Predictive Modeling for Site Detection Using Remotely Sensed Phenological Data

Scott Detrich Kirk, Amy E. Thompson, and Christopher D. Lippitt

Traditionally, archaeology has been a very tactile discipline. Excavation, and the acquisition of data on the material culture of past people, has been completely reliant on the ability of the archaeologist to get out into the field. In recent years, Geographic Information Systems (GIS) and remote sensing techniques have allowed archaeologists to acquire data without ever having to set foot outside. Through geospatial modeling and analysis, including, in particular, remote sensing, it is now possible to interpret archaeological data on settlement patterns and the environment in ways that had not been possible prior to the widespread availability of high powered personal computers and the geospatial processing programs that go along with them (Chase et

ABSTRACT

This paper examines the potential of remote sensing–derived metrics of vegetation phenology and a Multi-Layer Perceptron neural network to model the most likely locations of large, agglomerated archaeological sites. Focusing on two different environments in central New Mexico, the Galisteo Basin and the Sandia-Manzano Mountain range, this pilot study distinguishes between archaeological sites and their surroundings based on differential growth in vegetation. Using data derived from Landsat Thematic Mapper, a time series of Normalized Difference Vegetation Indices were created to characterize vegetation phenology in the study areas. Distinguishing between archaeological sites and their surroundings, the neural network was trained on a series of known sites to develop an output activation layer indicating the possible locations of other, previously unknown sites. This output activation layer, treated as a site suitability model, was validated using the receiver operating characteristic area under the curve using known sites excluded from the training procedure. Results show promise in large, open areas such as basin environments. While differences in vegetation type have relatively little effect, differences in elevation, or more directly the changes in phenology that go along with them, negatively impact the ability to infer the presence of archaeological sites using this approach.

Este artículo examina la potencial de métricas de teledetección de fenología vegetal y una red neural de Multi-Layer Perceptron para modelar las ubicaciones más probables de sitios arqueológicas grandes y aglomerados. Este estudio preliminar enfoque en dos localidades diferentes en el centro de Nueva México, el Cuenco de Galisteo y la cordillera de Sandia-Manzano y distingue entre sitios arqueológicos y sus entornos basado en crecimiento diferencial en vegetación. Un serie temporal del índice de vegetación diferencial normalizado (NDVI) fue creada de datos derivado de Landsat Thematic Mapper para caracterizar la fenología de las plantas en los áreas de estudio. La red neural distingue entre sitios arqueológicos y sus entornos y fue entrenando en un serie de sitios conocidos para desarrollar una capa de activación de salida que indique las ubicaciones posibles de sitios desconocidos. Tratado por un modelo de idoneidad del sitio, la capa de activación de salida fue validada con sitios conocidos excluidos del proceso de entrenamiento usando el área de operador receptor característico bajo de la curva. Los resultados son prometedores para áreas abiertas tal como cuencos. Diferencias en vegetación tienen relativamente poco efecto. Sin embargo, diferencias en elevación y los cambios concomitantes en fenología afectan negativamente la utilidad de este enfoque para inferir sitios arqueológicos.

Advances in Archaeological Practice 4(1), 2016, pp. 87–101

DOI: 10.7183/2326-3768.4.1.87

al. 2012). While ground truthing, or checking remotely sensed data for accuracy, is often critical to such analyses, the use of satellite imagery in archaeological studies has allowed research to continue in war-torn areas like Mesopotamia, where conventional archaeological work is no longer possible (Hritz 2014).

While references to archaeological data collection in politically unstable locations are perhaps the most extreme examples, the use of geospatial analysis and remotely sensed data does have practical applications for more traditional avenues of research as well. In a more mundane way, modern methods of remote sensing have allowed for the examination of large swaths of land that are impractical for traditional survey (Kennedy and Bishop 2011; Sadr and Rodier 2012). Using tools such as ArcGIS Online and Google Earth, which provide access to high resolution imagery of the earth's surface, as well as data layers characterizing waterways, urban areas, and major transportation routes, remote sensing data and techniques allow archaeologists to discover and analyze sites from a bird's eye view and gather information on settlement patterns on a broad scale (Kennedy and Bishop 2011; Sadr and Rodier 2012). However, the manual interpretation of aerial photography and satellite images for site detection is not the only means archaeologists have to discover new sites via remote sensing. Using multispectral satellite data, archaeologists can obtain information on aspects of the environment that correspond to anthropogenic activity (e.g., differential vegetation patterns due to building materials at or below the soil surface) with data that would not have adequate spatial resolution to use for visual assessment (Agapiou et al. 2013). When used in combination with information about how known archaeological sites manifest in terms of remotely sensed environmental variables (e.g., vegetation vigor, soil moisture, vegetation type), environmental information obtained from remote sensing can be used to model site distribution.

This paper examines the use of remotely sensed data indicative of differential growth in vegetation between archaeological sites and their surrounding areas to build site suitability models. Focusing on the Galisteo Basin and Sandia-Manzano Mountain Range of Central New Mexico, this study relies on the principle that anthropogenic modifications to the environment have a lasting effect. In areas of prior human habitation, buried stonework will have a depressive effect on vegetation growth, whereas ditches and middens will encourage the growth of plants (Bahn and Renfrew 2008; Lasaponara and Masini 2006). Findlow and Confeld (1980) have shown that in New Mexico, in particular, areas where humans once settled show changes in soil color. Using these differences in the growth of vegetation observed through remotely sensed data, the authors seek to derive a process by which the most likely locations for archaeological sites can be modeled in a relatively automated manner. In this vein, multispectral imagery was acquired from Landsats 5 and 8 (each having a 30m pixel resolution) and used to calculate Normalized Difference Vegetation Indices (NDVIs-calculated [NIR-Red]/[NIR+Red]) for a time-stacked series of images reflective of the rainy seasons in both 2009 and 2013 (Table 1). These NDVI images were then subtracted from one another to heighten differences in vegetation between sites and the vegetation of the surrounding area over time. The sets of time-stacked images were then used as input variables in a Multi-Layer Perceptron (MLP).

MLP is a neural network, a form of machine learning that excels at recognizing patterns through a computational abstraction of the human brain's structure (Lippitt, Rogan, Li, Eastman, and Jones 2008). Essentially, based on a set of inputs, in this case known archaeological sites, the computer uses a recursive learning procedure to understand the relationship between inputs and training data. It then uses the learned patterns to describe similarity to known samples, in this case archaeological sites, in the form of layers describing neuron activation (i.e., activation layers). In the case of presence/absence data like those used in this study, the output activation layer for the presence of a class (here archaeological sites) can be treated as a suitability layer for the likelihood of their existence. This information can then be used in the development of a random stratified sample for the purposes of survey. In this manner, high activation areas indicated by the model can be weighted more heavily than those with a low activation, allowing ground surveys to focus on areas most likely to contain archaeological remains.

This analysis was carried out in New Mexico for a variety of reasons. While not inaccessible to traditional archaeological methods and research, the vast amount of underdeveloped and undeveloped land in the state presents a major barrier to effective traditional archaeological survey. The amount of time and manpower needed to locate and analyze all the historic and prehistoric archaeological sites in the state, like many sparsely populated regions of the world, would be cost prohibitive. These factors make remote sensing the most practical method for undertaking extensive archaeological survey of expansive regions such as New Mexico.

BACKGROUND

While the basic principles of using aerial photography and vegetation indices for identifying archaeological sites are nothing new to archaeology and have been used before in archaeological location modeling (Agapiou et al. 2013; Current and Schilling 1990; Custer et al. 1986), the use of an artificial learning technique such as MLP to automate the process is a relatively novel approach. Indeed, even the use of a time-stacked series of images to find differences in vegetation indicative of anthropogenic alterations to the landscape is not commonly used in archaeological research (although it has been done at least once before, as seen in Agapiou et al. 2013).

Aerial Photography, Satellite Imagery, and Archaeology

Since the late nineteenth century, archaeologists have benefited from the use of aerial photography for the study of past landscapes (for discussions of this, see Bewley [2003]; Parry [1992]; Reeves [1936]). As far back as 1919, aerial photography was used to map Mesopotamian sites and, shortly thereafter, sites in Egypt, South America, Britain, and the United States (Bewley 2003; Reeves 1936). Soon after, variations in vegetation, such as positive and negative crop marks, were used to identify a variety

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

TABLE 1. Dates of Satellite Data Acquisitions and Correlated Time Periods Used in This Study.

Year	Time 1	Time 2	Time 3	Time 4	Time 5
2009	May 10	July 13	July 29	October 1	October 17
2013	May 30	June 22	September 26	October 12	-

of prehistoric and historic archaeological sites (Bahn and Renfrew 2008; Bewley 2003; Lasaponara and Masini 2005; St. Joseph 1945; Upex 1996).

With the advent of satellite technologies, these initial remote sensing observations were expanded upon in the 1970s and 80s. Prior to digital imaging, it was often difficult for archaeologists to detect large-scale changes in ecological diversity, along with other indicators of past anthropogenic activity. The development of satellite imagery allows for landscape wide anthropogenic activities to be detected with relatively little expense (Agapiou et al. 2013; Custer et al. 1986; Findlow and Confeld 1980; Menze and Ur 2004). Software such as Google Earth allows for archaeological research over areas too large to cover with traditional survey or aerial photography, not to mention in areas considered inaccessible to traditional study (Kennedy and Bishop 2011; Myers 2010; Sadr and Rodier 2012). Google Earth has already been used to locate stone-walled structures in South Africa (Sadr and Rodier 2012) and to locate archaeological remains across vast tracts of desert in Saudi Arabia (Kennedy and Bishop 2011). Furthermore, freely available Landsat imagery has been used to map extensive earthen features such as Roman field divisions on Mallorca (Montufo 1997) and those related to water management in Northeast Thailand (Parry 1992).

VEGETATION INDICES AND GEOSPATIAL MODELING FOR SITE DETECTION

This paper, like others specifically examining how archaeological remains manifest themselves in the growth of overlying vegetation, does not intend to model human behavior. It only models the likely places of habitation based on residual phenological markers. Combining the idea that satellite imagery can be used for site detection and the idea that crop markers and ecological diversity provide a useful proxy for the presence of anthropogenic remains, several recent studies have looked at changes in NDVI to determine differences between archaeological sites and their surroundings (Agapiou, Hadjimitsis, and Alexakis 2012; Agapiou, Hadjimitsis, Alexakis, and Sarris 2012; Lasaponara and Masini 2007). These studies have determined that archaeological sites in the agricultural fields of Cyprus, Neolithic Tells in Thessaly, and medieval sites in the rolling hills of Italy all have different NDVI returns than the areas immediately surrounding them (Agapiou, Hadjimitsis, and Alexakis 2012; Agapiou, Hadjimitsis, Alexakis, and Sarris 2012; Agapiou et al. 2013; Lasaponara and Masini 2007). Lasaponara and Masini (2005, 2006, 2007) have even taken this process one step further by examining spatial changes, as seen in linear deviations in NDVI in a process they call data fusion, to determine the layout of medieval sites using high-resolution satellite data from QuickBird. These studies demonstrate that, by using remotely sensed data, archaeologists can determine anomalies in vegetation patterns associated with the presence of archaeological sites. Taking this one step further, it should be possible to focus these efforts on the identification of probable sites based on differences in vegetation growth patterns for the purposes of archaeological survey. As high-resolution data from sources such as QuickBird and RapidEye come at a cost and cover limited area, this study will focus on the construction of a probability model for potential archaeological sites using data collected from Landsats 5 and 8, freely available from the United States Geological Service (USGS 2014a) *Earth Explorer* website.

While these data have a much lower spatial resolution than that acquired with satellites such as QuickBird and RapidEye, which offer roughly 2.8-m and 5-m pixel resolution, respectively (Heege et al. 2014; Lasaponara and Masini 2005), the key to this analysis is that it looks only for anomalies in the growth patterns of vegetation. Rather than seeking to understand the detailed pattern and structure of historic and prehistoric settlements on the landscape, it simply seeks to determine whether the patterns in vegetation growth are indicative of potential settlement or anthropogenic landscape modifications. For regional archaeological survey (i.e., identification of previously undiscovered sites), the ability to cover large swaths of land at little cost is more important than the ability to determine the pattern or layout of individual settlement features.

Modeling the Effects of Human Habitation

As discussed by Mehrer and Wescott (2006), predictive modeling using geospatial technologies is becoming increasingly common in archaeology. It has been used to understand settlement, environment, and the interplay between the two (Kohler et al. 2012; Legg and Anderton 2010). Unlike most American archaeological location models, with noted exceptions being Custer et al. (1986) and Findlow and Confeld (1980), this study does not seek to understand where a hypothetical site could have been located, but rather where sites are likely located, given observational determinatives that exist in satellite data. This perspective is common in academic studies focused on Europe (as demonstrated by Agapiou, Hadjimitsis, and Alexakis [2012]; Agapiou, Hadjimitsis, Alexakis, and Sarris [2012]; Agapiou et al. [2013]; Cavalli et al. [2007]; and Lasaponara and Masini [2006, 2007]) but stands in marked contrast to the approach of recent American settlement models that are typically based on a number of non-human influenced factors, such as seen in Kruse (2007) and Stirn (2014), or agent based modeling, such as seen in Kohler et al. (2012).

Indeed, it would appear that the trend within American research is, as one paper put it, to develop "plausible models that encompass the range of factors affecting… land use" (Hamilton et al. 2007:93). Thus, a rift can be seen as developing between the two sides of the Atlantic. Archaeological location modeling

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

in the American academic literature tends to focus on a set of features thought to be necessary for settlement, whereas many European studies seek to locate sites based on the effects of human habitation on the environment. The techniques presented here tend to follow the latter more than the former and represents modeling based on effect rather than cause.

Using the neural network MLP, this study uses differences over space and time in vegetation growth, characterized by NDVI, to model site location based on the known effects of anthropogenic features on the vegetation. It should be mentioned, however, that NDVI is not the only way to calculate a vegetation index; other indices do exist, including SR, TSAVI, and DVI, to name a few (see Table 1 in Agapiou, Hadjimitsis, Alexakis, and Sarris [2012:1500] for a complete list). NDVI was used due to the authors' familiarity and its broad application, including its successful use in a different geographic area by Agapiou et al. (2013) and Lasaponara and Masini (2005; 2006; 2007). Accounting for variation in environmental conditions (e.g., soil type), vegetation growth should be fairly uniform across both space and time within a localized area. Areas of concentrated human activity larger than 30 m^2 should show up as anomalies within any NDVI dataset based on imagery acquired from Landsat. Areas that register as having a phenological growth pattern most like that which covers other archaeological sites in the area can then be prioritized for survey over areas where vegetation growth is not reflective of known anthropogenic changes to the environment, ultimately enabling the discovery of new sites.

This method would not be able to identify smaller sites, such as those associated with hunters and gatherers or small tool production zones; this is in part due to the moderate resolution (30m) of Landsat data. The problem with using (primarily) free data is that it is typically coarser-grained (i.e., lower spatial resolution) than commercial data available for purchase. The large area covered by every pixel from moderate spatial resolution remote sensing data (30 x 30 m in the case of Landsat) makes it difficult to distinguish areas of human settlement due to every pixel being made up of multiple landscape features (i.e., the mixed pixel problem) (Lu et al. 2008). For archaeological data, poor resolution and the mixed pixel problem have traditionally meant that only sites of about one hectare (10,000 m^2) or more were detectible using satellite remote sensing data (Verhoeven and Dales 1994). Using anomalies in phenological growth over time to seek out patterns indicative of anthropogenic alterations to the environment, rather than the spatial pattern of vegetation anomalies from a single point in time, the method presented here has the potential to substantially reduce the minimum size of settlements that can be detected using remote sensing data of a given spatial resolution. When modeling the location of potential sites based on a comparison of the phenological growth of samples to that of known archaeological sites, the minimum size a site would need to be, hypothetically, is one pixel.

The effect of human activity on vegetation growth and the degree of similarity to the activities of known sites then become the primary determinants of the success of this method. Thus, a pixel or two containing a small adobe structure occupied over generations should appear similar to a pixel within a larger adobe structure of the same age due to similarities in construction, building materials, and possibly even the way in which those structures decayed. Similarly, Hejcman et al. (2013) have shown that specific use areas within a settlement promote the growth of specific species of plants. The identification of sites based on similarity in phenology, as characterized by vegetation indices, simultaneously considers both variation in vegetation species and growth rate. It relies fundamentally on the assumption that sites to be discovered exhibit similar anthropogenic effects on vegetation as those used to calibrate or train the model.

METHODS

Beginning with the assumption that the vegetation indices of plant life overlying archaeological sites differ, both spatially and temporally, from those of their surroundings (Agapiou, Hadjimitsis, and Alexakis 2012; Agapiou, Hadjimitsis, Alexakis, and Sarris 2012; Agapiou et al. 2013; Cavalli et al. 2007; Custer et al. 1986; Findlow and Confeld 1980; Lasaponara and Masini 2006, 2007), we present and test a phenology-based site suitability model that leverages the demonstrated pattern-learning capacity of MLP to differentiate between anthropomorphic/archaeological features and their surroundings (for more information on MLP analyses, see Eastman [2015:212] and Lippitt, Rogan, Li, Eastman, and Jones [2008:1203]). As MLP is a form of adaptive machine learning, it has the ability to tightly fit data with non-normal distributions (Lippitt, Rogan, Li, Eastman, and Jones 2008). The final result output from MLP, in this case, is known as an output activation layer describing the similarity of each location (i.e., raster grid cell) to the training data, which is here used as a site suitability model. MLP is known for having a higher degree of map accuracy than other parametric classifiers (Lippitt, Rogan, Li, Eastman, and Jones 2008) and for its high level of sensitivity to training parameters (Foody 2003), making it ideal for the type of modeling proposed here (Lippitt, Rogan, Toledano, Sangermano, Eastman, Mastro, and Sawyer 2008).

STUDY AREA

With the aim of locating previously unknown archaeological sites, the Sandia-Manzano Mountain Range was selected as an initial study area. The Manzano Mountains State Park (2004) has indicated that the area was in need of archaeological survey due to the high likelihood of previously unidentified archaeological sites existing within the park. Historically, the entire area was known to have held both early Spanish Colonial structures as well as a minimum of five Pueblo villages—the latter of which are all recorded as sites by the Laboratory of Anthropology in Santa Fe (Barrett 2012). Due to the high potential for archaeological discovery, the study area used in this analysis extended well beyond park boundaries across much of the mountain chain and surrounding area. Both of these locations can be found on the Path 33, Row 36, tile for LandSat but were analyzed separately to test the effectiveness of the model in the relatively homogeneous environment of the hilly Galisteo Basin versus the relatively heterogeneous environment in the highly varied elevations of the Sandia-Manzano Mountains. These areas were selected and categorized as homogeneous versus heterogeneous due to differences in flora; the Galisteo basin has more uniform vegetation, while the vegetation in the Sandia-Manzano mountains changes with elevation.

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

TABLE 2. The Calibration Sites That Went into the Basin MLP Analysis for the Galisteo Basin.

Basin Model 1	Basin Model 2	Basin Model 3
Pueblo Las Madres	Pueblo Las Madres	Galisteo Pueblo
San Lazaro	Galisteo Pueblo	San Lazaro
San Lazaro 3	San Lazaro 2	San Lazaro 2
San Cristobal	Ghost Town	San Lazaro 3
Pueblo Colorado	San Cristobal	Ghost Town
San Cristobal 2	San Cristobal 2	Pueblo Colorado
Pueblo Blanco	Pueblo She	Pueblo She
Pueblo She 2	Pueblo Blanco	Pueblo Planco 2
Pueblo She 3	Sueblo She 2	Sam Cristobal 3
Pueblo Blanco 2	Pueblo She 3	San Cristobal 4
Galisteo Pueblo 3	San Cristobal 3	Galisteo Pueblo 3
San Lazaro 4	San Cristobal 4	Galisteo Pueblo 3
San Lazaro 5	Galisteo Puueblo 4	San Lazaro 4
Pueblo Colorado 2	San Lazaro 5	Pueblo Colorado 2

TABLE 3. The Calibration Sites for the Sandia Manzano Mountain Range MLPs..

Mountain Model 1	Mountain Model 2	Mountain Model 3
Paa'ko	Gran Quivera	Gran Quivera
Tijeras 2	Abo Mission	Paa'ko
Abo Excavation	Abo Excavation	Abo Mission
Abo Buried	Abo Buried 2	Tijeras 2
Unknown Manzano Historic Site	Unknown Manzano Historic Site	Abo Buried
Quarai Mission	Tijeras	Abo Buried 2
Quarai Excavation	Quarai Mission	Tijeras 2
Paa'ko 2	Quarai Buried	Quarai Excavation
Unknown Manzano Historic Site 2	Paa'ko 2	Quarai Buried
Gran Quivera 2	Unknown Manzano Historic Site 2	Gran Quivera 3
Gran Quivera 3	Gran Quivera 2	Gran Quivera 4
Paa'ko 3	Gran Quivera 4	Paa'ko 3

Reference Data

Training sites were selected from Adams and Duff's (2004) *The Protohistoric Pueblo World, A.D. 1275–1600* and, to a lesser extent, Barrett's (2012) *The Spanish Colonial Settlement Landscape of New Mexico, 1598–1680* and Wilson's (1994) *Historic Resources Reconnaissance Survey of the Manzano and Sandia Mountain Villages.* Their names and relative locations can be seen in Tables 2 and 3, as well as in Figure 1. Since geographic coordinates for these sites were not readily available, a majority of the sites had to be located through the use of Google Earth. The locations of sites found on Google Earth were cross-referenced with maps found in Adams and Duff (2004) and Barrett (2012) that reflected their general locations. Similarly, site maps drawn during their initial excavations found on the Galisteo Basin Archaeological Sites Protection Act website (Center for New Mexico Archaeology 2014) were also used to locate these sites.

A total of only eight sites, with samples being one pixel (30 x 30 m) in size, were used in the Galisteo Basin, and five sites were used in the Sandia-Manzano Mountains. However, many of these sites were fairly large; therefore, some were broken down into two or three separate inputs so as to increase the sample size from which the MLP could train (i.e., calibrate), bringing the sample up to 23 archaeological locations (pixels) in the Galisteo Basin and 18 archaeological locations (pixels) in the Sandia-Manzano Mountains. This was determined to be viable due to work by Hejcman et al. (2013) showing that different parts of a site display different spectral signatures, given that

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

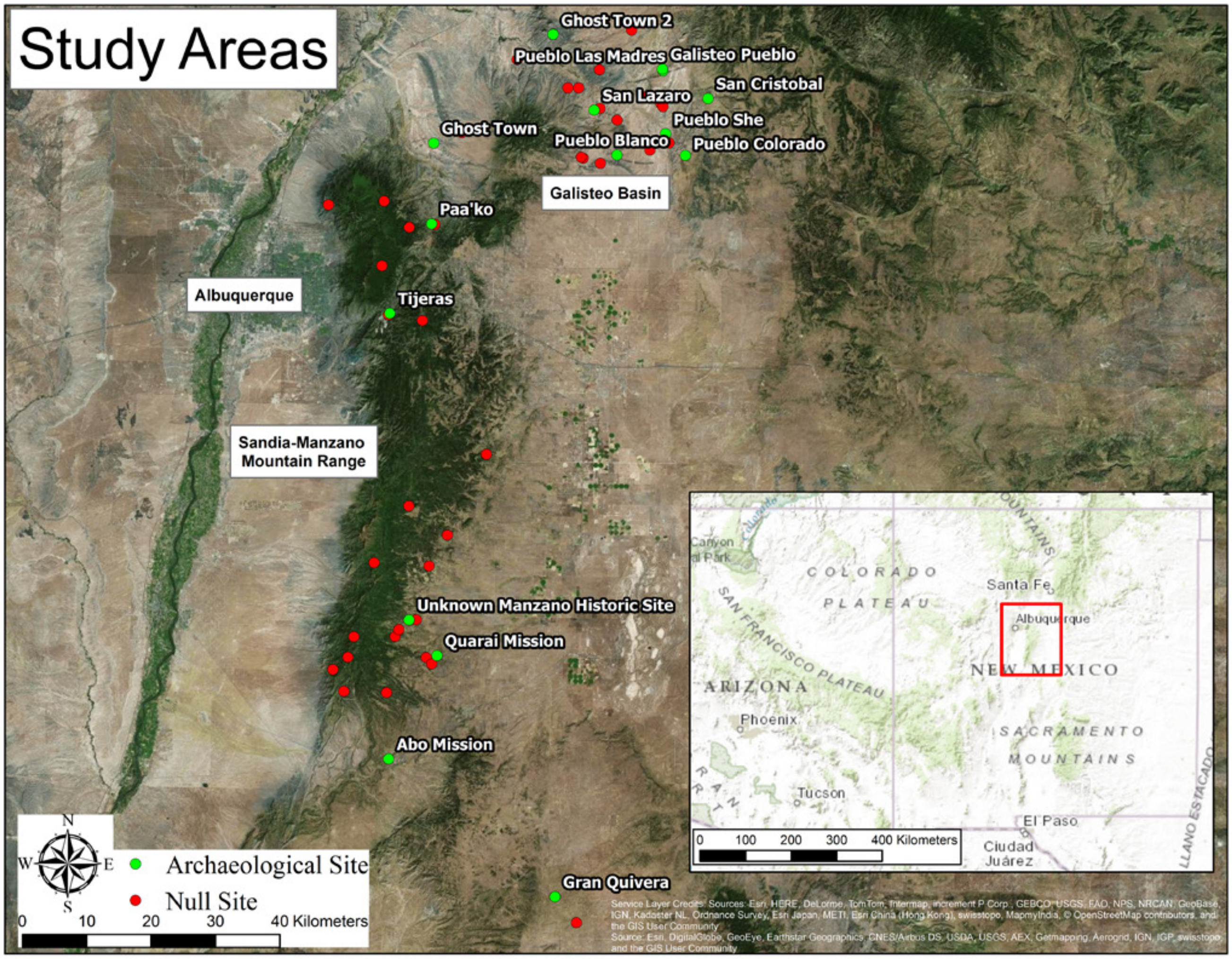

FIGURE 1. Study area and the inputs that went into MLP analysis. Known archaeological sites are shown as green dots, while null (non-sites) are designated by red dots. Large sites were further subdivided to increase the sample size for inputs.

they have different types of plants growing over them, often due to differences in activities performed across a site. This process was aided by maps available through the National Park Service (2014) website for the Salinas Pueblo Missions and, to a lesser extent, maps from the Galisteo Basin Archaeological Sites Protection Act (Center for New Mexico Archaeology 2014) website, showing the extent of these sites in cases where their full extent was not clearly visible from satellite imagery available on Google Earth. Regardless, due to the restrictive nature of archaeological site locations, the sample size in the area is still rather low. However, as this is a pilot study, it was seen that positive results would still be beneficial to the archaeological community as a whole and a larger, more representative sample will be tested in future research.

Remote Sensing Data

Landsat data was acquired from the USGS (2014a) *Earth Explorer* website for New Mexico's monsoon season in 2009, a relatively dry year, and 2013, a relatively wet year (see Table 1). This was done to gauge how much the amount of rain a site received in a year affected the results of the analysis. Furthermore, the monsoon season for each year was selected to heighten temporal changes in plant growth over archaeological sites during what is typically referred to as the "green up" period. Each image was downloaded as a Level 1 GEOTIFF that had already been preprocessed by the USGS for radiometric and geometric corrections (USGS 2014b). These images contained full-color spectra and Near Infra-Red (NIR) spectra. Using EDRISI, these images were converted into NDVI (calculated as [NIR-Red]/[NIR+Red]), using the image calculator. Each image was then clipped to reflect the different study areas, so that they could be analyzed separately.

Figure 2 shows the processing workflow. Following calculation of NDVI, a series of subtractive images were created for each year by taking earlier images and subtracting them from later ones (i.e., Time 5–Time 4, Time 5–Time 3, Time 5–Time 2, etc.). This served to heighten the differences in growth over time between archaeological sites and their surroundings, as well as to characterize the difference in growth patterns at different time lags. Once each set of images was made, values were extracted from

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

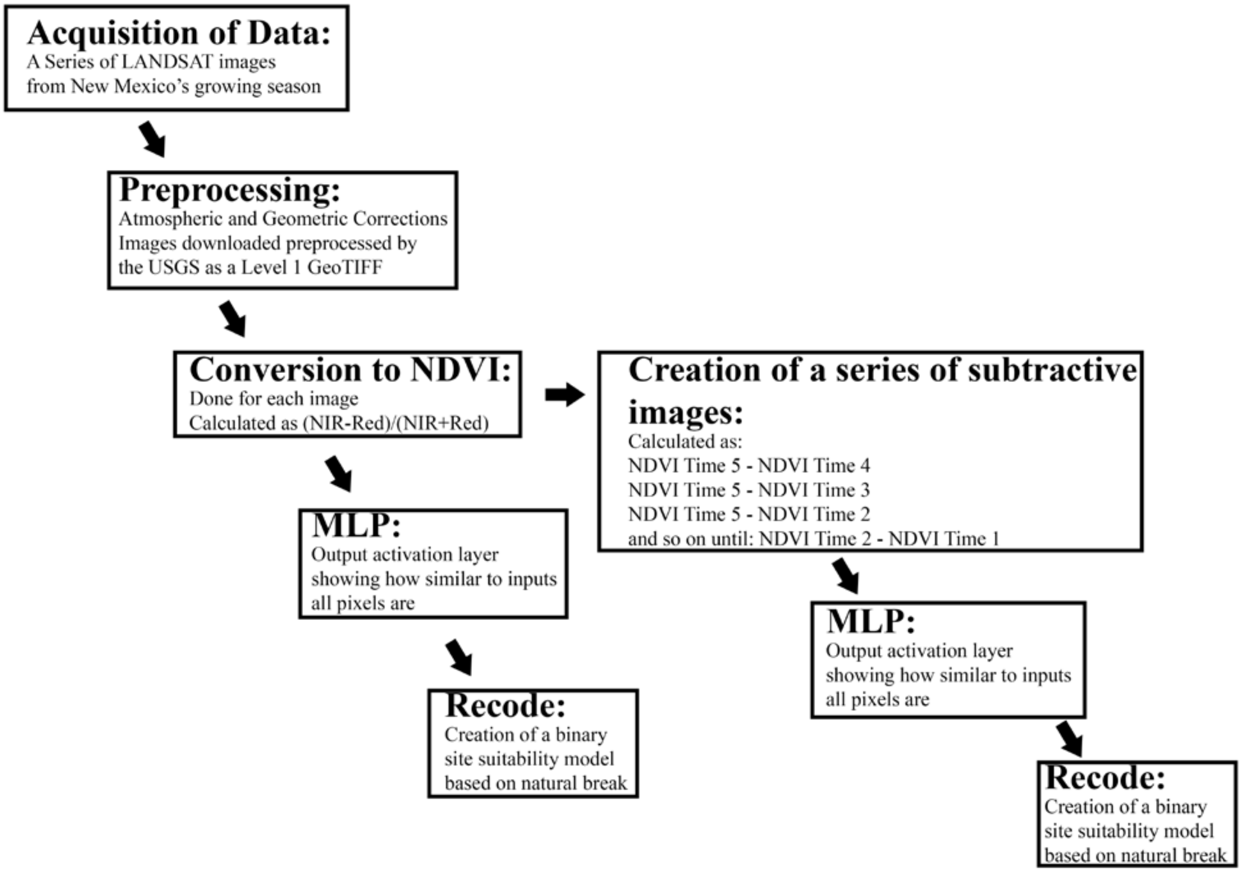

FIGURE 2. Model for using Landsat data for archaeological site detection using an MLP analysis.

a series of archaeological sites (sites that would later be used in the MLP analysis as inputs) in combination with a set of null sites (geographic points determined to not have any archaeological remains on them) to assess differences in phenological growth within the data. Previous iterations of this method concluded that within the Galisteo Basin, archaeological sites are more easily differentiated from non-sites in dry years, whereas in the Sandia-Manzano Mountain Range, sites are more easily differentiated from non-sites in wet years. Thus, for the Multi-Layer Perceptron Analysis, only the images from 2009 were used for the basin models, whereas only images from 2013 were used for the mountain models.

Analysis

MLP requires several inputs, including NDVI values for known archaeological sites, during what is known as the training procedure, where the algorithm "learns" how to identify patterns similar to the training (i.e., calibration) data it is shown. Iterations of the model were run for both the raw NDVI images and the subtractive images, in each, the Sandia-Manzano mountain range (i.e., heterogeneous vegetation environment) and the Galisteo Basin. Given the relative paucity of training data, a modified leave-one-out cross-validation technique was adopted. In each iteration, two-thirds of the total number of sites for each environment were used to construct the model, with one-third being put aside to test the accuracy of that model (Tables 2 and 3). The result was that three models were run for each set of data and each environment, leading to a total of 12 models. This was done to ensure that the procedure was both viable and repeatable based on the differences between sites and non-sites rather than the inputs being used.

Once the models were complete, accuracy was assessed using both a receiver operating characteristic area under the curve (ROC) and high-resolution satellite imagery from Google Earth. While ROC checked the validity of each model against training sites that were not included in the formation of the model, Google Earth allowed for a real test to see whether or not previously unknown anthropomorphic features were being identified. These results are discussed below.

RESULTS OF MLP: OUTPUT ACTIVATION LAYERS

For the purpose of successfully building a site suitability model, an important concept is to build a model that narrows down the number of potential sites into something manageable for survey. Using MLP's output activation layer, all models were built under this framework. The best results come from the use of the subtractive NDVI images, where all six of the models created drastically reduced the area that could potentially hold archaeological sites (Figure 3). The analysis worked significantly better in the Galisteo Basin (ROC = .99) than it did in the Sandia-Manzano

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

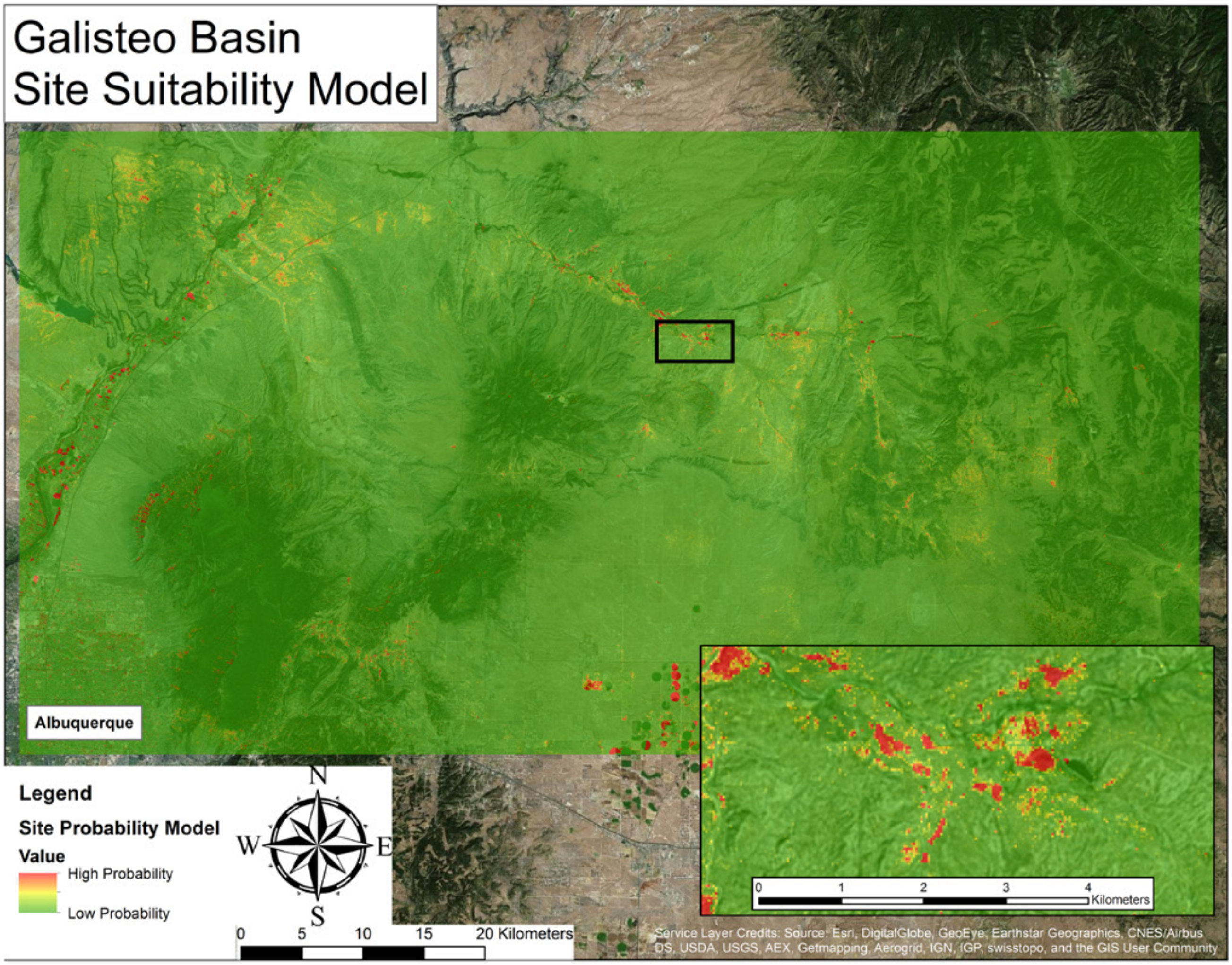

FIGURE 3: Site suitability model for the Galisteo Basin. Results from the second model run through MLP, showing how much like an input site every pixel is from high (red) to low (green). Areas of modern agricultural field near the Rio Grande and the circular fields in the center of the map. Albuquerque labeled for reference. Inset extent noted as black rectangle.

(ROC = .81–.98; average .895) Mountain Range. The output activation layers for the Galisteo Basin significantly narrowed down the number of potential sites to areas around water sources and places of logical human habitation. However, the models for the Sandia-Manzano Mountain Range narrowed down the locations that registered as potential sites only by about half; leaving an oversaturation of potential sites.

Models built for the Galisteo Basin display a higher degree of precision (i.e., a lower number of high activation areas) and accuracy (ROC = .99 as compared to ROC = .895 for the Sandia-Manzano mountain range). Model 2 (Figure 3), in particular, worked the best out of the three models for the Galisteo Basin (ROC = .99). Within Model 2, one can see the pixels most indicative of archaeological sites cluster around what appear to be arroyos and riverbeds, exactly where one would expect archaeological sites to be. Likewise, as New Mexico is an arid environment, large tracts of land appear to have no potential sites. Again, this is expected because it is unlikely that a site would be located in an area where the inhabitants would lack easy access to water.

In addition to picking out potential archaeological sites, this method also picked out agricultural fields and the urban sprawl of Albuquerque. Thus, rather than picking out only archaeological sites, the model identifies a variety of anthropogenic activity on the landscape. If one wishes to mask out currently built environments, it is entirely possible to do that using existing land cover datasets (e.g., the United States Geological Survey National Land Cover Dataset).

As both Foody (2003) and Lippitt, Rogan, Li, Eastman, and Jones (2008) have pointed out, MLP is highly sensitive to training data. If the researcher feels that the results of the MLP analysis are too sensitive, then there are ways to correct for it. Two patterns are apparent when the results of Model 2 in the Galisteo Basin are taken as a case study and the values are extracted and plotted onto a bar graph (Figure 4) for both archaeological sites and a set of null sites (areas that are selected randomly based on satellite imagery and, for a variety of reasons, are believed not to contain any archaeological remains). The first is that certain sites, such as Pueblo Las Madres, Pueblo Colorado, and Pueblo Blanco, do not share the same output values as other sites,

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

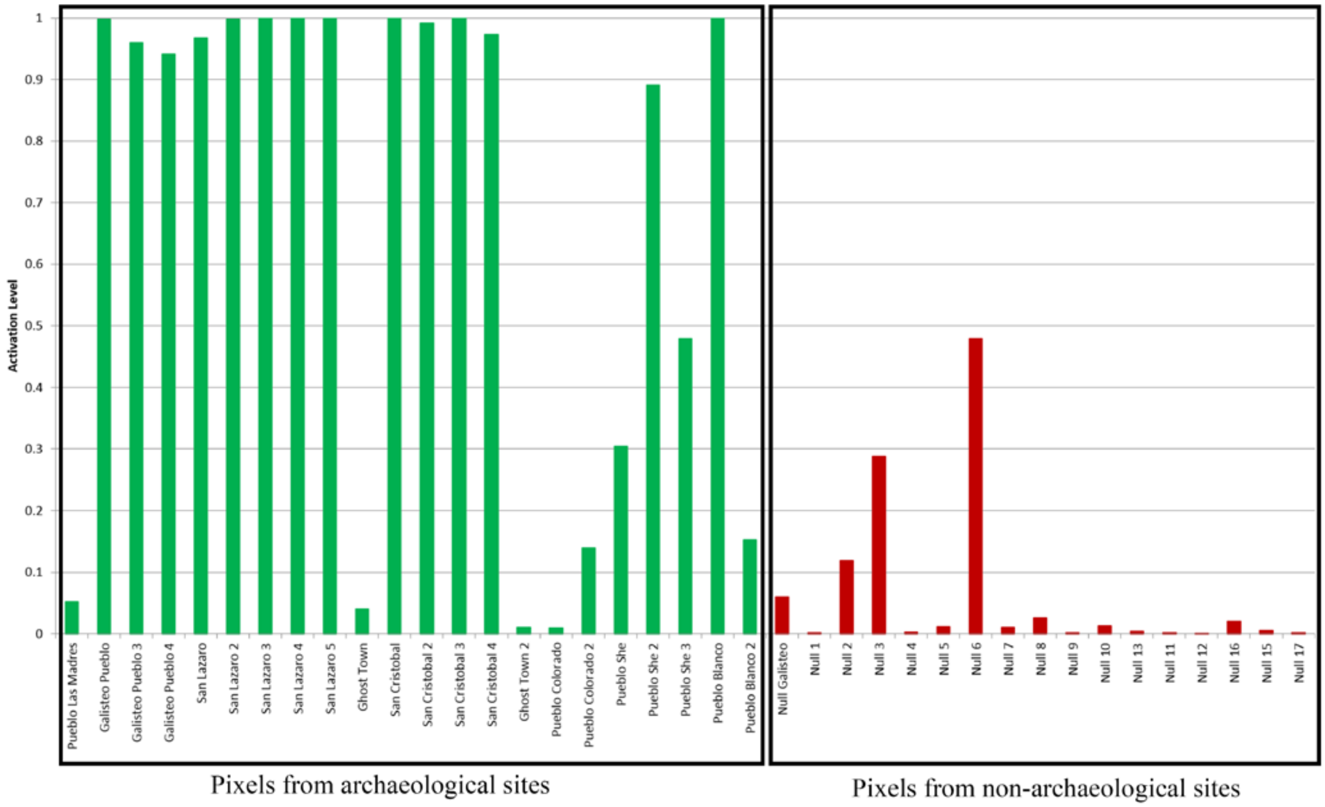

FIGURE 4. Differences in value results between archaeological sites and null sites used for inputs in the Galisteo Basin Suitability Model. Using this graph, it is possible to separate the values for sites and non-sites to create the binary image seen in Figure 5.

indicating that the pixels selected to represent them may have been poorly selected (i.e., a marker may have been placed on a pixel located on the periphery of the site, which would cause problems due to the mixed pixel problem (Lu et al. 2008)), and/or that the phenological growth pattern of these sites substantially differs from that of the other archaeological sites. The second is that there is a natural break between a majority of the archaeological sites and the null sites.

If one takes this natural break and creates a binary image (Figure 5), the number of pixels representing likely sites is greatly reduced in a process referred to as thresholding. This was done for both the Basin and Mountain environments by taking the pixel value just above the highest null site to the nearest tenth of a decimal and reclassifying higher values as potential sites and lower values as non-sites. Again, using Model 2 for the Galisteo Basin as a case study, every pixel value of .5 and above was reclassified as a potential site and every pixel of .4999 and below was reclassified as having a low likelihood of being an archaeological site. Thresholds here were set conservatively, but an equally viable option would be to set the threshold much higher, attempting to seek out only those pixels that most reflect the training data, which would, of course, narrow the target search area even further.

ROC and Satellite Validation

Additional validation of this method was completed by using an ROC based on the set of inputs withheld from model training. As previously stated, one-third of the potential sites were left

TABLE 4. ROC Validation Number for the MLP Results.

Model	AUC	Model	AUC
NDVI Basin 1	.998501	Subtractive Basin 1	.998159
NDVI Basin 2	.993501	Subtractive Basin 2	.998392
NDVI Basin 3	.797502	Subtractive Basin 3	.997734
NDVI Mountain 1	.909502	Subtractive Mountain 1	.817501
NDVI Mountain 2	.986503	Subtractive Mountain 2	.989502
NDVI Mountain 3	.899501	Subtractive Mountain 3	.893502

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

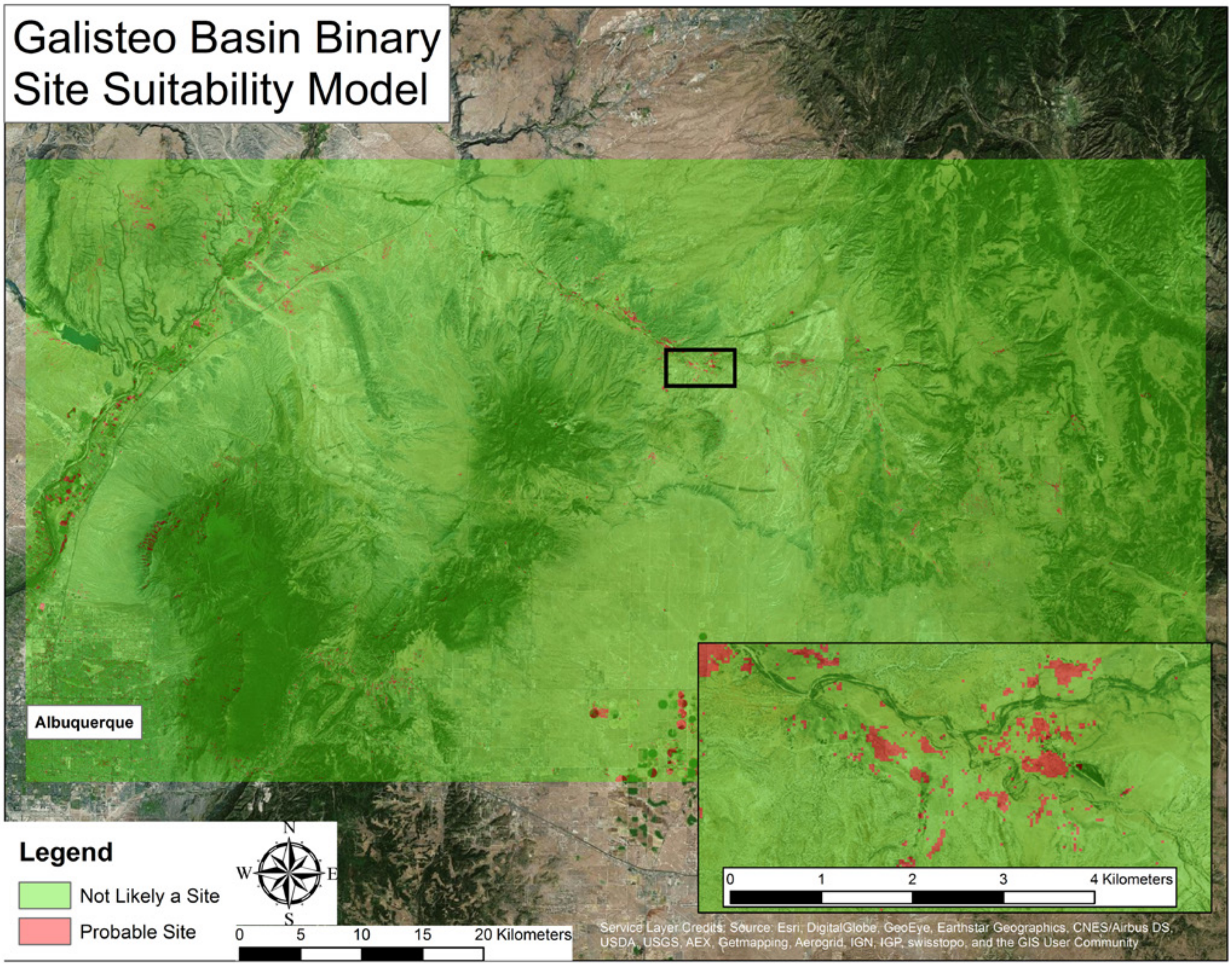

FIGURE 5. Binary image for the Galisteo Basin. Red indicates potential sites, and green indicates areas not likely to contain a site.

out of each iteration of the models created by MLP, so that this could be done. Results for ROC (Table 4) show that five of the 12 models created have a 99 percent accuracy rate with no model having an accuracy rate lower than 79 percent. However, this would indicate that the models created for the Sandia-Manzano Mountains also have a relatively high accuracy. Therefore, it was necessary to visually check a random sample of probable sites for accuracy as well.

Using the models based on the subtractive images showing the best results for both the Galisteo Basin and the Sandia-Manzano Mountain Range (Model 2 for both), a set of geographic coordinates were selected for this test (Figure 6). In the Galisteo Basin, 13 of17 randomly selected pixels from the binary image that registered as being a potential site showed some form of modern, historic, or prehistoric human modification when the coordinates were examined using Google Earth (Figure 7; Table 5). Human modification, here, is defined by the presence of straight lines or linear features visible from the air. However, within the coordinates selected for the mountains, only one out of five potential sites seemed to show evidence for human modification, though it could be that anthropogenic features may not be visible from the surface in these areas due to heavy vegetation cover (Figure 7; Table 5). Only five sites were chosen due to the oversaturation of potential sites in the study area. Therefore, it seems that, when using this technique to predict site location, uniformity in vegetation phenology (in part due to elevation) is key. We note that this study would benefit from a field-based follow-up in which areas that registered as potential sites in our analysis could be ground truthed or shovel tested to determine accuracy.

DISCUSSION OF THE DIFFERENCES BETWEEN BASIN AND MOUNTAIN ENVIRONMENTS

Within a large, open area with relatively homogeneous vegetation and a uniform topography, such as the Galisteo Basin, the use of a phenology-based suitability model for the identification of potential archaeological sites appears promising. The output

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

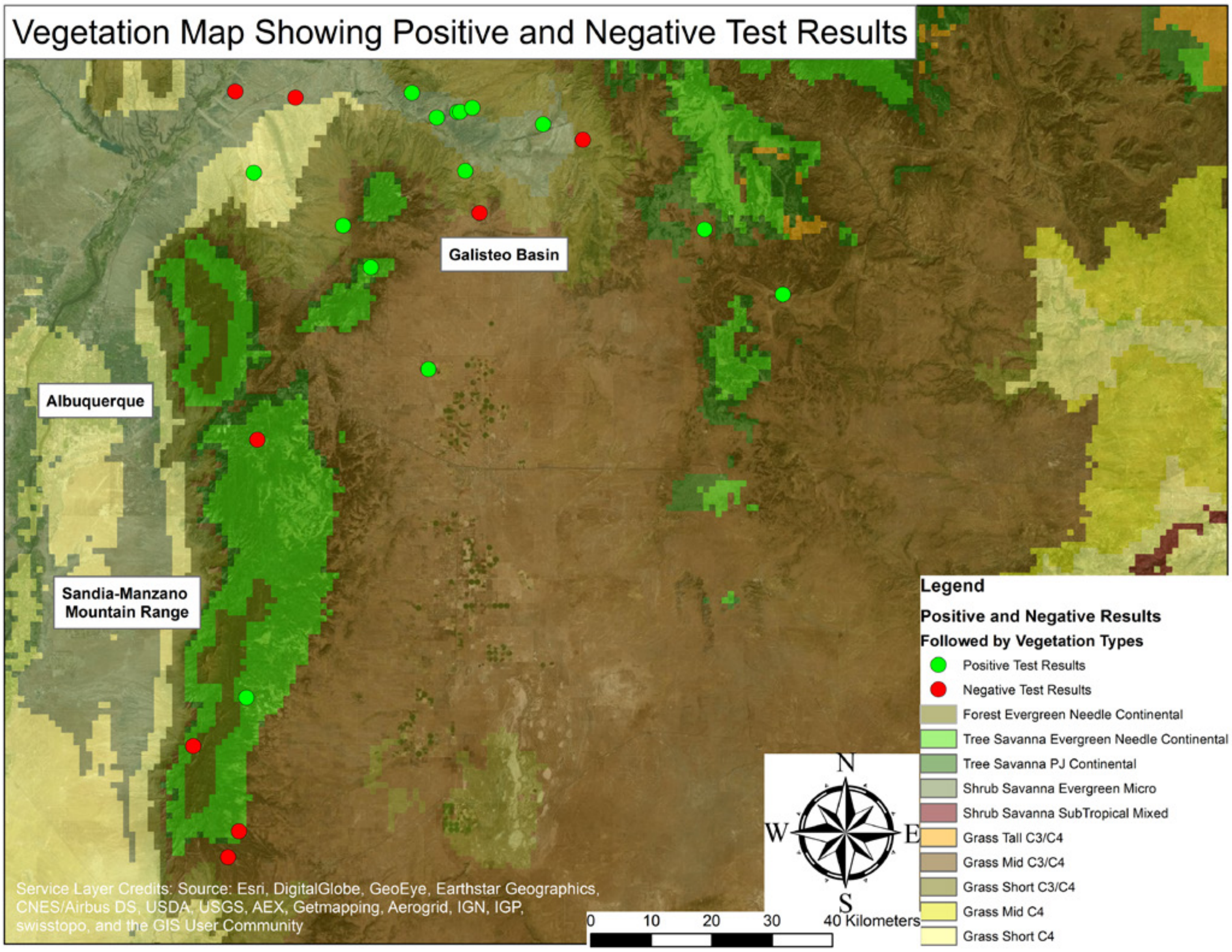

FIGURE 6. Locations that MLP characterized as being most likely to have sites were checked for signals of anthropogenic modification on Google Earth. These results were then plotted onto a vegetation map, with green dots indicating signs of anthropogenic modification and red dots indicating no visible signs.

activation layers produced by MLP can be used as a weight in the creation of a random stratified sample for the purposes of survey. However, results were not nearly as promising in the Sandia-Manzano Mountain Range (Figure 8). Looking at Model 2 for the subtractive images of the mountains, one can see that nearly half the pixels in the binary image were classified as potential sites, likely indicating an extremely high false positive rate. However, we hypothesize that this is likely due to changes in the ecological diversity that occur with changes in elevation. Nonetheless, if the models for the Sandia-Manzano Mountain Range do accurately narrow down what could be considered a potential site, then the model presented here succeeds in reducing the area that needs to be surveyed by half. Thus, this method still presents viable information for weighting a random stratified sample in survey.

Essentially, while the Galisteo Basin is characterized by more than one vegetation type, the environment itself is still fairly uniform. Within the mountains, however, as one ascends in elevation, the environment and ecological diversity has the potential for radical change. If the changes in phenology and vegetation caused by elevation are the cause of reduced inferential power in the mountain study site, it is likely that stratification of the model by elevation and environment type could substantially improve the usefulness of this technique in mountainous environments. However, this presents a challenge to the use of MLP; to replicate this study within the stratified layers of the Sandia-Manzano Mountains, there would need to be representative samples of archaeological sites in each stratigraphic layer to build the model. This is not the case in the Sandia-Manzanos.

POTENTIAL APPLICATIONS AND CONCLUSIONS

While the approach described here will allow for the detection of large, complex archaeological sites in relatively uniform environments such as basins and plains, using it across more heterogeneous environments, such as montane settings, presents certain challenges. In addition to its potential for locating and identifying previously unknown sites in large, open environ-

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

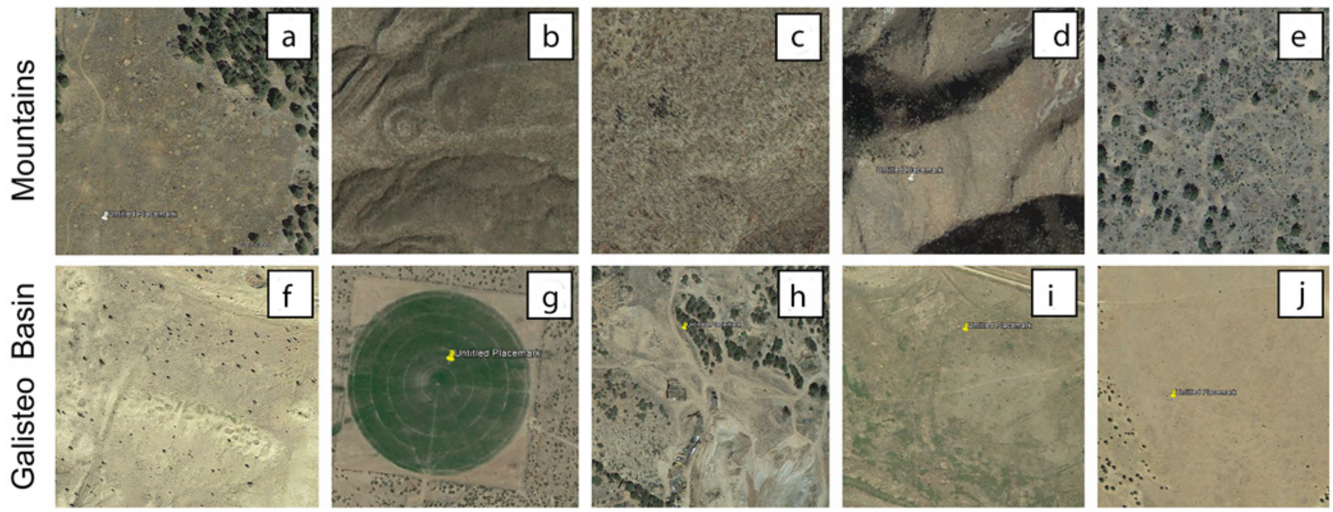

FIGURE 7. A subset of Google Earth images showing randomly selected high potential areas. Several areas reflect anthropogenic modifications: (a) and (b) Sandia Manzano Mountain Range; (f)–(i) Galisteo Basin. Other areas show no indications of human landscape alterations: (b)–(e) Sandia Manzano Mountain Range; (j) Galisteo Basin.

ments, in practical terms, this technique can also be used to weight a random, stratified sample for archaeological survey. Again, however, the size of the changes in sites would be an issue here as well. It should be reiterated that this technique cannot pick out smaller archaeological sites and features, though higher spatial resolution datasets may enable detection of smaller sites.

This pilot study focuses on the difference in vegetation over time within the study zones. Future iterations of this research will include inputs of elevation, soils, and rainfall, all of which directly impact the phenology of local flora. Additionally, NDVI is not the only vegetation index that can be run with MLP. Other indices (e.g., indexes of moisture, geology, other vegetation indices) should be run in the future to determine whether they are more sensitive to the presence and absence of archaeological sites. Similarly, the paired use of phenology with MLP may not be the only way to use this method in order to detect previously unknown sites. Other markers for site location (e.g., soil, water retention) may also be possible.

The phenology-based site detection model outlined here allows archaeologists to focus their efforts on areas with the greatest potential for archaeological sites, saving both time and money. Nonetheless, there are certain limitations. This approach can pick up only large, agglomerated sites, such as pueblos, when using the coarse-grained data provided by Landsat. Data from new and expected remote sensing systems (e.g., RapidEye) may allow this analysis to be applied at a finer grain, although the heterogeneity implicit in finer-grained data may present a challenge.

Even with more fine-grained data, it is unlikely that this method could ever pick out smaller sites, such as those left by hunters and gatherers, due to their ephemeral nature and the minimal effect they likely have on vegetation. This method should be seen as a complement to traditional archaeological survey rather than a replacement (Custer et al. 1986:583; Montufo 1997). It allows for the identification of larger sites but can in no way take the place of survey for the detection of smaller sites. There is a limit to what remotely sensed data can detect. Valuable data will still be left undiscovered without the use of the traditional detailed survey methods established by scholars, such as the contributors to Kent Flannery's (1982) *The Early Mesoamerican Village*. Nonetheless, because of its potential for finding previously unknown sites, and the implications this may have for the protection of national treasures from looters through cultural heritage management, this method does seem to be worth exploring.

Acknowledgments

We would like to thank the USGS and NASA for providing Landsat Data via the *Earth Explorer* website. Without this data, the project would not have been possible. Additionally, we would like to thank the Department of Anthropology and the Department of Geography and Environmental Studies at the University of New Mexico. A special thank you to Dr. James L. Boone, Eric C. Fries, Kristina Whitney, and Erin Hegberg for their insightful comments.

TABLE 5. Matrix Showing the Results of a Google Earth-Based Analysis of Randomly Selected High Potential Pixels.

Galisteo Basin			Sandia-Manzano Mountains	
Yes	No	Model	Yes	No
N/A	N/A	No	N/A	N/A
13 (76%)	4 (24%)	Yes	1 (20%)	4 (80%)

Note: Pixels that did not register as sites were not checked on Google Earth for potential sites.

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

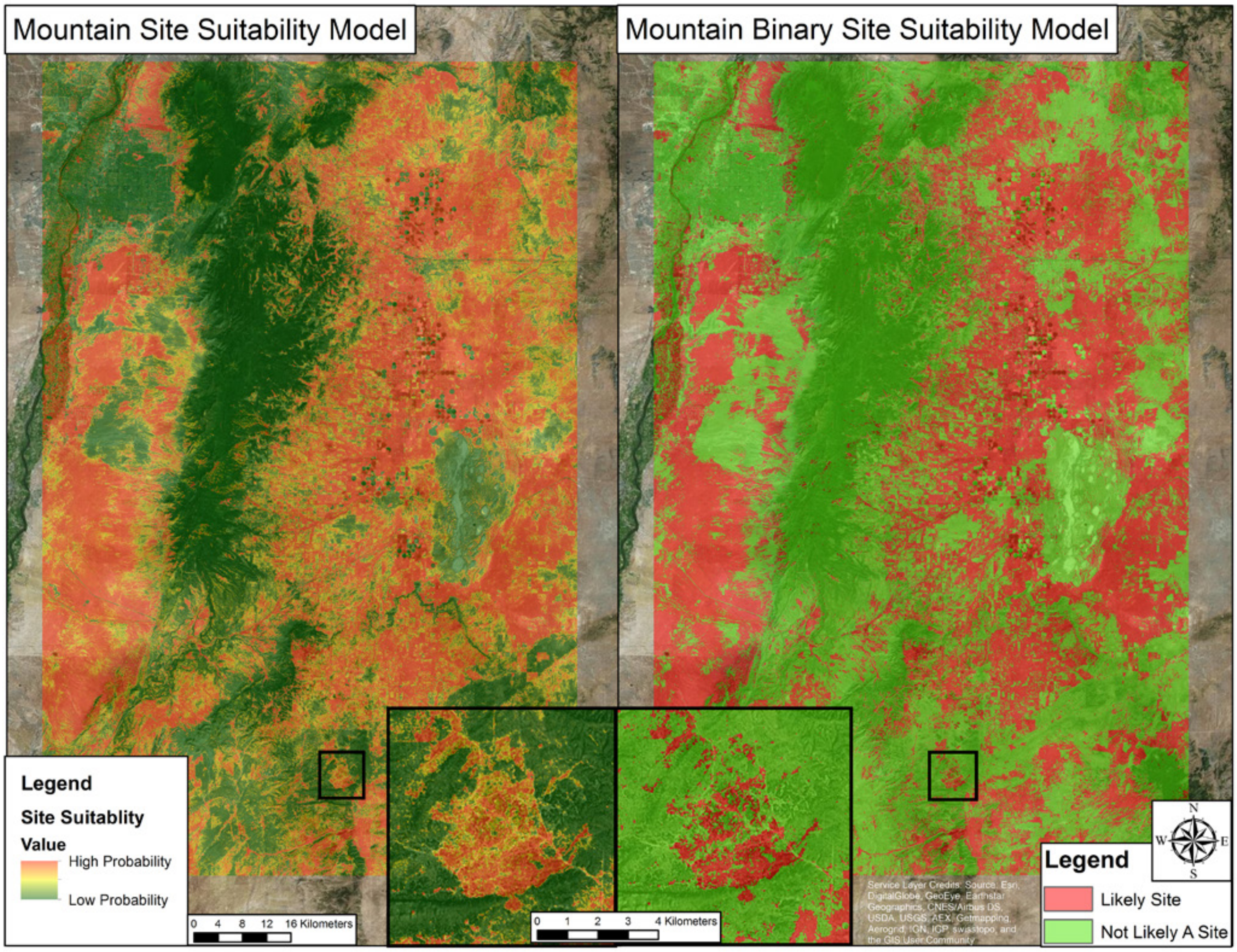

FIGURE 8. Results for the Sandia-Manzano Mountain Range.

Data Availability Statement

Original data were not used in the preparation of this manuscript, and data that were used can be readily obtained from the Internet. The satellite imagery used in this study can be downloaded for free from http://earthexplorer.usgs.gov/ by looking at the correct tile (Path 33, Row 36) under the Landsat Archive dataset. Coordinates for archaeological sites used in this experiment can be acquired through the New Mexico Department of Cultural Affairs at http://www.nmhistoricpreservation.org/arms/faqs.html or by searching out the coordinates for individual sites on Google. For this study, the coordinates of sites were found by searching Google Earth for their location.

REFERENCES CITED

Adams, E. Charles, and Andrew I. Duff (editors)
2004 *The Protohistoric Pueblo World, A.D. 1275–1600*. University of Arizona Press, Tucson.

Agapiou, Athos, Diofantos G. Hadjimitsis, and Dimitrios D. Alexakis
2012 Evaluation of Broadband and Narrowband Vegetation Indices for the Identification of Archaeological Crop Marks. *Remote Sensing* 4:3892–3919.
2013 Development of an Image-Based Method for the Detection of Archaeological Buried Relics Using Multi-temporal Satellite Imagery. *International Journal of Remote Sensing*. 34:5979–5996.

Agapiou, Athos, Diofantos G. Hadjimitsis, Dimitrios Alexakis, and Apostolos Sarris
2012 Observatory Validation of Neolithic Tells (Magoules) in the Thessalian Plain, Central Greece, Using Hyperspectral Spectroradiometric Data. *Journal of Archaeological Science* 39:1499–1512.

Bahn, Paul, and Colin Renfrew
2008 *Archaeology: Theories, Methods, and Practice*. 5th ed. Thames and Hudson, London.

Barrett, Elinore M.
2012 *The Spanish Colonial Settlement Landscape of New Mexico, 1598–1680*. University of New Mexico Press, Albuquerque.

Bewley, Robert H.
2003 Aerial Survey for Archaeology. *The Photogrammetric Record* 18:273–292.

Cavalli, Rosa Maria, Francesca Colosi, Angelo Polombo, Steffano Pignatti, and Maurizio Poscolieri
2007 Remote Hyperspectral Imagery as a Support to Archaeological Prospection. *Journal of Cultural Heritage* 8:272–283.

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press

Center for New Mexico Archaeology
2014 *Galisteo Basin Archaeological Sites Protection Act.* Electronic resource, http://galisteo.nmarchaeology.org/index.html, accessed November 13, 2015.

Chase, Arlen F., Diane Z. Chase, Christopher T. Fisher, Stephen J. Leisz, and John F. Weishampel
2012 Geospectral Revolution and Remote Sensing LiDAR in Mesoamerican Archaeology. *PNAS* 109:12916–12921.

Current, John R., and David A. Schilling
1990 Location Modeling: Perspective and Overview. *Geographical Analysis* 22:1–3.

Custer, Jay F., Timothy Eveleigh, Vytautas Klemas, and Ian 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* 51:572–588.

Eastman, J. Ronald
2015 *TerrSet Manual.* Clark Labs, Worcester.

Findlow, Frank, and Linda Confeld
1980 Landsat Imagery and the Analysis of Archaeological Catchment Territories: A Test of the Method of Catchment Analysis. In *Catchment Analysis: Essays on Prehistoric Resource Space,* edited by Frank J. Findlow and Jonathon E. Ericson, pp. 31–52. University of California, Los Angeles.

Flannery, Kent (editor)
1982 *The Early Mesoamerican Village.* Academic Press, Waltham.

Foody, G.M.
2003 Uncertainty, Knowledge Discovery and Datamining in GIS. *Progress in Physical Geography* 27:113–121.

Hamilton, Scott, James Graham, and B.A. Nicholson
2007 Archaeological Site Distributions and Contents: Modeling Late Precontact Blackduck Land Use in the Northeastern Plains. *Canadian Journal of Archaeology/ Journal Canadien d'Archéologie* 31(3):93–136.

Heege, Thomas, Viacheslav Kiseleva, Magnus Wettlea, and Nguyen Nghia Hungb
2014 Operational Multi-Sensor Monitoring of Turbidity for the Entire Mekong Delta. *International Journal of Remote Sensing* 35:2910–2926.

Hejcman, M., P. Karlík, J. Ondráček, and T. Klír
2013 Short-Term Medieval Settlement Activities Irreversibly Changed Forest Soils and Vegetation in Central Europe. *Ecosystems* 16:652–663.

Hritz, Carrie
2014 Contributions of GIS and Satellite-based Remote Sensing to Landscape Archaeology in the Middle East. *Journal of Archaeological Research* 22:229–276.

Kennedy, David, and M.C. Bishop
2011 Google and the Archaeology of Saudi Arabia. A Case Study from the Jeddah Area. *Journal of Archaeological Science* 38:1284–1293.

Kohler, Timothy A., R. Kyle Bocinsky, Denton Cockburn, Stefani A. Crabtree, Mark D. Varien, Kenneth E. Kolm, Schaun Smith, Scott G. Ortman, and Ziad Kobti
2012 Modelling Prehispanic Pueblo Societies in Their Ecosystems. *Ecological Modeling* 241:30–41.

Kruse, Melissa
2007 The Agricultural Landscape of Perry Mesa: Modeling Residential Site Location in Relation to Arable Land. *Kiva* 73(1):85–102.

Lasaponara, Rosa, and Nicola Masini
2005 Quickbird-Based Analysis for the Spatial Characterization of Archaeological Sites: Case Study of the Monte Serico Medieval Village. *Geophysical Research Letters* 32(12):L12313.
2006 Identification of Archaeological Buried Remains Based on the Normalized Difference Vegetation Index (NDVI) from QuickBird Satellite Data. *IEEE Geoscience and Remote Sensing Letters* 3:325–328.
2007 Detection of Archaeological Crop Marks by Using Satellite QuickBird Multispectral Imagery. *Journal of Archaeological Science* 24:214–221.

Legg, Robert J., and John B. Anderton
2010 Using Paleoshoreline and Site Location Modeling in the Northern Great Lakes: Geoarchaeological Approaches to Prehistoric Archaeological Survey in the Pictured Rocks National Lakeshore. *Geoarchaeology* 25:772–783.

Lippitt, Christopher D., John Rogan, Zhe Li, J. Ronald Eastman, and Trevor G. Jones
2008 Mapping Selective Logging in Mixed Deciduous Forest: A Comparison of Machine Learning Algorithms. *Photogrammetric Engineering and Remote Sensing* 74:1201–1211.

Lippitt, Christopher D., John Rogan, James Toledano, Florencia Sangermano, J. Ronald Eastman, Victor Mastro, and Alan Sawyer
2008 Incorporating Anthropogenic Variables into a Species Distribution Model to Map Gypsy Moth Risk. *Ecological Modeling* 210:339–350.

Lu, Dengsheng, Hanqin Tian, Guomo Zhou, and Hongli Ge
2008 Regional Mapping of Human Settlements in Southeastern China with Multisensory Remotely Sensed Data. *Remote Sensing of Environment* 112:3668–36679.

Manzano Mountains State Park
2004 *Manzano Mountains State Park Management Plan 2004–2008.* Electronic document, http://www.emnrd.state.nm.us/SPD/documents/ManzanoStatePark_000.pdf, accessed November 13, 2015.

Mehrer, Mark, and Konnie Wescott
2006 *GIS and Archaeological Site Location Modeling.* Taylor and Francis, Boca Raton.

Menze, Bjoern H., and Jason A. Ur
2014 Multitemporal Fusion for the Detection of Static Spatial Patterns in Multispectral Satellite Images—With Application to Archaeological Survey. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7:3513–3524.

Montufo, Antonio M.
1997 The Use of Satellite Imagery and Digital Image Processing in Landscape Archaeology: A Case Study from the Island of Mallorca, Spain. *Geoarchaeology* 12:71–85.

Myers, Adrian
2010 Camp Delta, Google Earth, and the Ethics of Remote Sensing in Archaeology. *World Archaeology* 42:455–467.

National Park Service
2014 *Salinas Pueblo Missions.* Electronic document, http://www.nps.gov/sapu/index.htm, accessed November 13, 2015.

Parry, John T.
1992 The Investigative Role of Landsat-TM in the Examination of Pre and Proto Historic Water Management sites in Northeast Thailand. *Geocarto International* 7(4):5–24.

Reeves, Dache M.
1936 Aerial Photography and Archaeology. *American Antiquity* 2:102–107.

Sadr, Karin, and Xavier Rodier
2012 Google Earth, GIS and Stone-walled Structures in Southern Gauteng, South Africa. *Journal of Archaeological Science* 39:1034–1042.

St. Joseph, J.K.
1945 Air Photography and Archaeology. *The Geographical Journal* 105:47–59.

Stirn, Matthew
2014 Modeling Site Location Patterns Amongst Late-Prehistoric Villages in the Wind River Range, Wyoming. *Journal of Archaeological Science* 41:523–532.

United States Geological Service (USGS)
2014a USGS Earth Explorer. *USGS.* 01 July 2014. Electronic document, http://earthexplorer.usgs.gov/, accessed November 13, 2015.
2014b Landsat Processing Details. *USGS.* 08 December 2014. Electronic document, http://landsat.usgs.gov/Landsat_Processing_Details.php, accessed November 13, 2015.

Upex, Stephen

1996 Leicestershire Headlands: Some Cropmarks in the Midlands. *Current Archaeology* 13:191–193.

Verhoeven, Kris, and L. Dales

1994 Remote Sensing and Geographical Information Systems (GIS for Archaeological Research (Applied in Mesopotamia). In *Cinquante-deux reflexions sur le Proche-Orient ancien*, edited by Hermann Gasche, Michel Tanret, C. Janssen, and A. Degraeve, pp. 519–539. Peeters, Ghent.

Wilson, Chris

1994 *Historic Resources Reconnaissance Survey of the Manzano and Sandia Mountain Villages*. State Historic Preservation Division of Cultural Affairs, Santa Fe.

AUTHOR INFORMATION

Scott Detrich Kirk and **Amy E. Thompson** ■ Department of Anthropology, University of New Mexico, Albuquerque, NM 87131 (kirks@unm.edu; athomp04@unm.edu)

Christopher D. Lippitt ■ Department of Geography, University of New Mexico, Albuquerque, NM 87131 (clippitt@unm.edu)

https://doi.org/10.7183/2326-3768.4.1.87 Published online by Cambridge University Press