Black

Mesa's unique transitional geography and its density of archaeological sites which, according to the current Zone Archaeologist (2022), includes over 52% of the cultural resources in the ASNF. To the southeast of the BMRD is the Lakeside District, which is pockmarked with cinder cones and contains an underlying layer of rough igneous rock, primarily basalt. East and south of Lakeside are the Springerville and Alpine Districts which become dominated by deep gullies and canyons. To the north of BMRD is a plateau of comparatively flat grassland with occasional sandstone or limestone outcroppings. The BMRD itself contains sandy soils with alluvial gravels and its surface is mainly striated with gentle washes/drainages. (Kanter 2004)

Multivariate statistics, such as binary logistic regression, have been a traditional method of predictive mapping used since the late 1970s and gained popularity in the late '80s as GIS technology became available (Ebert 2004, Kvamme 1992, Richards 1998, Verhagen and Whitley 2012). In a typical APM, factors are examined for their relationship to site placement and weighted according to their significance. These weights are then used to build a model that attempts to explain site likelihood. With logistic models, each factor can be assumed to be independent and its data to have non-Gaussian distribution. This allows the researcher some flexibility to add or remove factors to the model as they attempt to explain a greater percentage of variability in site location (Diwan 2020, Taliaferro 2021).

Frequency Ratios and SI (W_i) are more recent additions to the archaeological researcher's toolkit. Both have been used for decades to aid in Landslide Susceptibility Mapping (LSM), but have only just begun to be applied to archaeology (Diwan 2020, Nicu et al. 2019, Nsanziyera et al. 2018, Regmi et al. 2014, Zare et al. 2021). They

primarily differ from BLR, which collects discreet variables, by using fuzzy logic to apply a range of values to observations that vary in their geographic isolation.

2. Study Area

The ASNF is in the White Mountains of eastern Arizona bordering New Mexico. Portions of the Forest sit along a geographic region known as the Mogollon Rim, an area of dramatic elevation change overlooking the low-lying southern and central regions of the state. The Black Mesa Ranger District is at the far western end of the Sitgreaves side of the ASNF, where it meets the Tonto National Forest, and encompasses the town of Heber-Overgaard (Figure 1).

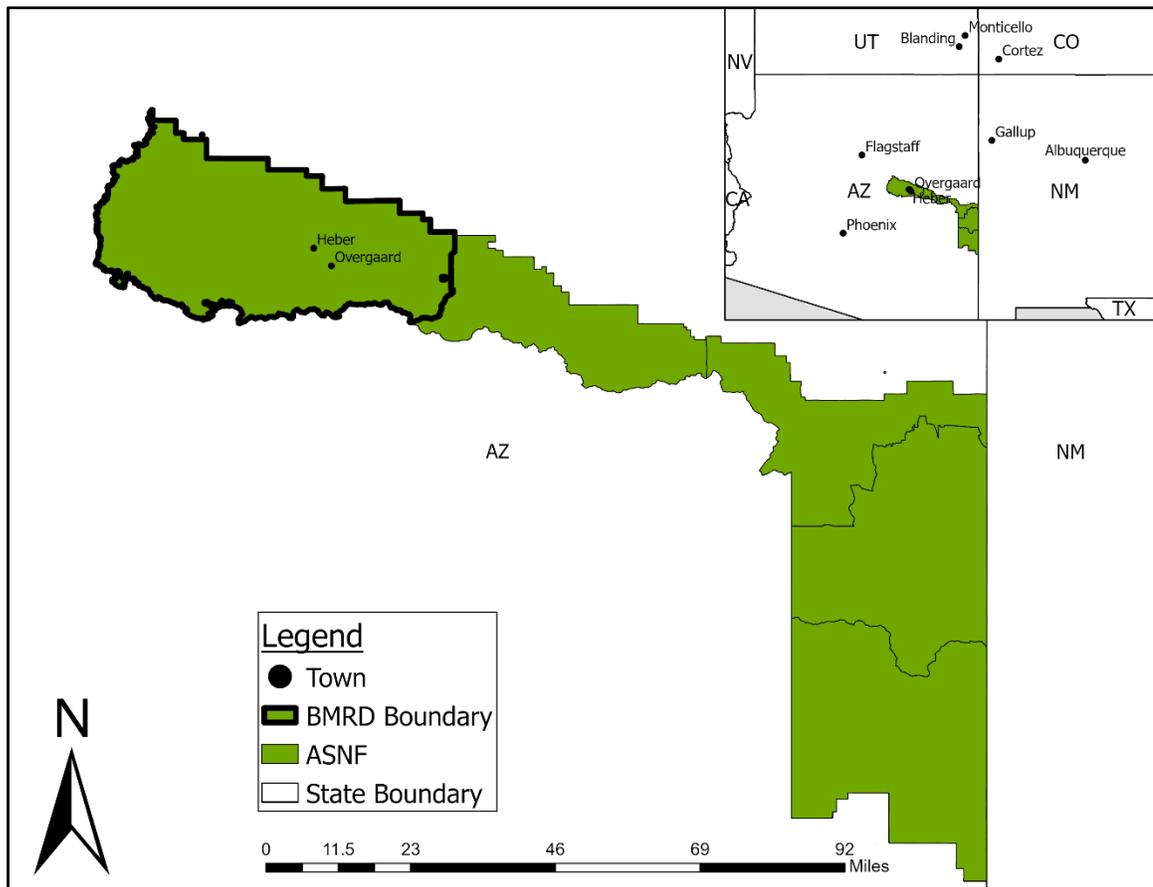


Figure 1. Apache-Sitgreaves National Forest. Inset shows the Forest’s extent in relation to the Four Corners area.

Southern Arizona and portions of New Mexico both have records of human occupation from the Paleoindian (9500 - 6000 BCE) and Archaic Periods (6000 BCE - 200 CE) (Dokter et al 2020, Kantner 2004). Traces of these periods are found on the ASNF as well, but the archeological record in the White Mountains begins in earnest with the Mogollon Tradition (200 CE - 1385 CE). This era observes a shift from early Basketmaker II hunter-gather practices (Hilltop to Cottonwood phases 200-600 CE) to increased reliance on agriculture in the late Basketmaker III Period (Forestdale phase 600 - 800 CE). This transition also includes the introduction of brown ware ceramics, kivas and a progression in architecture for pit/slabhouses. (Dokter et al 2020, Kantner 2004)

Table 1. North’s (North et al. 2003) Proposed Revisions to Haury’s Original Chronology and Mills and Herr’s (1999:279) Revised Chronology for the Mogollon Rim Region as provided by Dokter et al. 2020.

Pecos Classification	Phase	Haury's Chronology (years AD)	Mills and Herr's Revised Chronology (years AD)	North's Rodeo-Chediski Revised Chronology (years AD)
Basketmaker II	Hilltop	200-400	200-400	200-400
Basketmaker II/III	Cottonwood	400-600	400-600	400-600
Basketmaker III/ Pueblo I	Forestdale	600-800	600-800	600-800
Pueblo I	Corduroy	800-900	800-900	800-900
Pueblo II	Dry Valley	900-1000	900-1080	900-1030
Pueblo II/III	Carrizo	1000-1100	Early: 1080-1150 Late: 1150-1200	Early: 1030-1150 Late: 1150-1200
Pueblo III	Linden	1100-1200	1200-1275	1200-1275
Pueblo III/IV	Pinedale	1200-1300	1275-1325	1275-1325
Pueblo IV	Canyon Creek	1300-1400 or 1450	1325-1385	1325-1385

The Pueblo Periods (I-IV), beginning with the Dry Valley Phase (900 - 1000 CE), are characterized by a steady increase in population through immigration, as evident in the influx of varied designs for black-on-white ceramics, the construction of great kivas, and small multi-roomed masonry structures. This continues until the Pinedale Phase (1275 - 1325 CE) when large, 100+ room pueblos are built to accommodate large populations. Well known pueblos include Pinedale Ruin and Bailey Ruin, both of which had over 200 masonry rooms constructed around plazas (Kantner 2004). Around 1300 CE, there appears to be a decrease in population due to mass emigration or abandonment, continuing into the Canyon Creek Phase (1325 - 1385 CE) until much of the region is uninhabited. The White Mountains did not remain empty, however, as several Apache tribes moved into the area following its abandonment, raiding villages that remained and even reusing structures left behind by the Pueblo (Herr, North & Wood 2009).

Historic occupation of the White Mountains began with Spanish expeditions in 1539 as they explored the inlands of the New World, although European settlement would remain sparse until the late 1800s. U.S military control of the area would begin after the Gadsen Purchase in 1854, which included all present-day Arizona. This would spur America's efforts to make the land safe for new settlers, and in 1871 Lieutenant Colonel George Crook was sent to force the Apache bands onto reservations. This subjugation was followed by the establishment of several farming communities (including Heber) and, in 1882, the Atlantic and Pacific Railroad. Timber resources quickly became a major business along with sheep and cattle ranching. (Dokter et al. 2020)

3. Data

3.1 Archaeological Sites

Archaeological site location data was obtained in the Fall of 2021 from the District Archaeologist for the Black Mesa Ranger District of the ASNF, and all following work was done using ArcGIS Pro. Site descriptions included with the feature layer first allowed sites to be organized according to their time period. The Select By Attributes tool was used to search for key words contained within the site description field of the feature attribute table to sort sites into Historic, Prehistoric, Multi-component, or Unknown categories. As some site descriptions occasionally surpassed the maximum character limit for the attribute table, it was sometimes difficult to ascertain the exact category to which each site belonged, and in these cases the best choice was made using the prior knowledge of the researcher. Also at the discretion of the researcher was the exclusion of some historical components from prehistoric sites. These would have otherwise been categorized as “Multi-component”, however, in these cases, the historical artifacts present did not appear to represent an actual occupation of the site but rather an incidental visit. These visits were perceived due to site descriptions that included phrases such as “...a couple of cans...” or “...a few bits of glass...”. This would be the modern-day equivalent of leaving a plastic water bottle behind after hiking around in the Forest and did not merit inclusion in the predictive models which primarily aim to identify permanent or semi-permanent land use.

A total of 4,698 sites were included in the geodatabase received from the BMRD District Archaeologist. After exhaustive sorting using key words, 4,088 were categorized

as Prehistoric, 245 as Historic, 40 as Multi-component, and 289 as Unknown due to a lack of site description. Multi-component sites were included in both Prehistoric and Historic site statistics. Thirty-six of the Historic sites were linear trails or roads that make the collection of discreet site statistics impossible, and so were removed from the current research along with all 289 sites of Unknown component.

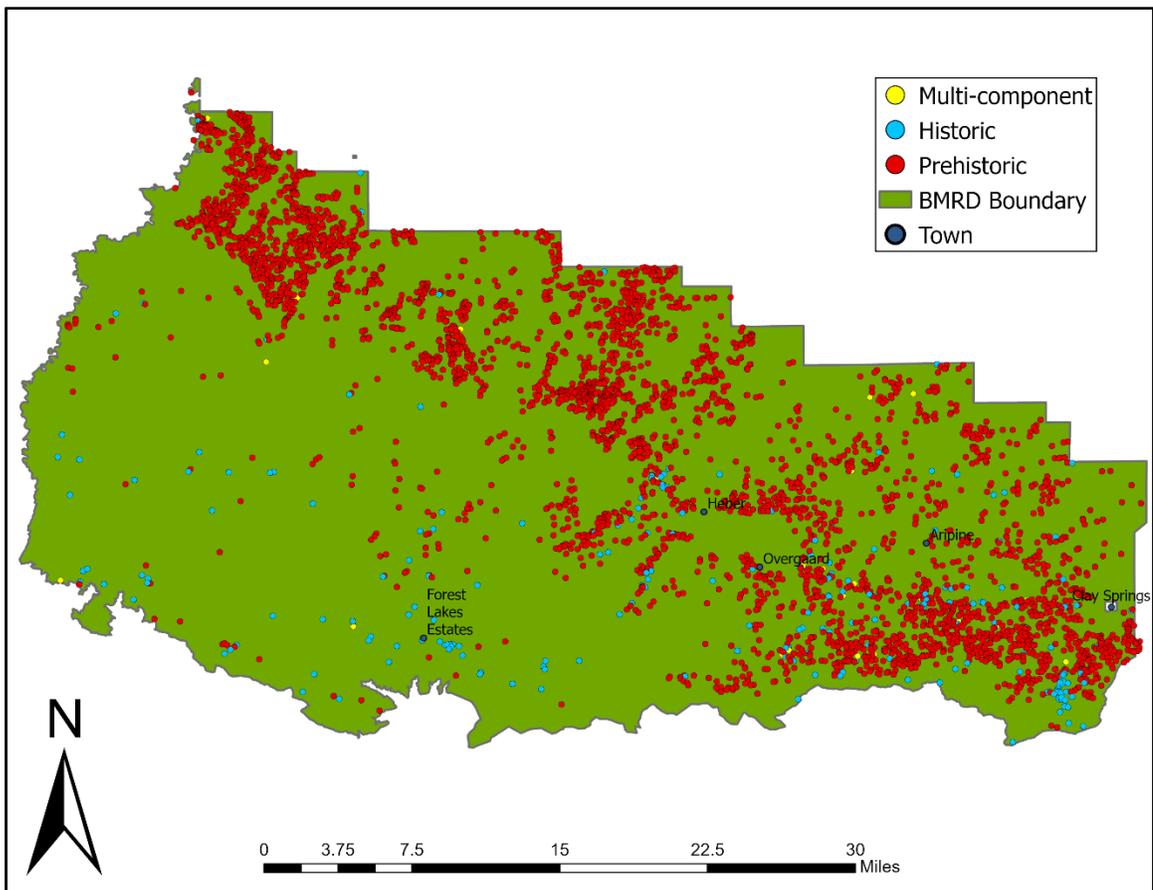


Figure 2. Black Mesa Ranger District.

3.2 Elevation

Two 1/3 arc-second Digital Elevation Models were downloaded from the US Geological Survey National Elevation Dataset (~10m resolution) which displayed the elevation range for all the Sitgreaves side of the ASNF. A Ranger District boundary shapefile from the US Department of Agriculture was used to mask the DEMs after they were made into a single raster using the Mosaic To New Raster Data Management tool in ArcGIS. Elevation on the BMRD ranged between 1739 – 2422 m (5705 – 7946 ft).

3.3 Slope

Surface level is used as an indication of site fitness due to relative ease of use for daily activities (Diwan 2020, Regmi et al. 2014, Taliaferro 2021). Flat surfaces provide more comfortable living conditions and convenient agricultural space. The Slope Spatial Analyst Tool was used to derive a raster layer from the DEM that calculated the angle of slope in degrees (0-90°) for each 10m x 10m cell.

3.4 Aspect

Aspect describes the dominate direction that a ground surface (i.e. slope) faces, measured in cardinal degrees (0-360°). By default, the Aspect Spatial Analyst tool results in 8 classifications consisting of the 4 cardinal directions (North, South, East, West) and the 4 ordinal directions (NE, SE, SW, NW), with each encompassing 45°, and a ninth class representing flat ground with no dominant viewshed. Aspect is used as an indication that site inhabitants were seeking shelter from adverse weather such as wind,

or else using locations exposed to sunlight to their benefit for warmth or agriculture. (Diwan 2020, Nsanziyera et al. 2018, Regmi et al. 2014, Taliaferro 2021).

3.5 Sun Exposure

To gauge the relationship between Aspect and sunlight, the Area Solar Radiation Spatial Analyst tool was used to investigate the amount of sunlight the land surface receives over an entire year. Three factors were measured in the amount of direct radiation in WH/m² (Solar Direct), the amount of total radiation from direct and diffuse light in WH/m² (Solar Total) and the amount of time spent in direct sunlight in hours (Solar Hours). The amount of sunlight received by any given section of land would have been an important factor given its necessity for crop cultivation and the warmth that it provides to both people and plants (Kanter 2004, Salzer 2000, Salzer and Kipfmüller 2005)

3.6 Distance to Water

A basic assumption that can be made is that all humans throughout time have had a need for regular access to water. Given the agricultural lifestyle lead by Mogollon peoples, accessible water would have needed to be close by for crop irrigation during at least part of the year (Kantner 2004). Indeed, over 120 sites from the BMRD dataset contained mention of some form of water control feature, commonly referred to as a “check dam”. These are typically small linear features a few meters in length that are created by stacking local stone across drainages.

To understand this relationship with water, the linear distance from major water sources was calculated for each site point. A hydrology shapefile was downloaded from AZGeo, and its relevant features were selected using the Clip Analyst tool to select features within the BMRD boundary. These features were used in the Euclidean Distance Spatial Analyst tool to create a raster that measured the distance of each cell from the nearest stream feature (Diwan 2020, Nsanziyera et al. 2018).

3.7 Precipitation Runoff

Another important factor for agricultural purposes would be the inclusion of precipitation data. Kantner (2004) notes that the Mogollon Highlands receive more precipitation than lower elevations “with yearly totals commonly [measuring] 350-500 mm”, or 13.8-19.7 in. Historic (1950 – present) annual averages from the National Oceanic and Atmospheric Administration (NOAA) range from 230-760 mm (9-30 in) of precipitation for the Black Mesa Ranger Station in Heber. WorldClim 2.1 provides historical climate data in 30s arc resolution rasters, which is roughly 5 km², however, without finer resolution information it is difficult to perceive meaningful differences in precipitation per site. Therefore, a proxy for measuring the amount of rainfall that would pass through each site as it collected in the myriad drainages that pass through the BMRD was created. Following the example of Taliaferro 2021, Flow Direction and subsequent Flow Accumulation rasters were created, using their respective Spatial Analyst tools, to simulate the movement of water across the topography. An equation was then applied to the Flow Accumulation raster via the Raster Calculator Spatial Analyst tool to estimate the amount of water that would enter the next cell after a small amount (15%) is presumed absorbed by soil. This model was based on research by Toney (2012) and

Schollmeyer (2009), as well as climate data from Salzer and Kipfmueller (2005) from whom an annual mean precipitation of 27.4 cm was calculated. As Taliaferro 2021 focused on the entire ASNF, the current research used the NOAA data from Heber, AZ as a more refined precipitation estimate for just the BMRD, and substituted the annual mean precipitation value for 20.99 cm which resulted in the equation:

$$\textit{Water entering cell} = 0.15 * (0.85 * (\textit{Flow_Acc} * (0.2 * \textit{pp}))) + (0.8 * \textit{pp})$$

3.8 Local Relief

Local Relief describes the amount of variation in elevation by calculating the difference in range within a user defined area. The Focal Statistics Spatial Analyst tool was used to set a radius of 0.5 km around each cell in the DEM and produce a range with smaller results representing less variability and higher results representing greater variability (Diwan 2020, Taliaferro 2021). This factor is used as a representation of how much uniformity in terrain was preferred for site location. A large range might indicate that access to multiple biomes was important to the inhabitants, while a low range might indicate focus on agriculture or ease of pedestrian travel.

3.9 Land Texture

Land Texture also examines surface variation, this time using the Focal Statistics tool to measure the standard deviation within a user defined area, which was set as a 0.5 km radius around each cell. This indicates where sites are located in elevation within a given range (Diwan 2020, Taliaferro 2021). It also provides an indication of how flat the surrounding landscape is.

3.10 Land Shelter

Another derivative of the Focal Statistics tool, the amount of shelter that a site receives from its environment is calculated by finding the mean elevation within a user defined area (0.5 km radius) and subtracting that mean from the local cell elevation via the Raster Calculator tool. This factor indicates the “level of protection from aspects of the environments offered by a location in relation to the surrounding landscape” (Taliaferro 2021).

3.11 Topographic Positional Index

Topographic Positional Index (TPI) is used to measure whether a location is placed high or low relative to the surrounding variation. If a pattern of favoritism for a given range of the TPI were observed, it may suggest a selection of habitation site based on ready access to specific landforms or natural resources (Taliaferro 2021). TPI was calculated using a 0.5 km radius as the local in the Focal Statistics tool and the equation:

$$TPI = \frac{(Elevation - Focal Elevation_{Min})}{(Focal Elevation_{Max} - Focal Elevation_{Min})}$$

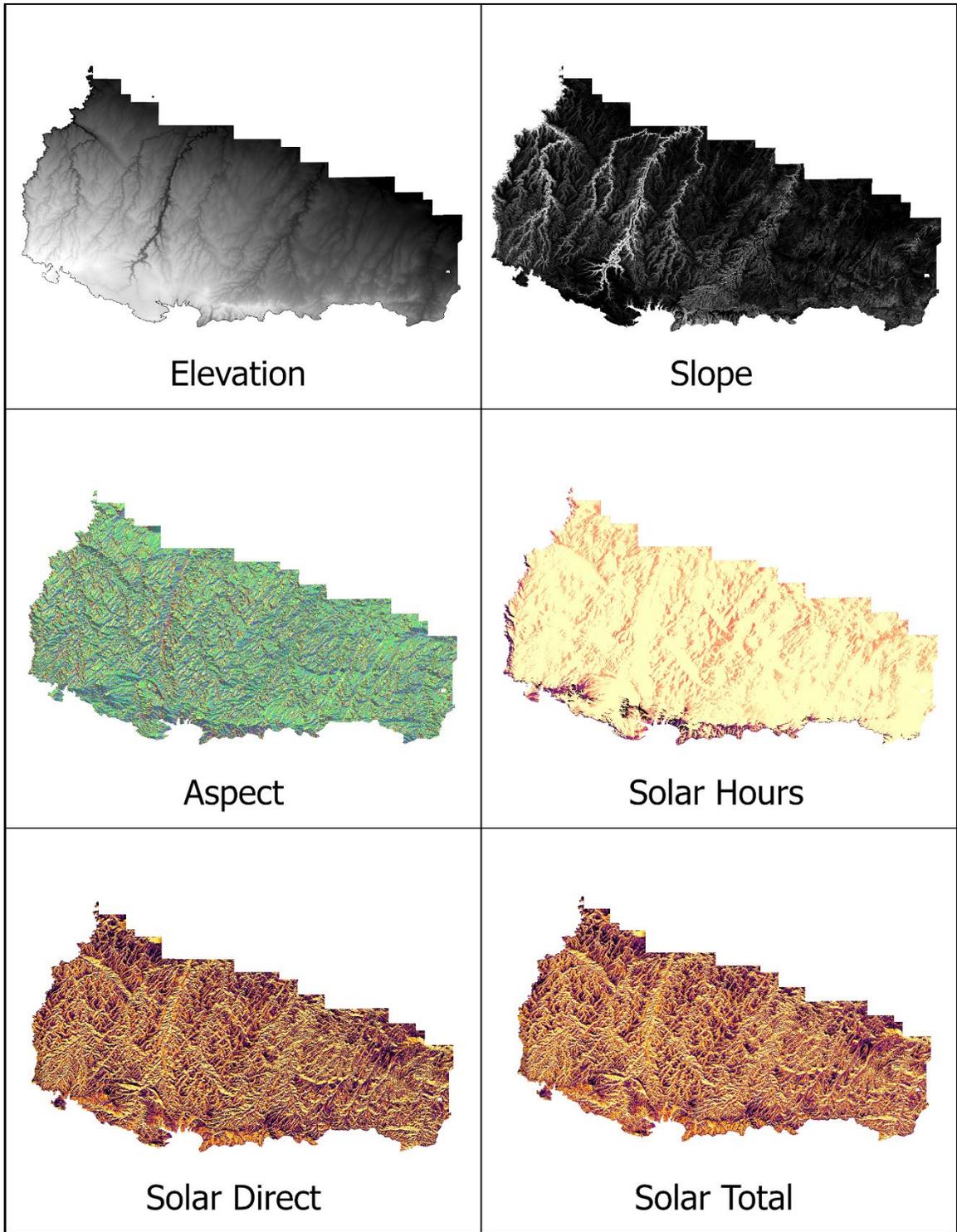


Figure 3. Factor Maps.

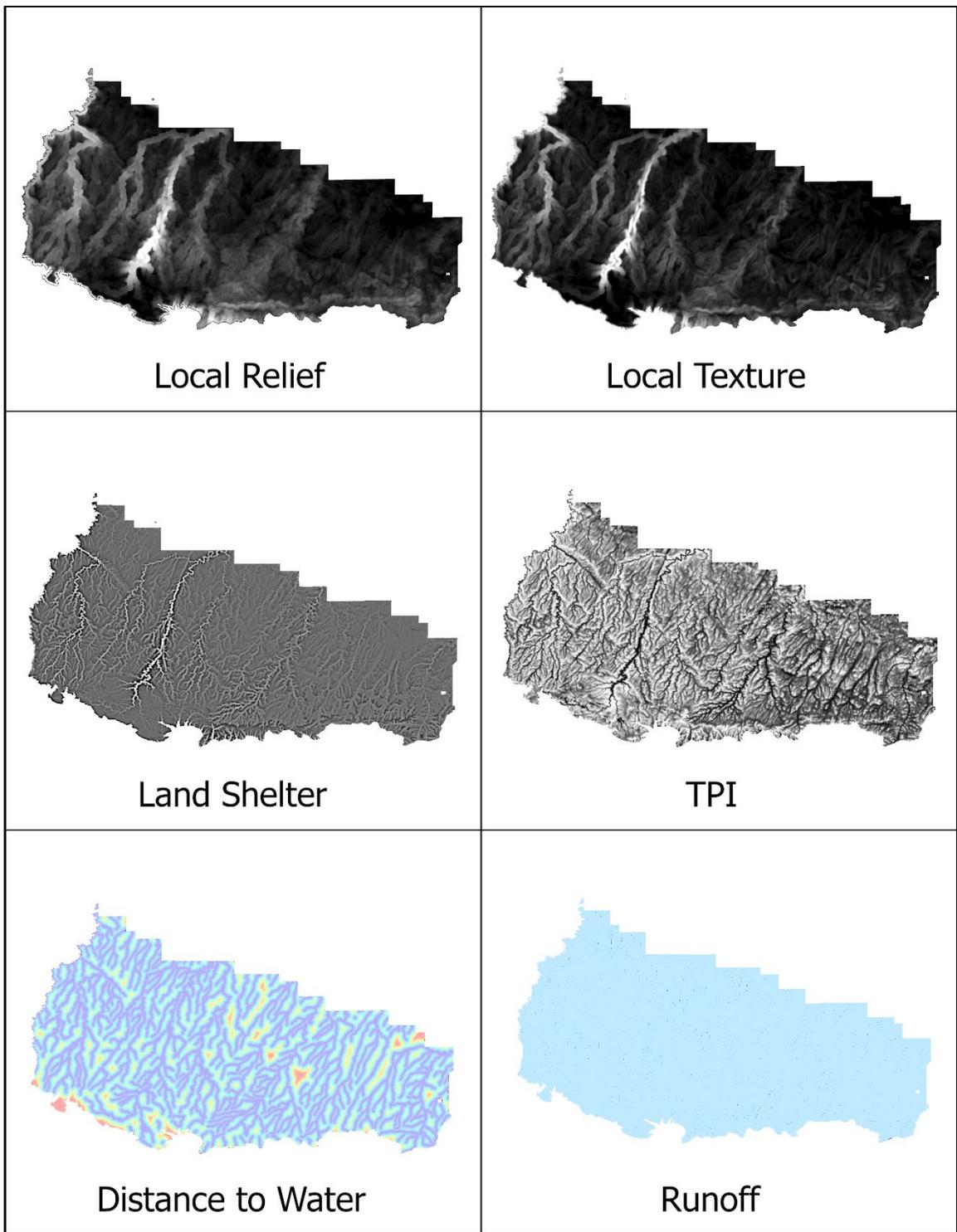


Figure 4. Factor Maps cont'd.

4 Methodology

4.1 Binary Logistic Regression

To facilitate the collection of site statistics, the site boundary polygons received from the District Archaeologist were converted using the Feature to Point Data Management tool to create a centroid that represented the middle of each site. These site points were considered “true sites” but accounted for only one-half of the binary outcome necessary for logistic regression, and so sets of random points were created to act as “false sites”. It was the intent of the researcher to use a portion of the true sites as training data and the remaining portion as verification data, which meant that the population of false sites needed to be equal to the population of the percentage of true sites chosen for training the predictive model. Due to the low number of Multicomponent sites (40), a threshold of 75% was chosen to ensure that a training population of at least 30 true sites was included in each component model. This goal was derived from the Central Limit Theorem which indicates that 30 is a useful sample size for obtaining data that reflects a general population (Urdan 2017, pg. 61). Similar divisions of data were used by Nicu et al. (2019) and Koohpayama et al. (2021) who chose an 80%/20% and 70%/30% split respectively. Thus, the Create Random Points Data Management tool in ArcGIS was used to generate the desired populations of training points that were constrained to true site locations for each component. An equal number of truly random false site points were also constructed and were then merged into unique feature classes for each component.

Once the factor rasters above were created, the Extract Multi Value Spatial Analyst tool was run to add raster values to all site points in each of the three component features classes. The attribute table for each component could then be exported to Excel (Table to Excel Data Management), whereby they could be opened in JMP Pro for analysis. In JMP there is a Fit Model function that allows the user to select the type of regression analysis they want to perform. The column representing true sites was set as a nominal variable and displayed on the y-axis while all remaining factors were considered independent variables along the x axis. Two component models were created: a Prehistoric model containing both prehistoric and multicomponent sites, and a Historic model containing both historic and multicomponent sites. Each component model was run multiple times using both Forwards and Backwards Stepwise methods to perform logistic regression (Tables 2 and 3). Model selection was based on the Omnibus test for significant P-values of $P < 0.05$ and on the amount of variability explained in the Area Under the Curve (AUC). The confusion matrix generated by JMP Pro was also used to ascertain the number of True Positive and True Negative observations that were correctly predicted, as summarized in the % Accuracy Training Sites column of Table 3.

With the selected models, the Raster Calculator tool was run in ArcGIS to apply a probability to each cell in the study area. As the resulting probabilities were not necessarily a perfect range from 0-1, each raster was reclassified into tercile groups so that each group contained roughly 33.3% of total cells. For the Prehistoric model, the lowest range of probabilities were classified as Low (value of 1), the middle range as Moderate (value of 2) and the highest as High (value of 3). For the Historic model, the same method was applied using values of 10, 20 and 30. Raster Calculator was used a

second time to combine the two models into values of 11, 12, 13, 21, 22, 23, 31, 32, and 33 following the example of Taliaferro 2021. These final values represent combinations of Low, Moderate and High (Table 2) and were used for a final classification scheme of Very Low, Low, Moderate, High, and Very High.

Table 2. Classification Scheme for Combined Models. This dictates which combination of values from the Component models receive a final classification of 1-5.

Combined Model Classification Scheme					
10	1	11	L-L	1	Very Low
	2	12	L-M	2	Low
	3	13	L-H	3	Mod
20	1	21	M-L	2	Low
	2	22	M-M	3	Mod
	3	23	M-H	4	High
30	1	31	H-L	3	Mod
	2	32	H-M	4	High
	3	33	H-H	5	Very High

Table 3. BLR Model Methods and JMP Pro Results.

Component Type	Model Number	Model Personality	Stopping Rule	Omnibus	AUC	% Accuracy Training Sites
Prehistoric	1	Full	-	no	0.812	76.3
	2	All Factors with P<0.05	-	yes	0.812	76.4
	3	Stepwise - Forwards	AICc	yes	0.805	75.4
	4	Stepwise - Backwards	AICc	no	0.717	65.6
	5	Stepwise - Forwards	P-value < 0.05	no	0.812	75.9
	6	Stepwise - Backwards	P-value < 0.05	yes	0.718	65.4
Historic	7	Full	-	no	0.748	68.8
	8	All Factors with P<0.05	-	yes	0.745	67.7
	9	Stepwise - Forwards	AICc	yes	0.745	67.7
	10	Stepwise - Backwards	AICc	yes	0.641	58.3
	11	Stepwise - Forwards	P-value < 0.05	yes	0.745	67.7
	12	Stepwise - Backwards	P-value < 0.05	yes	0.641	58.3

Table 4. BLR Model Equations.

Component Type	Model Number	z =
Prehistoric	1	$(-24.8862) + (0.0294 * LocalTexture) + (0.0049 * LocalRelief) + (0.0217 * LandShelter) + (0.0004 * Sun_time) + (-1.2462e-6 * Sun_direct) + (8.6348e-6 * Sun_total) + (0.0355 * Slope) + (0.0104 * Elevation) + (0.0002e-1 * Runoff) + (-0.0001e-1 * Dis_Water) + (1.1616 * TPI) + (0.0645 * Aspect)$
	2	$(-25.1203) + (0.0486 * LocalTexture) + (0.0219 * LandShelter) + (0.0004 * Sun_time) + (-0.0001e-2 * Sun_direct) + (0.0001e-2 * Sun_total) + (0.0351 * Slope) + (0.0105 * Elevation) + (1.1361 * TPI) + (0.064 * Aspect)$
	3	$(-24.6825) + (0.0475 * LocalTexture) + (0.0005 * Sun_time) + (0.0371 * Slope) + (0.0106 * Elevation)$
	4	$(-0.8808) + (-0.0233 * LocalTexture) + (0.0326 * LocalRelief) + (-0.0002 * Sun_time) + (-1.4618e-8 * Sun_direct) + (1.2213e-9 * Sun_total)$
	5	$(-24.3049) + (0.0271 * LocalTexture) + (0.0054 * LocalRelief) + (0.0004 * Sun_time) + (-0.0001e-2 * Sun_direct) + (0.0007e-3 * Sun_total) + (0.0395 * Slope) + (0.0104 * Elevation) + (0.0693 * Aspect)$
	6	$(-0.8158) + (0.0271 * LocalRelief) + (-0.0002 * Sun_time) + (-0.0001e-2 * Sun_direct) + (0.000001 * Sun_total)$
Historic	7	$Sun_time + (-2.957e-11 * Sun_direct) + (1.1064e-10 * Sun_total) + (0.1929 * Slope) + (-0.003 * Elevation) + (-1.0003e-8 * Runoff) + (-1.0003e-8 * Dis_Water) + (4.4430 * TPI) + (0.0466 * Aspect)$
	8	$2.7816 + (0.0387 * LandShelter) + (0.1922 * Slope) + (-0.0029 * Elevation) + (4.6247 * TPI)$
	9	$2.7816 + (0.0387 * LandShelter) + (0.1922 * Slope) + (-0.0029 * Elevation) + (4.6247 * TPI)$
	10	$0.5049 + (0.0102 * LocalRelief) + (-0.0245 * LandShelter)$
	11	$2.7816 + (0.0387 * LandShelter) + (0.1922 * Slope) + (-0.0029 * Elevation) + (4.6247 * TPI)$
	12	$0.5049 + (0.0102 * LocalRelief) + (-0.0245 * LandShelter)$

4.2 Frequency Ratios

Frequency Ratios compare the occurrence of a phenomenon in relation to its non-occurrence (Diwan 2020, Nicu et al. 2019). The formula for FRs is clearly explained by Regmi et al. 2014 as:

$$FR = \frac{\left(\frac{N_{pix}(SX_i)}{\sum_{i=1}^m (SX_i)} \right)}{\left(\frac{N_{pix}(X_j)}{\sum_{j=1}^n N_{pix}(X_j)} \right)}$$

Where:

$N_{pix}(SX_i)$: number of pixels with sites within class i of factor X_i ,

$N_{pix}(X_j)$: number of pixels within factor X_j ,

m : number of classes in the parameter X_i ,

n : number of factors in the study area.

This compares the site occurrence in a class to the sum of site occurrence in the factor, and the factor extent to the sum of all factor extents. The two ratios are then themselves used to create a final ratio, the FR (Table 4). By using the Lookup Spatial Analyst tool in ArcGIS Pro, a new raster can be created for each factor that assigns the appropriate FR to every cell (pixel). The sum of FRs for all factors is obtained through the Raster Calculator and results in the final APM, so that:

$$APM = FR_1 + FR_2 \dots + FR_n$$

Initially, each of the original factors was reclassified into ten classes using the Natural Breaks (Jenks) method with the Reclassify Spatial Analyst tool. A count of the occurrence for each class was then recorded. The Clip Raster Data Management tool used two

feature classes containing Prehistoric/Multi-component and Historic/Multi-component site polygons to isolate the values for archaeological site occurrence. The FR equation could then be calculated and applied to new rasters as described.

4.3 Statistical Index Model (W_i)

As with FR models, the W_i method has only recently transferred to archaeological study from geological research and uses all cells within sites to build its statistics (Diwan 2020, Nsanziyera et al. 2018, Regmi et al. 2014). Whereas FR is a comparison of site occurrence to factor occurrence, W_i uses site density in a class compared to site density in the factor. Each classes' weight value is calculated using a natural logarithmic function of the density ratio which can result in both positive and negative weights, or the estimation of occurrence or non-occurrence respectively (Table 4). W_i values close to zero indicate a weaker estimation than values more distant from zero.

Outlined in Diwan 2020, W_i was calculated for each class following:

$$\ln W_i = \ln \left(\frac{\text{Densclas}}{\text{Densmap}} \right) = \frac{\left(\frac{N_{pix}(S_i)}{N_{pix}(N_i)} \right)}{\left(\frac{\sum N_{pix}(S_i)}{\sum N_{pix}(N_i)} \right)}$$

and the resulting weights were combined into the APM using the Lookup and Raster Calculator tools as with the FR method.

Table 5. Pixel Counts for all Factors and Site Components. FR and W_i for every class.

Factor	Class	Class Range		$N_{pix}(N_i)$	$N_{pix}(S_i)$	$N_{pix}(S_i)$	FR		W_i	
		Min	Max	Class	Prehistoric	Historic	Prehistoric	Historic	Prehistoric	Historic
Aspect	0	Flat		16328	0	0	0.000	0.000	0.000	0.000
	1	337.5°	22.5°	4930245	35478	14925	2.059	1.843	-0.009	-0.119
	2	22.5°	67.5°	4733380	40091	17280	2.327	2.134	0.154	0.068
	3	67.5°	112.5°	4299560	39058	12886	2.267	1.591	0.225	-0.129
	4	112.5°	157.5°	3363902	28924	10736	1.678	1.326	0.170	-0.066
	5	157.5°	202.5°	2119743	16014	9627	0.929	1.189	0.040	0.286
	6	202.5°	247.5°	1848907	10006	8644	0.581	1.067	-0.293	0.315
	7	247.5°	292.5°	2750595	13386	9322	0.777	1.151	-0.400	-0.006
	8	292.5°	337.5°	4456470	24020	13839	1.394	1.709	-0.298	-0.094
Elevation	1	1739.56	1891.92	1548618	20070	204	1.163	0.025	0.580	-3.253
	2	1891.92	1940.03	3174642	26541	763	1.538	0.094	0.142	-2.651
	3	1940.03	1988.15	3252208	32646	1309	1.891	0.161	0.325	-2.136
	4	1988.15	2036.26	4136762	59031	5695	3.420	0.703	0.677	-0.906
	5	2036.26	2087.05	3698727	43237	3111	2.505	0.384	0.477	-1.399
	6	2087.05	2140.51	3163446	17421	12497	1.009	1.542	-0.275	0.148
	7	2140.51	2199.31	2748843	3725	794	0.216	0.098	-1.677	-2.468
	8	2199.31	2263.46	2615357	990	840	0.057	0.104	-2.953	-2.361
	9	2263.46	2322.27	2860922	1762	62011	0.102	7.649	-2.466	1.850
	10	2322.27	2421.16	1351362	1653	10037	0.096	1.238	-1.780	0.779
Distance to Water	1	0.00	152.23	6968815	40709	17914	2.358	2.210	-0.217	-0.281
	2	152.23	334.91	7382455	50672	22154	2.935	2.734	-0.056	-0.127
	3	334.91	502.36	5581798	44118	20717	2.555	2.556	0.085	0.086
	4	502.36	669.82	3899013	31910	18172	1.848	2.242	0.120	0.314
	5	669.82	852.49	2507132	21581	14104	1.250	1.740	0.171	0.502
	6	852.49	1065.61	1307517	11937	3770	0.691	0.465	0.230	-0.167
	7	1065.61	1370.08	611482	5448	377	0.316	0.047	0.205	-1.709
	8	1370.08	1841.99	206455	784	36	0.045	0.004	-0.648	-2.972
	9	1841.99	2648.81	58868	0	0	0.000	0.000	0.000	0.000
	10	2648.81	3881.88	24989	0	0	0.000	0.000	0.000	0.000
Local Relief	1	5.16	31.99	6598978	81222	35722	4.706	4.406	0.529	0.463
	2	31.99	50.36	8699200	76028	49306	4.405	6.082	0.186	0.509
	3	50.36	68.72	5395691	28599	5103	1.657	0.629	-0.314	-1.281
	4	68.72	87.08	3403365	10812	3919	0.626	0.483	-0.825	-1.085
	5	87.08	108.26	1991732	4810	859	0.279	0.106	-1.100	-2.067
	6	108.26	133.68	953419	3231	1219	0.187	0.150	-0.761	-0.980
	7	133.68	164.75	724255	1131	66	0.066	0.008	-1.536	-3.621
	8	164.75	204.30	530976	192	41	0.011	0.005	-2.999	-3.787
	9	204.30	256.55	220444	305	280	0.018	0.035	-1.657	-0.987
	10	256.55	365.30	33287	746	746	0.043	0.092	1.128	1.884
Land Shelter	1	-145.43	-49.28	91564	213	155	0.012	0.019	-1.137	-0.699
	2	-49.28	-29.84	315930	1345	479	0.078	0.059	-0.533	-0.810
	3	-29.84	-17.57	1011629	4970	970	0.288	0.120	-0.390	-1.268
	4	-17.57	-9.39	2676070	17656	5296	1.023	0.653	-0.095	-0.543
	5	-9.39	-2.23	7888778	72705	33557	4.212	4.139	0.240	0.222
	6	-2.23	3.91	9315196	73960	35819	4.285	4.418	0.090	0.121
	7	3.91	13.12	4738770	29180	18554	1.691	2.289	-0.164	0.139
	8	13.12	28.46	1775575	5411	2199	0.314	0.271	-0.867	-1.012
	9	28.46	53.01	545396	1339	232	0.078	0.029	-1.083	-2.080
	10	53.01	115.40	192439	297	0	0.017	0.000	-1.547	0.000
Local Texture	1	0.96	7.12	7863804	93177	41560	5.398	5.127	0.491	0.439
	2	7.12	11.95	9364161	71738	46591	4.156	5.747	0.055	0.379
	3	11.95	17.23	5184822	26892	4672	1.558	0.576	-0.335	-1.330
	4	17.23	22.94	2883193	8432	1806	0.489	0.223	-0.908	-1.693
	5	22.94	29.97	1531249	2932	997	0.170	0.123	-1.332	-1.655
	6	29.97	38.76	675549	2374	734	0.138	0.091	-0.725	-1.143
	7	38.76	48.43	490362	732	208	0.042	0.026	-1.581	-2.083
	8	48.43	59.42	325371	310	204	0.018	0.025	-2.030	-1.693
	9	59.42	72.61	167342	238	238	0.014	0.029	-1.629	-0.873
	10	72.61	113.04	65494	251	251	0.015	0.031	-0.638	0.118

Table 5. continued.

Factor	Class	Class Range		$N_{pix}(N_i)$	$N_{pix}(S_i)$	$N_{pix}(S_i)$	FR		W_i	
		Min	Max	Class	Prehistoric	Historic	Prehistoric	Historic	Prehistoric	Historic
Runoff	1	16.79	444.23	28085650	205138	96096	11.885	11.854	0.007	0.004
	2	444.23	1940.27	300164	1264	921	0.073	0.114	-0.544	-0.105
	3	1940.27	4291.20	86250	304	144	0.018	0.018	-0.722	-0.713
	4	4291.20	7283.28	39188	127	36	0.007	0.004	-0.806	-1.311
	5	7283.28	11130.24	19777	107	24	0.006	0.003	-0.293	-1.032
	6	11130.24	15832.09	9691	40	5	0.002	0.001	-0.564	-1.887
	7	15832.09	21816.25	5249	31	18	0.002	0.002	-0.205	0.007
	8	21816.25	29937.62	3160	11	17	0.001	0.002	-0.734	0.457
	9	29937.62	40837.35	1328	31	0	0.002	0.000	1.169	0.000
	10	40837.35	54515.44	430	23	0	0.001	0.000	1.998	0.000
Solar Hours	1	0.00	510.06	36922	18	1006	0.001	0.124	-2.700	2.079
	2	510.06	1122.13	59101	26	802	0.002	0.099	-2.803	1.382
	3	1122.13	1496.17	96811	89	2180	0.005	0.269	-2.066	1.889
	4	1496.17	1887.22	121515	121	284	0.007	0.035	-1.986	-0.377
	5	1887.22	2346.27	94394	83	602	0.005	0.074	-2.110	0.627
	6	2346.27	2805.32	135914	49	214	0.003	0.026	-3.002	-0.772
	7	2805.32	3196.37	222241	287	518	0.017	0.064	-1.726	-0.379
	8	3196.37	3587.41	558712	7110	3603	0.412	0.444	0.562	0.638
	9	3587.41	4046.47	3985585	42967	14675	2.489	1.810	0.396	0.078
	10	4046.47	4335.50	23239692	156326	73377	9.057	9.052	-0.075	-0.076
Solar Direct	1	0.00	350378.77	4904637	39283	15841	2.276	1.954	0.099	-0.053
	2	350378.77	644696.95	3600768	24934	13099	1.445	1.616	-0.046	0.066
	3	644696.95	967045.42	2958205	19688	10293	1.141	1.270	-0.086	0.021
	4	967045.42	1303409.04	2685958	19716	9362	1.142	1.155	0.012	0.023
	5	1303409.04	1625757.51	2647906	19272	9309	1.117	1.148	0.003	0.032
	6	1625757.51	1920075.68	2585014	17808	9386	1.032	1.158	-0.051	0.064
	7	1920075.68	2242424.16	2648946	18867	10447	1.093	1.289	-0.018	0.146
	8	2242424.16	2620833.23	6503367	47508	19524	2.753	2.408	0.007	-0.126
	9	2620833.23	2929166.55	501	0	0	0.000	0.000	0.000	0.000
	10	2929166.55	3573863.50	15585	0	0	0.000	0.000	0.000	0.000
Solar Total	1	7.10	596470.97	2504137	25218	7422	1.461	0.916	0.328	-0.139
	2	596470.97	1043818.86	4153526	31247	14205	1.810	1.752	0.037	0.004
	3	1043818.86	1491166.76	3922220	25043	14035	1.451	1.731	-0.127	0.049
	4	1491166.76	1975793.65	3328820	22866	11848	1.325	1.462	-0.054	0.044
	5	1975793.65	2441781.04	3298793	23006	11699	1.333	1.443	-0.039	0.040
	6	2441781.04	2833210.45	3722950	25014	13818	1.449	1.705	-0.076	0.086
	7	2833210.45	3261918.85	7603869	54680	24234	3.168	2.989	-0.009	-0.067
	8	3261918.85	3765185.23	770	2	0	0.000	0.000	-1.027	0.000
	9	3765185.23	4175254.13	279	0	0	0.000	0.000	0.000	0.000
	10	4175254.13	4753078.50	15523	0	0	0.000	0.000	0.000	0.000
Slope	1	0.00	2.57	10774606	107044	48850	6.212	0.419	0.314	0.285
	2	2.57	5.44	8184562	64635	32824	3.751	0.254	0.084	0.162
	3	5.44	8.87	4025772	22128	11737	1.284	0.037	-0.278	-0.157
	4	8.87	12.59	2286502	7889	2811	0.458	-0.054	-0.744	-1.020
	5	12.59	16.88	1352045	2800	707	0.162	-0.076	-1.254	-1.875
	6	16.88	21.74	783097	1233	265	0.072	-0.081	-1.528	-2.310
	7	21.74	27.47	476422	575	59	0.033	-0.083	-1.794	-3.316
	8	27.47	33.76	359952	394	6	0.023	-0.083	-1.892	-5.321
	9	33.76	41.77	209031	205	0	0.012	-0.083	-2.002	0.000
	10	41.77	72.96	67141	74	0	0.004	-0.083	-1.885	0.000
TPI	1	0.00	0.15	740674	3931	1162	0.228	0.143	-0.313	-0.776
	2	0.15	0.25	1593532	11152	6419	0.646	0.791	-0.036	0.167
	3	0.25	0.34	2279318	17517	6989	1.015	0.862	0.058	-0.106
	4	0.34	0.43	2877416	21935	8716	1.270	1.075	0.049	-0.118
	5	0.43	0.51	3360284	26273	10489	1.522	1.293	0.075	-0.088
	6	0.51	0.60	3737739	27188	10646	1.575	1.313	0.003	-0.180
	7	0.60	0.69	3935532	28511	12792	1.651	1.577	-0.001	-0.047
	8	0.69	0.78	4029977	27098	16153	1.570	1.991	-0.076	0.162
	9	0.78	0.87	3590387	24053	13953	1.393	1.720	-0.080	0.131
	10	0.87	1.00	2406028	19487	9995	1.129	1.232	0.110	0.198

5 Results

Applying the three model methods to the BRMD resulted in a series of susceptibility maps that highlighted Low, Moderate and High probability for both Prehistoric and Historic components. The combined models display Very Low, Low, Moderate, High and Very High probability ranges that suggest the least and most suitable geographic regions for discovering archaeological resources. Before these maps can be used, however, some semblance of their accuracy needs to be known. While this is a somewhat problematic subject (Grøn 2018, Yaworsky et al. 2020) we first look at the amount of variation that is predicted to be explained by our model. This is done using Receiver Operating Characteristic curves (ROC). The ROC plots a model's ability to accurately predict a true positive or a true negative site, with the percentage of true values described in the AUC (Nicu et al. 2019, Regmi et al. 2014, Zare et al. 2021). While this is a natural product of logistic regression in JMP Pro, the process had to be done again for the FR and W_i models. The Extract Multi Values tool was used to retrieve the predictor value from both models for each of the two components, and these values were exported to JMP Pro. Binary logistic regression analysis was run using the predictor value as the independent variable and an ROC generated for every component model. The resulting AUC percentages were then recorded in Table 6.

Use of Kvamme's Gain as a measure of model significance is another common choice amongst researchers (Diwan 2020, Nicu et al. 2019, Nsanziyera et al. 2018, Tiarafarro 2021) and is simple to calculate. Gain is expressed with the equation:

$$G = 1 - \left(\frac{\%PS}{\%GS} \right)$$

Where Gain (G) is a value between 0 and 1, derived from the ratio of highest probability area (%PS) to percent of sites in that area (%GS) (Diwan 2020). Gain statistics for each model can also be found in Table 6. According to Diwan 2020 and others (Nicu et al. 2019, Nsanziyera et al. 2018, Tiarafarro 2021) a successful Gain value of 0.5 or higher is preferred to verify model reliability. Several of the current models come close with BLR consistently scoring 0.45 or higher, but all three FR models, along with the Historic W_i model, fall short of 0.4. The combined W_i model scores the best at 0.56 which suggest that it has better predictive power.

Table 6. Results of Kvamme’s Gain. All true-site points included in Gain results. The AUC column describes variation explained in the training data. Due to classification methods, cell specific predictor values were not included in the final models.

Kvamme's Gain				AUC
	PS	GS	Gain	
	% Area Highest Probability	% Sites in Area	1 - (PS/GS)	
BLR_P	33.35	60.47	0.45	0.81
BLR_H	32.99	61.97	0.47	0.745
FR_P	32.55	48.55	0.33	0.7
FR_H	32.79	42.46	0.23	0.529
W_i _P	33.07	63.94	0.48	0.799
W_i _H	32.05	45.26	0.29	0.577
BLR_C	14.42	26.72	0.46	
FR_C	24.70	33.85	0.27	
W_i _C	16.45	37.29	0.56	

A summation of model efficacy by exploring the percentage of verification sites ($s = 0.25n$) captured by each probability class can be seen in Table 7. For all three methods there is a greater than 45% inclusion of verification sites in the High probability class for both components. Prehistoric site models for BLR and W_i both exceed 60%, along with the Historic site BLR model. The combined models all exceeded 50% inclusion of sites

in the High and Very High range with 54.5%, 64.8%, and 53% for BLR, Wi, and FR respectively.

Table 7. Verification Data Classification.

Model Coverage of Verification Data - 25% Total Component Population											
Component	Binary Logistic Regression - Components										
	Low	%-Model	Moderate	%-Model	High	%-Model				Total	
Prehistoric	29	2.81	378	36.63	625	60.56				1032	
Historic	8	11.27	19	26.76	44	61.97				71	
Statistical Index Model (Wi) - Components											
Component	Low	%-Model	Moderate	%-Model	High	%-Model				Total	
	Prehistoric	71	6.88	319	30.91	642	62.21				1032
Historic	17	23.94	22	30.99	32	45.07				71	
Frequency Ratios - Components											
Component	Low	%-Model	Moderate	%-Model	High	%-Model				Total	
	Prehistoric	145	14.05	389	37.69	498	48.26				1032
Historic	19	26.76	20	28.17	32	45.07				71	
Binary Logistic Regression - Combined											
Component	Very Low	%-Model	Low	%-Model	Moderate	%-Model	High	%-Model	Very High	%-Model	Total
	Prehistoric	10		119		273		349		271	1022
Historic	17		19		19		4		2	61	
Multi-component	4	2.84	3	12.90	1	26.79	2	32.50	0	24.97	10
Statistical Index Model (Wi) - Combined											
Component	Very Low	%-Model	Low	%-Model	Moderate	%-Model	High	%-Model	Very High	%-Model	Total
	Prehistoric	66		124		165		285		382	1022
Historic	11		8		6		16		20	61	
Multi-component	0	7.04	1	12.17	4	16.01	3	27.82	2	36.96	10
Frequency Ratios - Combined											
Component	Very Low	%-Model	Low	%-Model	Moderate	%-Model	High	%-Model	Very High	%-Model	Total
	Prehistoric	125		116		242		207		332	1022
Historic	13		4		9		13		22	61	
Multi-component	3	12.90	0	10.98	2	23.15	1	20.22	4	32.75	10

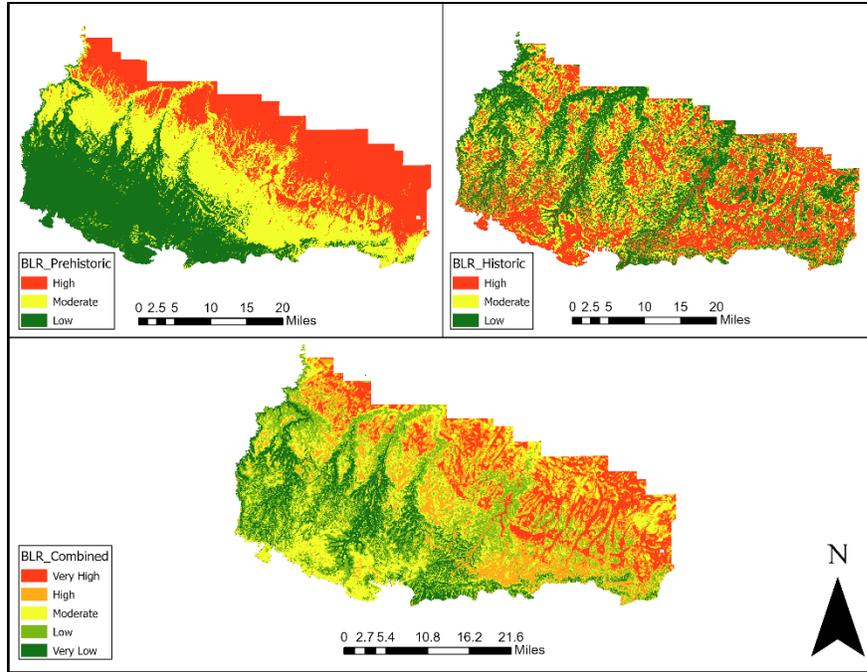


Figure 5. Results for the Binary Logistic Regression Model.

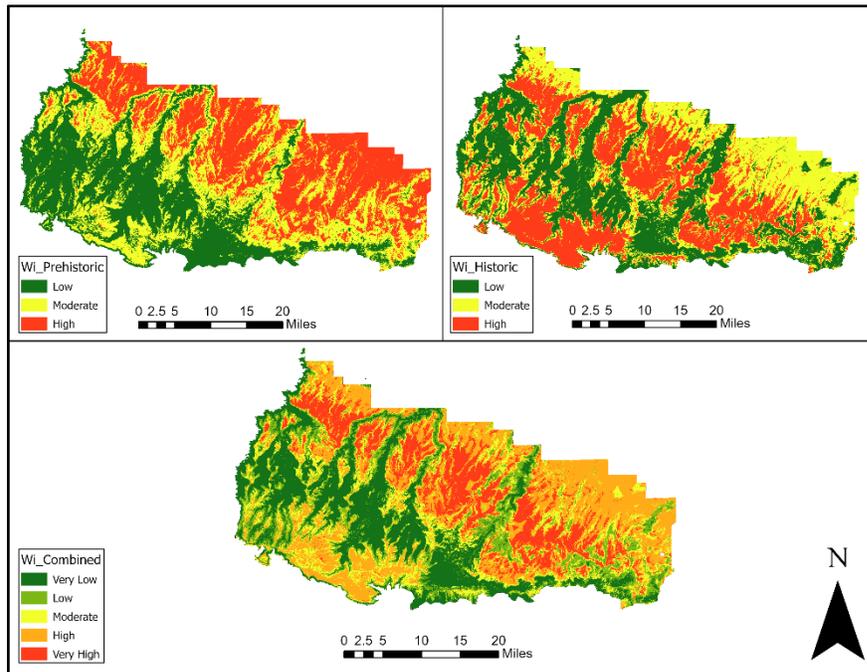


Figure 6. Results for the Statistical Index Model (W_i).

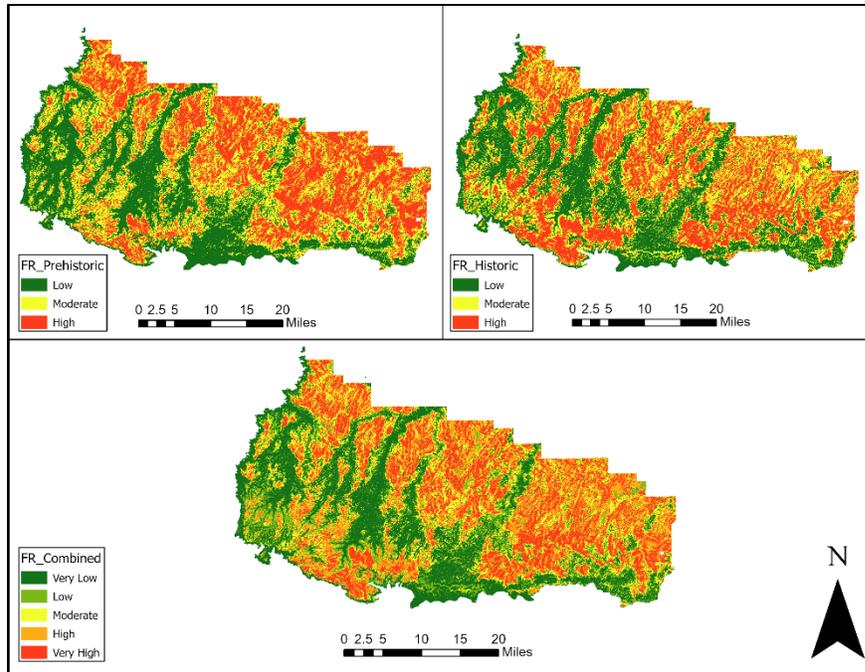


Figure 7. Results for the Frequency Ratio Model.

6 Discussion

While the Gain values for the final models would suggest that the W_i model is the highest performing method, there is a higher concentration of sites in the Moderate – (Very) High categories for both BLR component models and the combined BLR model. Additionally, the three watershed models from Taliaferro 2021 that had low Gain values (Chevelon Canyon/Upper Salt – Historic, Silver Creek – Historic, Silver Creek – Prehistoric) all fall fully or partially within the BMRD. This suggests that the predictive ability of the current BLR model marginally outperforms both the old BLR model from Taliaferro 2022, and the ratio methods in the current research. Compared to results generated by either FR or W_i , multivariate statistics appear to have captured sites of higher interest as well. For example, the 200+ room pueblo of Bailey Ruin (Figure 8) is almost entirely classified as Low/Very Low probability using ratio methods, but is mostly

classified as High probability with BLR. Kivas, ceremonial structures built to accommodate large groups, are an important feature found in only a few sites, including Bailey Ruin. Of only 19 total kiva sites on the BMRD, 6 were classified as less than High probability by the FR model, 2 by the W_i model, and 4 by the BLR model. However, only the BLR model included every single kiva in at least Moderate probability (Figure 9).

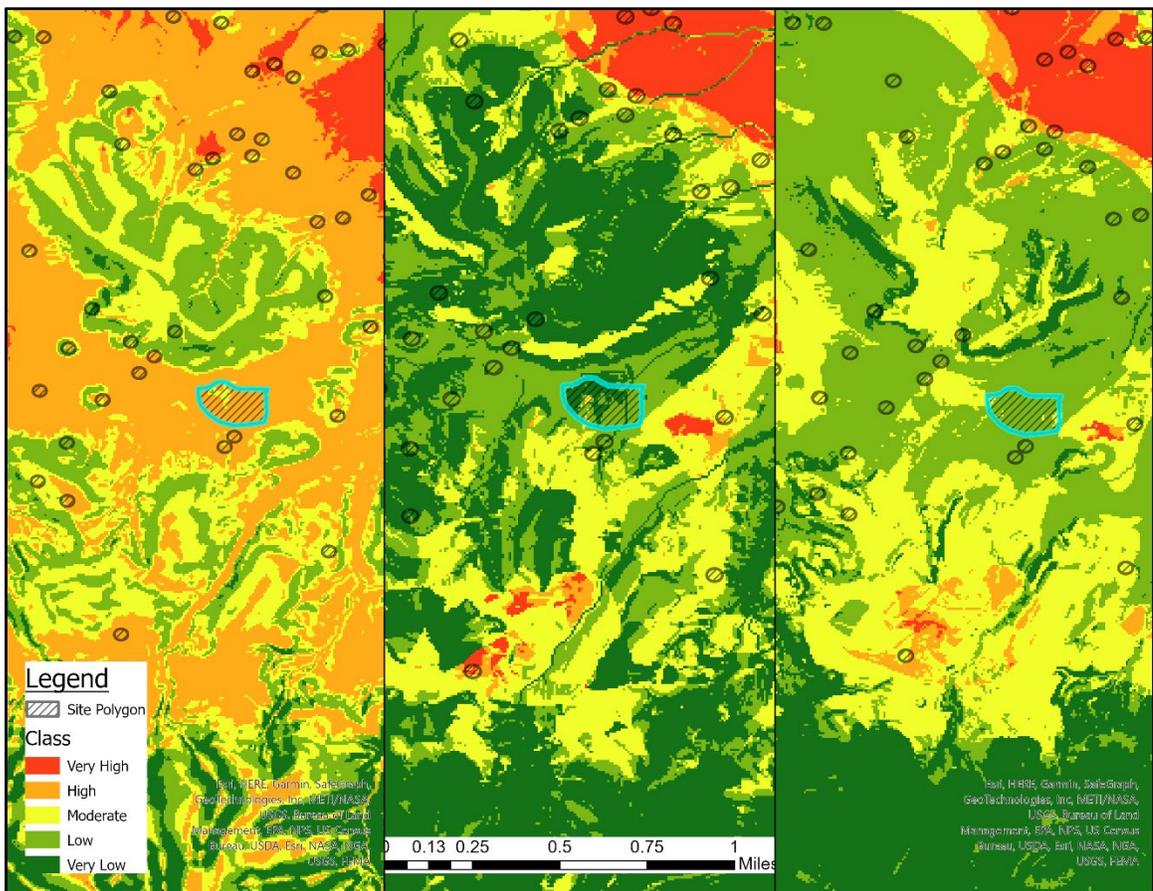


Figure 8. Bailey Ruin highlighted in blue. From left to right: BLR, FR, W_i .

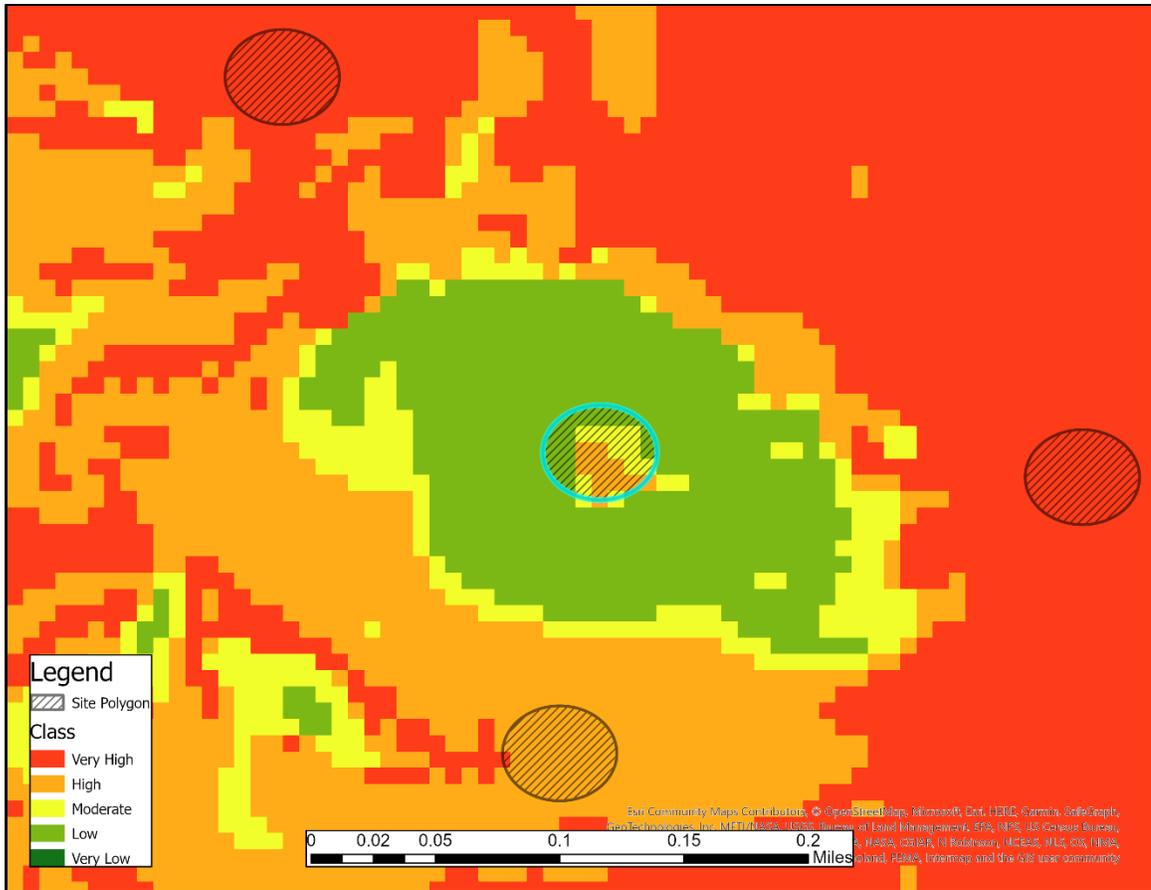


Figure 9. Site 03010201892. Centered, kiva site 1892 includes high classification with the BLR model, despite its surroundings. The FR and W_i models capture this important site as only Low-Moderate.

The results of this paper could perhaps be improved upon by incorporating further modelling methods such as species distribution (Franklin et al. 2014) to predict the availability of game animals in relation to habitation sites, Point Process Models (Davis et al. 2020) that create probability rasters around observation clustering, and Geographically Weighted Regression (GWR) in areas of low survey coverage (Shen et al. 2022). GWR and similar concepts such as Maximum Entropy (Yarwosky 2020) are intriguing modeling techniques that attempt to overcome the absence of an observation by weighting factors to proximity with known observations. The predictive power of that factor is therefore lessened as distance from known observations increases. This would be

especially useful in context of the BMRD where sandy soils are prone to shift over time, occasionally revealing sites in areas previously surveyed and thought to be barren of cultural remains. Such instability causes the reliability of past survey to falter with age.

Yaworsky et al. 2020 does raise some concerns in regard to the reliability of predictive models, namely that small observation populations, time scale, and the inductive nature of many APMs cause too much uncertainty. This dataset contains a large population of site data ($n = 4373$) with sites that have a fairly constrained time period of 200 – 1400 AD, though this could be improved by creating sub-models based on site features. 1013 prehistoric site descriptions mention either masonry architecture or black-on-white ceramics which are both indicators of habitation after 800 AD. It was the original intent of this research to include models for individual phases of occupation. However, the categorization of sites into specific periods of time is complicated by multi-phase occupations that can make old sites appear more recent, poor or inconsistent site descriptions, and the ephemeral nature of semi-subterranean features (pithouses, kivas, etc.) that can cause them to go unidentified.

7 Conclusions

The effectiveness of any environmentally based APM is influenced by the geography of its study region and the amount of known archaeological site information available to the researcher. Although binary logistic regression appears to have been a more successful method than Frequency Ratios or the Statistical Index method in this context, it does not necessarily indicate that these methods are inferior in all situations. Additionally, it is an unfortunate reality that no single model is likely

to encapsulate the entirety of a region's cultural resources with perfect accuracy as there may be social conditions beyond the physical landscape that influence site location.

As demonstrated in this work, the techniques available to those in the GIS industry and in Cultural Resource Management are continuing to evolve along with our understanding of their shortcomings. It is the purpose of the research presented here to add to a growing body of comparison between common techniques to further the understanding of their abilities. Refinement of these methods offers valuable day to day use for the protection of unique, unrenowable resources like our archeological and historical sites, and for the efficiency of other conservation and land management practices.

REFERENCES:

Brandt, R., Groenewoudt, B. J., & Kvamme, K. L. (1992). *An Experiment in Archaeological Site Location: Modeling in the Netherlands using GIS Techniques*. 24(2).

Caracausi, S., Berruti, G. L. F., Daffara, S., Bertè, D., & Rubat Borel, F. (2018). Use of a GIS predictive model for the identification of high altitude prehistoric human frequentations. Results of the Sessera valley project (Piedmont, Italy). *Quaternary International*, 490, 10–20. <https://doi.org/10.1016/j.quaint.2018.05.038>

Davis, D. S., DiNapoli, R. J., & Douglass, K. (2020). Integrating Point Process Models, Evolutionary Ecology and Traditional Knowledge Improves Landscape Archaeology—A Case from Southwest Madagascar. *Geosciences*, 10(8), 287. <https://doi.org/10.3390/geosciences10080287>

Diwan, A. G. (2020). *GIS-BASED COMPARATIVE ARCHAEOLOGICAL PREDICTIVE MODELS: A FIRST APPLICATION TO IRON AGE SITES IN THE BEKAA (LEBANON)*. <https://doi.org/10.5281/ZENODO.3819601>

Dockter, A., Mead, K., & Tenen, L. B. (2020). *A Cultural Resources Survey. Of 8,446 Acres of the Black Mesa and Lakeside Ranger Districts for the Rodeo-Chediski Prescribed Burn Phase 4, Apache-Sitgreaves National Forest, Navajo County, Arizona*. Harris Environmental Group, Inc.

Duke, D., & King, J. (2014). A GIS model for predicting wetland habitat in the Great Basin at the Pleistocene–Holocene transition and implications for Paleoindian archaeology. *Journal of Archaeological Science*, 49, 276–291. <https://doi.org/10.1016/j.jas.2014.05.012>

Ebert, D. (2004). Applications of Archaeological GIS. *Canadian Journal of Archaeology*, 28(2), 319–341.

Franklin, J., Potts, A. J., Fisher, E. C., Cowling, R. M., & Marean, C. W. (2015). Paleodistribution modeling in archaeology and paleoanthropology. *Quaternary Science Reviews*, 110, 1–14. <https://doi.org/10.1016/j.quascirev.2014.12.015>

- Goings, C. A. (2003). A Predictive Model for Lithic Resources in Iowa. *Plains Anthropologist*, 48(184), 53–67.
- Grøn, O. (2018). Some problems with modelling the positions of prehistoric hunter-gatherer settlements on the basis of landscape topography. *Journal of Archaeological Science: Reports*, 20, 192–199. <https://doi.org/10.1016/j.jasrep.2018.04.034>
- Hamilton, S. (2000). Archaeological Predictive Modelling in the Boreal Forest: No Easy Answers. *Canadian Journal of Archaeology*, 24(1), 41–76.
- Heilen, M., Leckman, P. O., Byrd, A., Homburg, J. A., & Heckman, R. A. (2013). *Archaeological Sensitivity Modeling in Southern New Mexico: Automated Tools and Models for Planning and Management. Report prepared for U.S. Bureau of Land Management, by Statistical Research, Inc., Albuquerque, New Mexico. SRI Technical Report 11-26.* Statistical Research, Inc.
- Herr, S., North, C., & Wood, S. J. (2009). Scouting for Apache Archaeology in the Sub-Mogollon Rim Region. *Kiva*, 75(1), 35–62.
- Kantner, J. (2004). *Ancient Puebloan Southwest.* Cambridge University Press.
- Koohpayma, J., Makki, M., Lentschke, J., & AlaviPanah, S. K. (2021). Predicting potential locations of ancient settlements using GIS and Weights-Of-Evidence method (case study: North-East of Iran). *Journal of Archaeological Science: Reports*, 40, 103229. <https://doi.org/10.1016/j.jasrep.2021.103229>
- Kvamme, K. L. (1992). A Predictive Site Location Model on the High Plains: An Example with an Independent Test. *Plains Anthropologist*, 37(138), 19–40.
- Layzell, A. L., & Mandel, R. D. (2019). Using soil survey data as a predictive tool for locating deeply buried archaeological deposits in stream valleys of the Midwest, United States. *Geoarchaeology*, 34(1), 80–99. <https://doi.org/10.1002/gea.21707>
- Leloch, M., Kot, M., Pavlenok, G., Szymczak, K., Khudjanazarov, M., & Pavlenok, K. (2022). Tracing the Palaeolithic settlement patterns in the Western Tian Shan piedmont: An example of predictive GIS modelling use. *Journal of Quaternary Science*, 37(3), 527–542. <https://doi.org/10.1002/jqs.3393>

Magnini, L., & Bettineschi, C. (2021). Object-Based Predictive Modeling (OBPM) for Archaeology: Finding Control Places in Mountainous Environments. *Remote Sensing*, 13(6), 1197. <https://doi.org/10.3390/rs13061197>

Malaperdas, G., & Zacharias, N. (2019). The habitation Model Trend Calculation (MTC): A new effective tool for predictive modeling in archaeology. *Geo-Spatial Information Science*, 22(4), 314–331. <https://doi.org/10.1080/10095020.2019.1634320>

Mertel, A., Ondrejka, P., & Šabatová, K. (2018). Spatial predictive modeling of prehistoric sites in the Bohemian-Moravian Highlands based on graph similarity analysis. *Open Geosciences*, 10(1), 261–274. <https://doi.org/10.1515/geo-2018-0020>

Mills, B. J., & Herr, S. A. (n.d.). *Living on the Edge of the Rim: Excavations and Analysis of the Silver Creek Archaeological Research Project, 1993-1998* (Vol. 1). Arizona State Museum.

Nicu, I., Mișu-Pintilie, A., & Williamson, J. (2019). GIS-Based and Statistical Approaches in Archaeological Predictive Modelling (NE Romania). *Sustainability*, 11(21), 5969. <https://doi.org/10.3390/su11215969>

North, C., Foster, M. S., & Senior, L. (2003). *Archaeology at the Edge of the Rim: An Area Heritage Resource Survey of Portions of the RodeoChediski Burn Area, Black Mesa and Lakeside Ranger Districts, Apache-Sitgreaves National Forests, Navajo County, Arizona*. (Cultural Resources Report No.2003-56). SWCA.

Nsanziyera, A. F., Lechgar, H., Fal, S., Maanan, M., Saddiqi, O., Oujaa, A., & Rhinane, H. (2018). Remote-sensing data-based Archaeological Predictive Model (APM) for archaeological site mapping in desert area, South Morocco. *Comptes Rendus Geoscience*, 350(6), 319–330. <https://doi.org/10.1016/j.crte.2018.06.010>

Ozdemir, A., & Altural, T. (2013). A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. *Journal of Asian Earth Sciences*, 64, 180–197. <https://doi.org/10.1016/j.jseaes.2012.12.014>

Ozsahin, E., & Ozdes, M. (2022). Agricultural land suitability assessment for agricultural productivity based on GIS modeling and multi-criteria decision analysis: The case of

Tekirdağ province. *Environmental Monitoring and Assessment*, 194(1), 41.
<https://doi.org/10.1007/s10661-021-09663-1>

Plog, F. (1981). *Managing Archeology: A Background Document for Cultural Resource Management on the Apache-Sitgreaves National Forests, Arizona*. USDA Forest Service, Southwestern Region.

Regmi, A. D., Devkota, K. C., Yoshida, K., Pradhan, B., Pourghasemi, H. R., Kumamoto, T., & Akgun, A. (2014). Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arabian Journal of Geosciences*, 7(2), 725–742.
<https://doi.org/10.1007/s12517-012-0807-z>

Richards, J. D. (1998). Recent Trends in Computer Applications in Archaeology. *Journal of Archaeological Research*, 6(4), 331–382.

Salzer, M. W. (2000). Temperature Variability and the Northern Anasazi: Possible Implications for Regional Abandonment. *Kiva*, 65(4), 295–318.

Salzer, M. W., & Kipfmueller, K. F. (2005). Reconstructed Temperature And Precipitation On A Millennial Timescale From Tree-Rings In The Southern Colorado Plateau, U.S.A. *Climatic Change*, 70(3), 465–487. <https://doi.org/10.1007/s10584-005-5922-3>

Shen, Y., de Hoogh, K., Schmitz, O., Clinton, N., Tuxen-Bettman, K., Brandt, J., Christensen, J. H., Frohn, L. M., Geels, C., Karssenber, D., Vermeulen, R., & Hoek, G. (2022). Europe-wide air pollution modeling from 2000 to 2019 using geographically weighted regression. *Environment International*, 168, 107485.
<https://doi.org/10.1016/j.envint.2022.107485>

Taliaferro, M. (2021). *TMR Sampling Plan*. Unpublished.

Urdan, T. C. (2017). *Statistics in plain English* (Fourth edition). Routledge, Taylor & Francis Group.

Verhagen, P. (2007). Case Studies in Archaeological Predictive Modelling. *Archaeological Studies*.

Verhagen, P., & Whitley, T. G. (2012). Integrating Archaeological Theory and Predictive Modeling: A Live Report from the Scene. *Journal of Archaeological Method and Theory*, 19(1), 49–100.

Yaworsky, P. M., Vernon, K. B., Spangler, J. D., Brewer, S. C., & Coddling, B. F. (2020). Advancing predictive modeling in archaeology: An evaluation of regression and machine learning methods on the Grand Staircase-Escalante National Monument. *PLOS ONE*, 15(10), e0239424. <https://doi.org/10.1371/journal.pone.0239424>

Zare, N., Hosseini, S. A. O., Hafizi, M. K., Najafi, A., Majnounian, B., & Geertsema, M. (2021). A Comparison of an Adaptive Neuro-Fuzzy and Frequency Ratio Model to Landslide-Susceptibility Mapping along Forest Road Networks. *Forests*, 12(8), 1087. <https://doi.org/10.3390/f12081087>

Data:

BMRD Boundary:

<https://www.fs.usda.gov/detail/r3/landmanagement/gis/?cid=stelprdb5201889>

Heber Precipitation: <https://www.weather.gov/wrh/Climate?wfo=fgz>

Hydrology: <https://azgeo-open-data-agic.hub.arcgis.com/datasets/azgeo::streams-ephemeral-and-perennial/explore?location=34.158683%2C-111.971281%2C7.69>

WorldClim 2.1: <https://www.worldclim.org/data/worldclim21.html>