Past and Prediction: Archaeology and ArcGIS in Cultural Resource Management

Seth Johnstone

ABSTRACT

The utilization of ESRI's ArcGIS software greatly facilitates the management of cultural resources on Federal lands. Correlation between archaeological sites and environmental variables can be rapidly accomplished through the analysis of both vector and raster data sets. The ability to join multiple feature classes based on spatial location has greatly aided the identification of areas of high, medium, and low probability for prehistoric and historic site locations. Applications of relatively simple statistical techniques, such as chi-squared evaluations, can then be applied to this spatial data to weight different environmental and cultural variables for final model production using the Spatial Analyst extension. This paper explores the procedures involved in developing a predictive model within a GIS-based environment utilizing data from Fort Stewart Military Installation, Georgia. Implications for the greater body of anthropological knowledge concerning human settlement in the Southeast United States will also be briefly touched upon.

INTRODUCTION

Balancing the training requirements of military installations with concerns for the preservation of cultural resources under their care is facilitated by the use of site probability models that can accurately depict the nature of archaeological distributions. Limited budgets, narrow work schedules, and the evolving need for larger and more diversified training areas can limit the ability of cultural resource managers to inventory and determine eligibility of archaeological resources before they are impacted. Predictive models are created in an effort to prevent this by delineating areas that can undergo less rigorous survey and testing in order to better allocate limited funds. In the case of Fort Stewart Military Installation near Savannah, Georgia (Figure 1), the current model is used to determine areas that can undergo low probability testing on a 45-x-45 m grid, a practice that requires anywhere from one-half to two-thirds the funding of investigating the same area on a high probability 30-x-30 m grid.

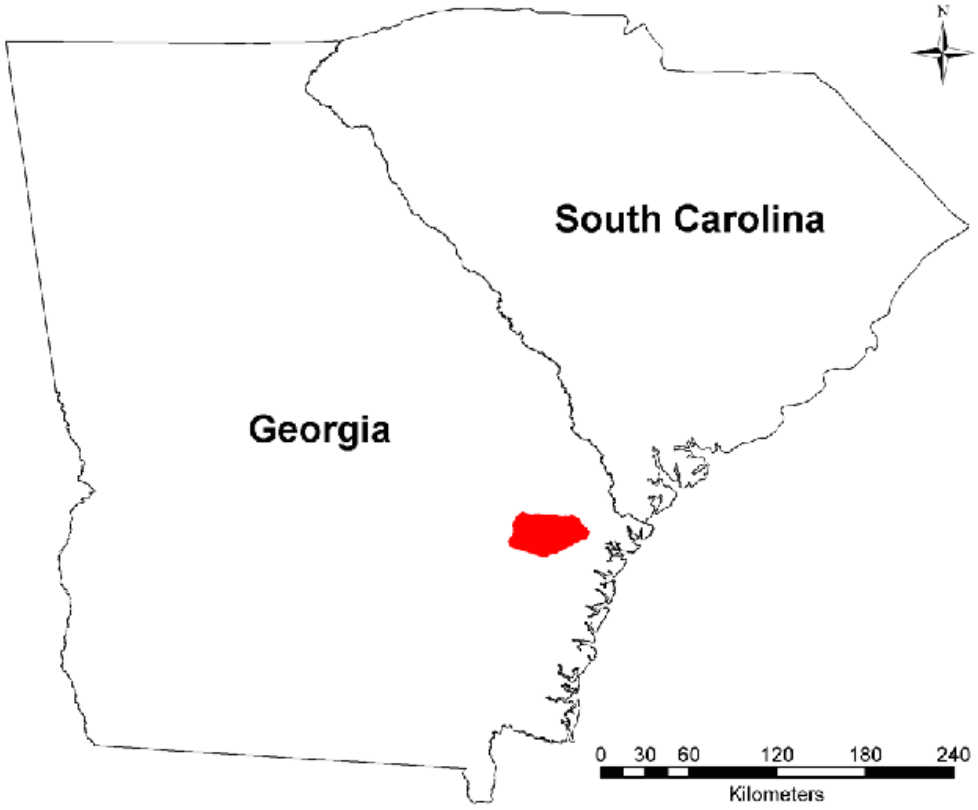

Figure 1. Location of Fort Stewart in Red.

The current model in use is solely dependent on the presence of specific soil types for determining areas of high and low probability for containing archaeological sites. Some of the dangers in utilizing a single variable for designating probability zones should be apparent, however some discussion is still warranted. Two primary questions come to mind when using only soil type as an indicator of site location. First, were soil types salient enough in either the conscious or subconscious minds of prehistoric and historic settlers of the area to choose one type over another to settle upon and exploit? Second, is a large enough area of each soil type represented in our sample surveys to be capable of showing significant correlation with sites, and thereby support the proposition that specific soil types were selected for or against?

With regards to the first question, data gathered during the course of the current project suggests that it was a combination of factors that contributed to the favorability of certain areas over others for settlement or resource exploitation. While soil type may have been important to some extent, it is more likely that certain geomorphological and environmental characteristics that contribute to the creation and maintenance of the soil type played the primary role in the decision making process regarding site location. Concerning the second question, only five out of the 71 different soil types represented at the installation have sample areas large enough to run chi-squared evaluations on site location. The difficulty in conducting statistical analyses on such a diverse sample universe negates the ability to offer any meaningful justification for site probability, except for these five well-represented soil types. For these two reasons alone, differential survey occurring solely on the basis of soil type designations should be considered insufficient.

The limitations of the old model in conjunction with the need to develop a new model that takes multiple factors into consideration for determining areas that are likely to contain prehistoric or historic resources acted as the catalyst for the current project. Using ArcGIS 8.2 in combination with multiple cultural and environmental data sets, a more accurate representation of the nature of prehistoric and historic inhabitation at the installation was developed. In addition, research questions generated by the illustration of high and low probability zones should aid in the determination of National Register of Historic Places (NRHP) eligibility. The following section will focus on the actual implementation of a program utilizing ArcGIS 8.2, Spatial Analyst, 3D Analyst, Statistical Analyst, and actual statistical analysis in creating a probability, or correlative, model for the management of cultural resources.

METHODOLOGY

Fort Stewart covers an area of approximately 1130 km^2, and is situated within the Coastal Plain of Georgia in a region that exhibits a relatively homogenous topographic landscape. Despite this limited variation, measurements depicting even minor changes across the study area have helped to better illustrate the nature of site locations. Using digitized contours obtained from the installation, it was possible to generate a TIN elevation surface of the study area (Figure 2) with the 3D Analyst extension, which was extrapolated to derive rasters of the following environmental attributes: elevation, slope (Figure 3), local relief, and aspect. Local relief, which refers to the difference between the highest and lowest points within a fixed radius of a sample point, was evaluated at a 50-m (Figure 4), 150-m (Figure 5), and 250-m radius (Figure 6). Aspect is dependent on the direction of slope, and is measured in degrees with 0° representing flat areas and 360° representing areas facing north (Figure 7). Using the raster calculator in Spatial Analyst, each of these six variables was converted from raster to vector format for the actual site distribution analysis.

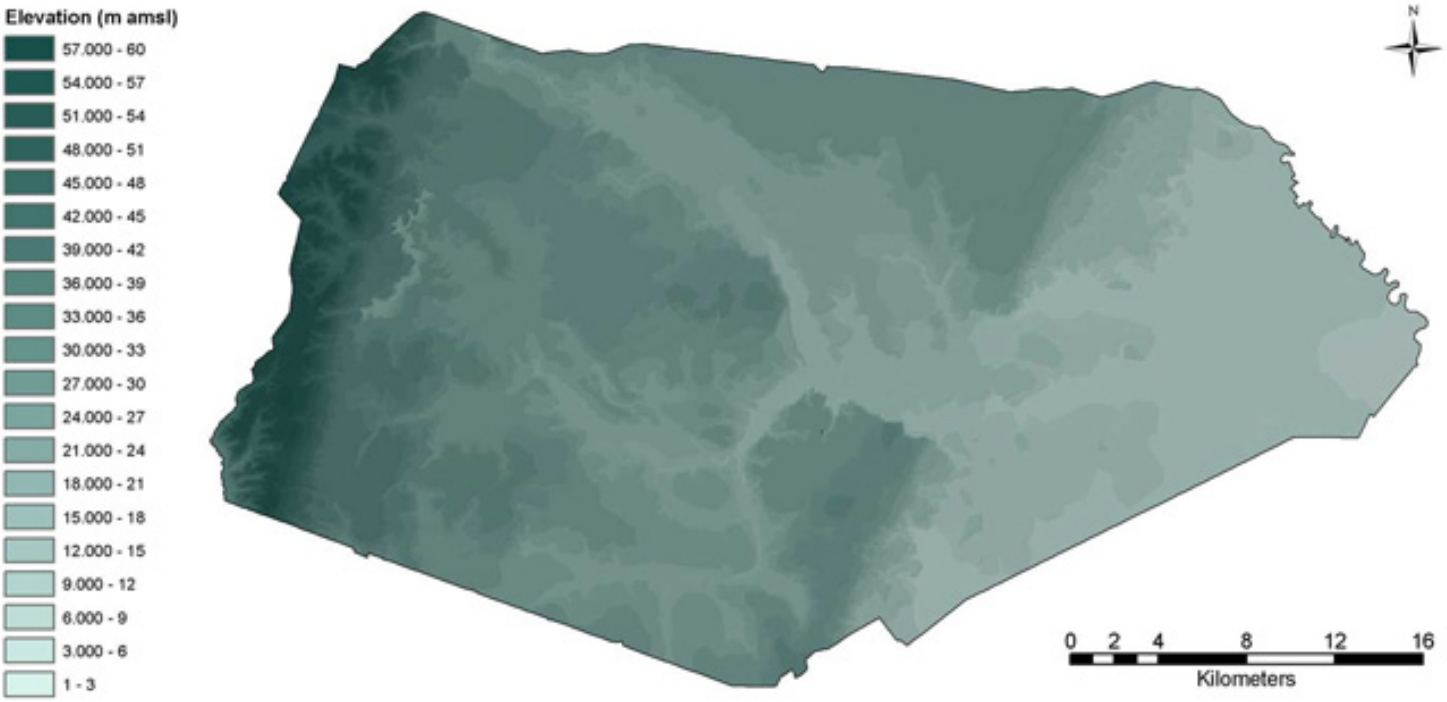

Figure 2. Elevation TIN.

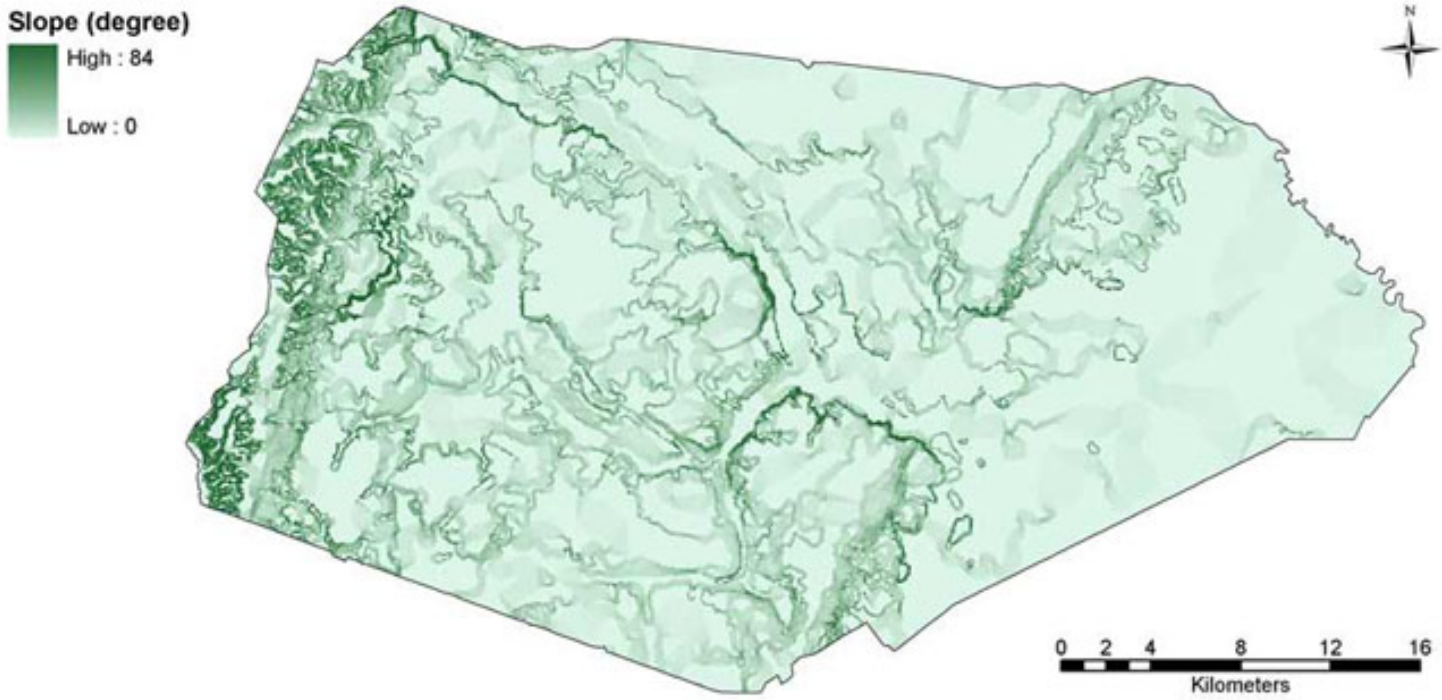

Figure 3. Slope.

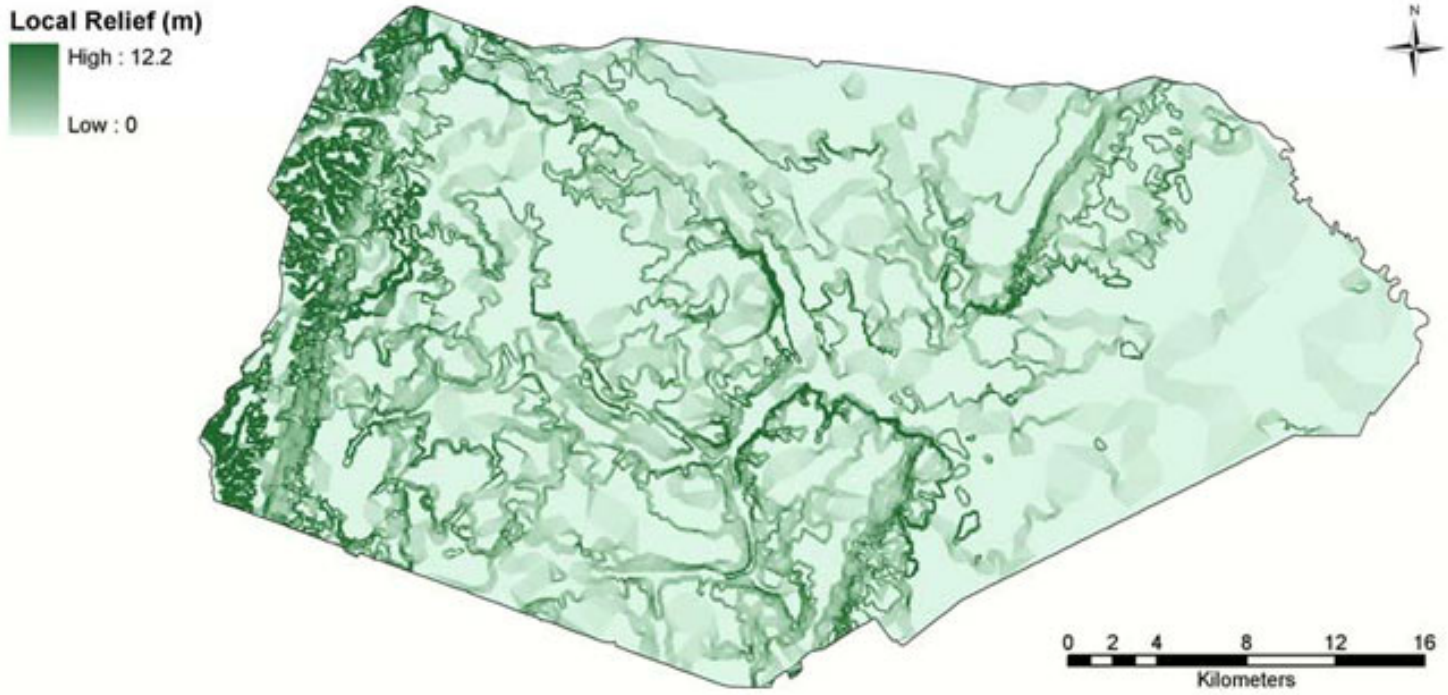

Figure 4. Local Relief (50-m Radius).

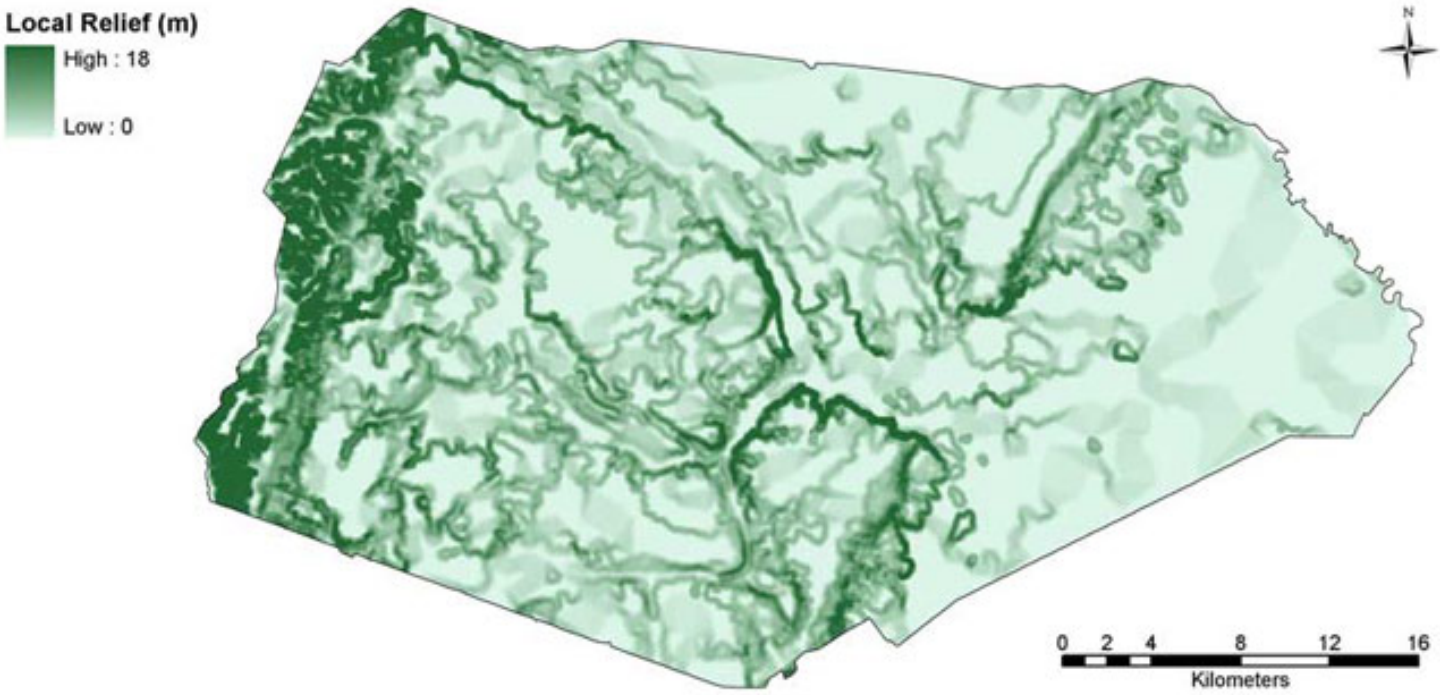

Figure 5. Local Relief (150-m Radius).

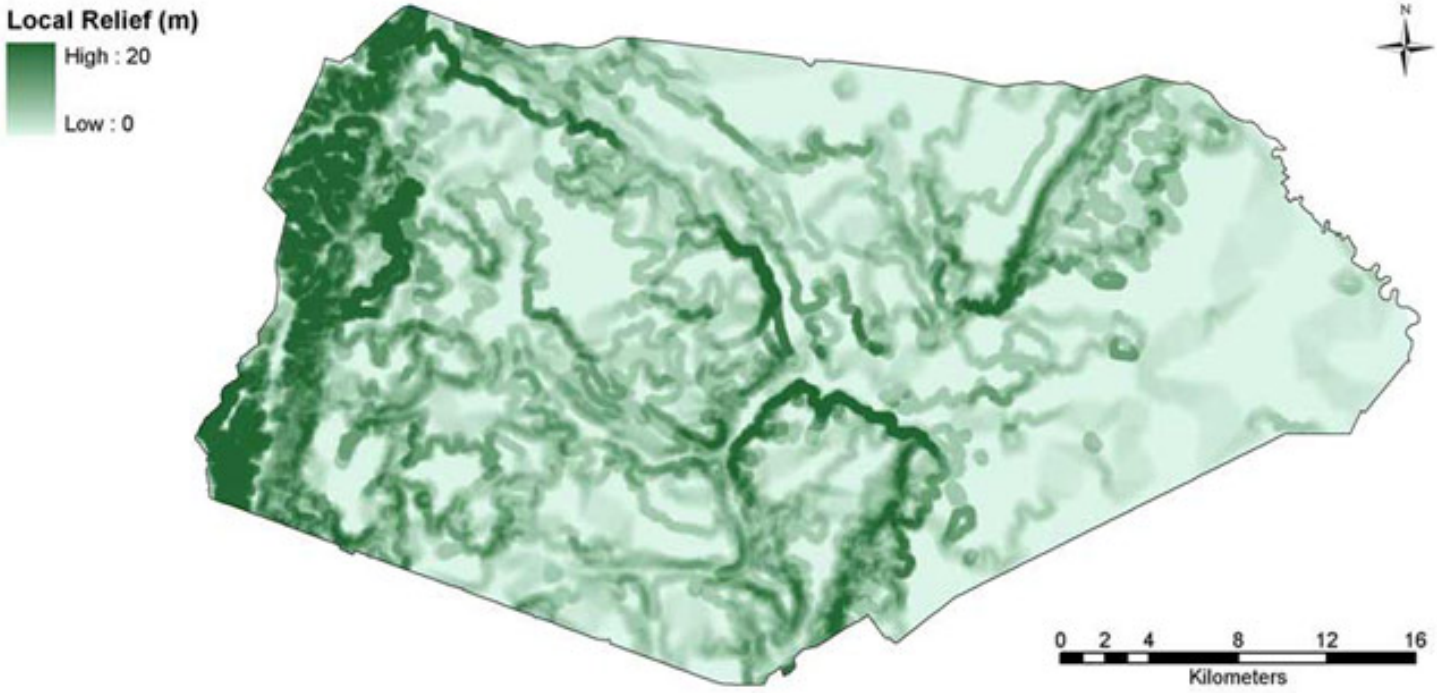

Figure 6. Local Relief (250-m Radius).

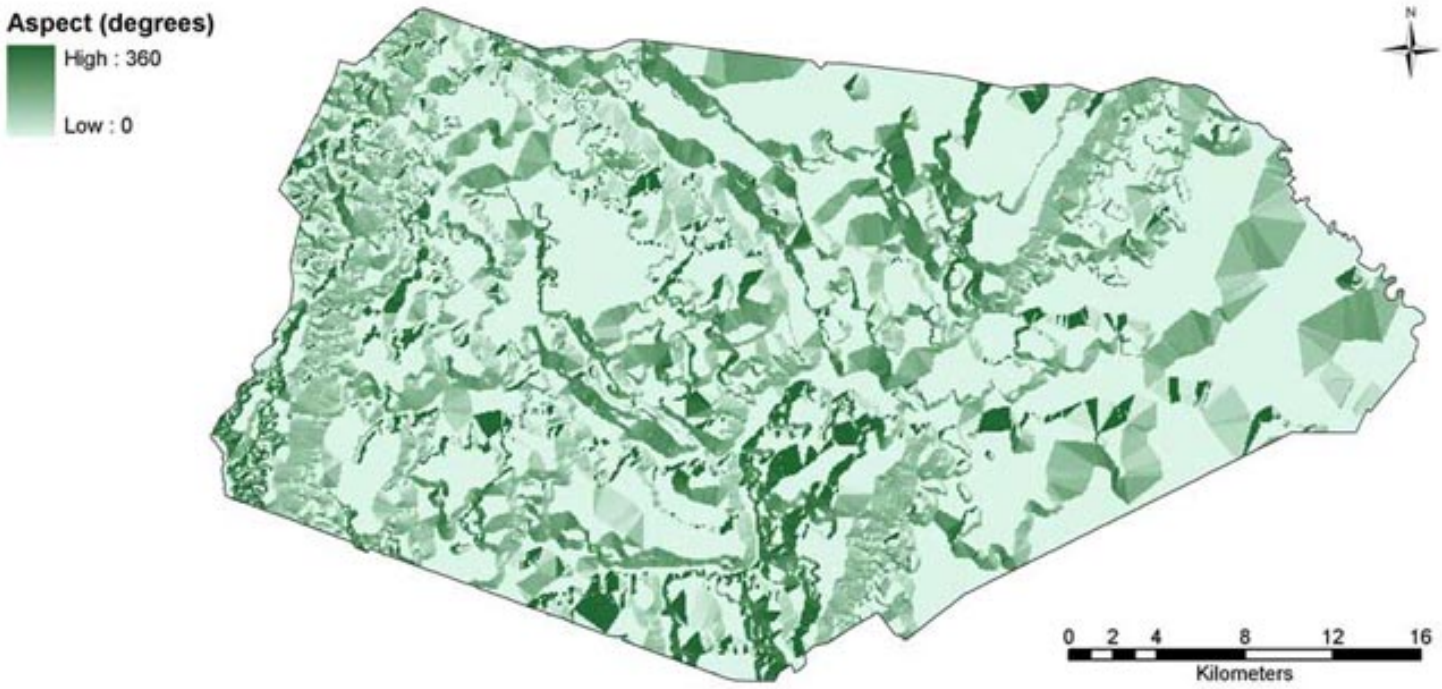

Figure 7. Aspect.

In addition to the elevation, slope, aspect, and local relief, two other environmental variables were used in the analysis of prehistoric and historic site distributions. The first was digitized using the editor within ArcMap, and depicts the major surface waters of the installation that were thought to have been present for at least the last 5000 years (Figure 8). Water channels where more than 2 m of downcutting occurred were selected on the TIN map and located on georeferenced aerial photographs and UTM maps of the installation, all of which were used as guides for digitizing the boundaries of this polygon feature class. The second variable, soil permeability, was derived from digitized Department of Agriculture (USDA) soil maps provided by the installation (Figure 9). In order to obtain permeability data, the actual soil survey manuals for the study area were obtained from the USDA's Soil Conservation Service, and the values were entered into a specially created attribute field for each soil type. Using the Geoprocessing Wizard's Dissolve Tool in ArcMap, soil types were merged based on eight different permeability classifications, which ranged from Excessively Drained to Very Poorly Drained.

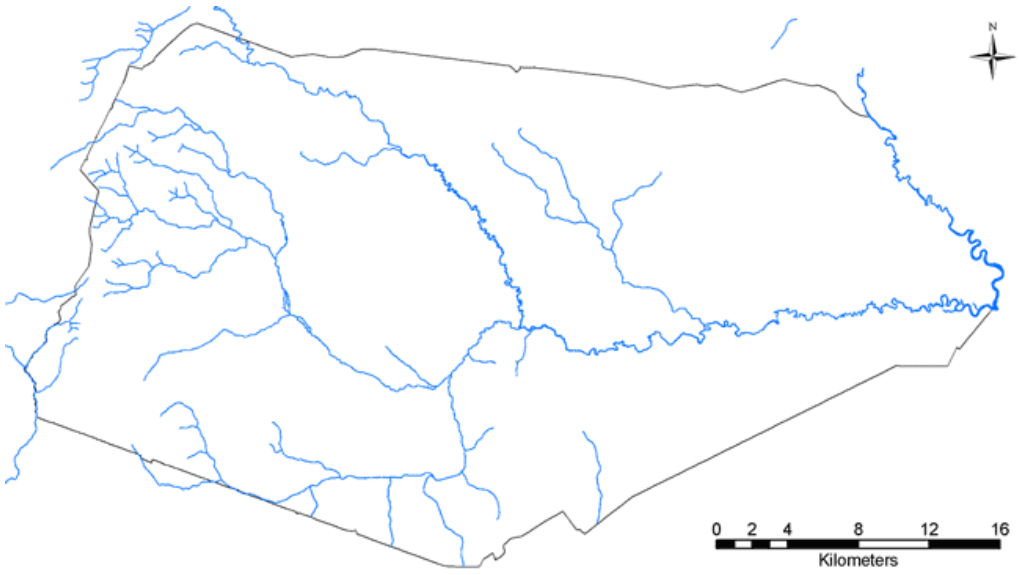

Figure 8. Major Surface Waters.

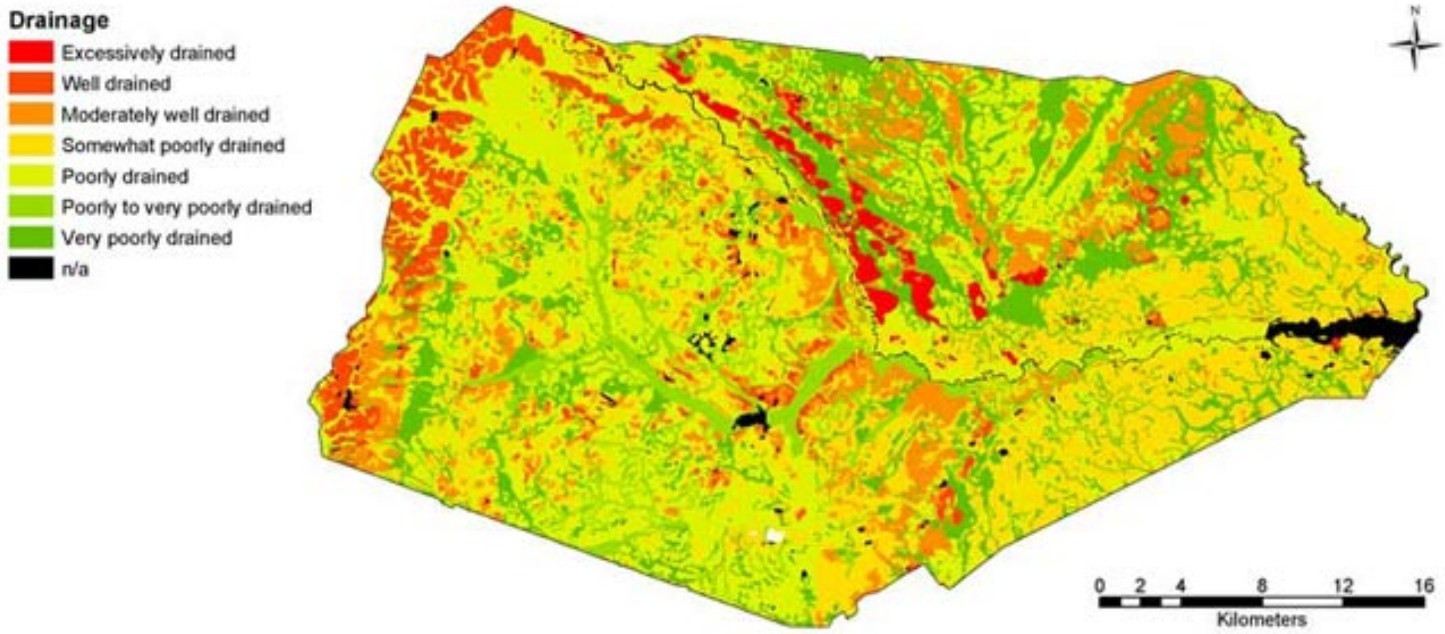

Figure 9. Soil Permeability.

Historic site distributions were also measured with regards to three cultural variables. These included distance to roads, distance to historic communities, and distance to historic schools. Figure 10 illustrates the location of historic communities and historic schools within installation boundaries, overlain on a density map of all historic sites. Roads are depicted in Figure 11. All of these feature classes were buffered using the Buffer Wizard in order to create distance intervals for historic site distributions.

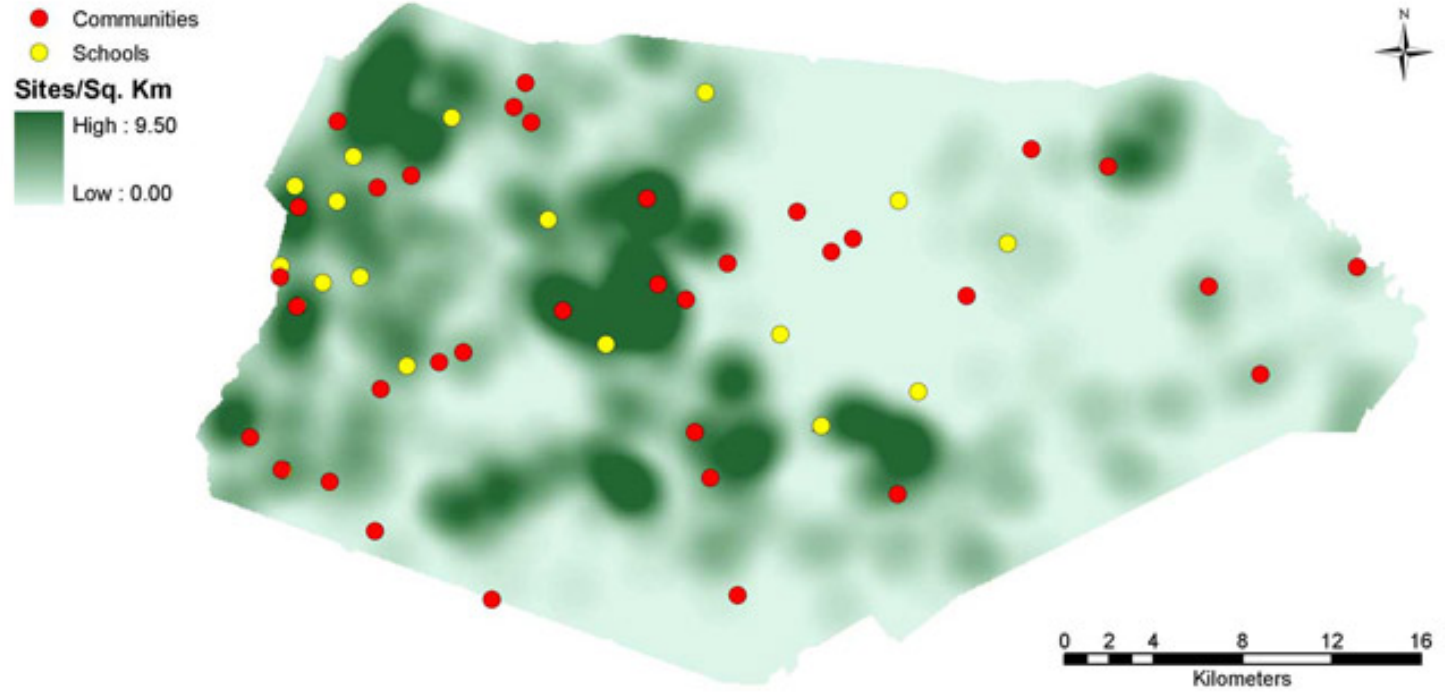

Figure 10. Historic Communities and Schools.

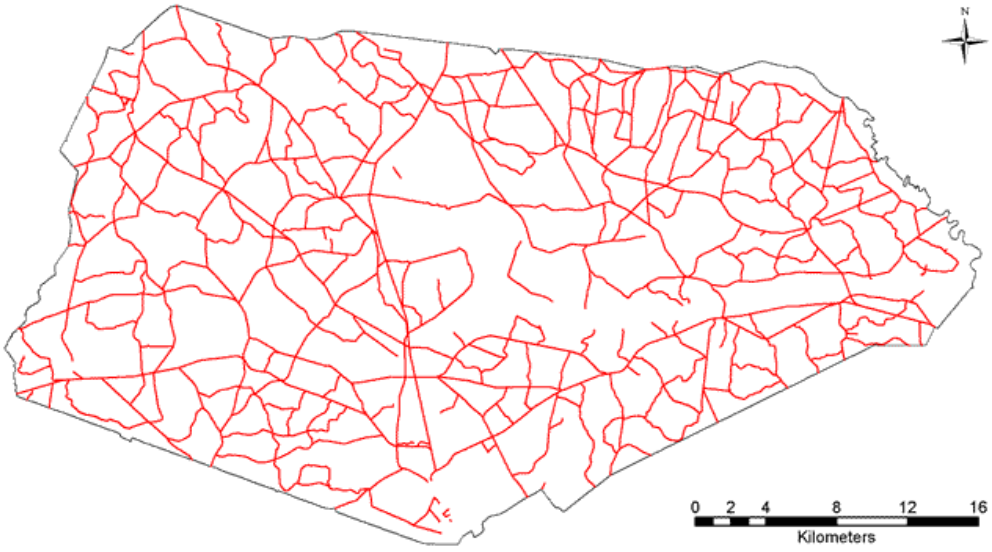

Figure 11. Roads.

In order to determine the areas favored by prehistoric and historic populations, a sample data set of actual sites located at the installation was used. To eliminate some of the bias inherent in conducting archaeological surveys, only zones that were investigated using shovel tests on a 30-x-30 m grid were included in the final analysis (Figure 12). The included area covered approximately 202 km^2 or 18% of the installation, and provided a good sample of all environmental and cultural variables under investigation. Within the shovel tested region, 336 prehistoric and 646 historic sites were located (Figure 13). Each of these data sets was kept separate, with the 336 prehistoric sites then being stratified into eight different non-exclusive feature classes based on the presence of lithic tools, ceramics, and multiple temporal components. This was done using the query builder in ArcMap with the hope of providing some level of explanation for the final model while still maintaining a large enough sample for actual analysis. Historic sites were not separated into multiple analysis layers since the majority of historic materials at the installation are from late 19th and early 20th century occupation.

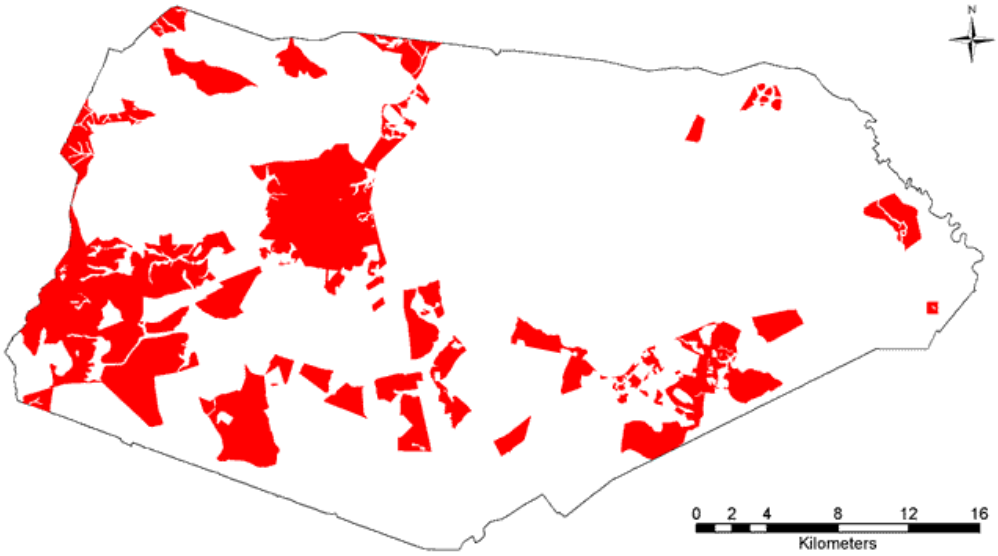

Figure 12. Surveyed Areas.

Figure 13. Prehistoric and Historic Sites.

The process of correlating site locations with environmental attributes was accomplished using two strategies. First, spatial joins were run in ArcMap connecting each prehistoric site to each of the eight environmental variables and each historic site to the eight environmental and three cultural variables. With regards to distance to water, roads, communities, and schools, the spatial join appended a distance field in addition to all of the attribute fields of the nearest relevant point, line, or polygon. For the remaining seven environmental variables, the spatial join appended the information from the polygon each site fell within.

The second strategy focused on the determination of the horizontal extent of environmental variable intervals. Values and distances obtained from the spatial join process discussed previously were studied using the Histogram, Trend Analysis, and General QQ Plot functions within the Statistical Analyst extension, with relationships between variables and spatial patterns being noted. From this information, intervals within each environmental and cultural variable were delineated based on natural breaks within the site distribution data (Figure 14). Using a polygon layer depicting the area surveyed by shovel tests on a 30-x-30 m grid (Figure 12), each environmental and cultural feature class was clipped to the actual survey boundaries by the Clip Tool in the Geoprocessing Wizard. Values within the environmental layers were merged into discrete intervals using the Dissolve Tool, and areas of each interval were then calculated to derive expected site numbers for the chi-squared analysis discussed below.

Distance (m) Water	Soil Permeability	Elevation (m)	Slope (deg)	Aspect (deg)	Local Relief 50 m	Local Relief 150 m	Local Relief 250 m
11	4	11	8	9	9	9	9
0-149.99	Excessively to well	2-6	0	none	0	0	0
150-249.99	Moderately well	6.01-11	0.01-0.184	0.01-45	0.01-0.2	0.01-0.6	0.01-0.9
250-399.99	Somewhat poorly	11.01-16	0.18401-0.396	45.01-90	0.21-0.7	0.61-1.6	0.91-2.3
400-549.99	Poorly to very poorly	16.01-21	0.39601-0.649	90.01-135	0.71-1.2	1.61-2.6	2.31-3.6
550-749.99		21.01-26	0.64901-0.931	135.01-180	1.21-1.7	2.61-3.7	3.61-5.0
750-1049.99		26.01-31	0.93101-1.370	180.01-225	1.71-2.4	3.71-5.1	5.01-6.6
1050-1549.99		31.01-36	1.37001-1.880	225.01-270	2.41-3.4	5.11-7.2	6.61-8.3
1550-1999.99		36.01-41	>1.880	270.01-315	3.41-5.4	7.21-9.5	8.31-10.1
2000-2799.99		41.01-46		315.01-360	>5.4	>9.5	>10.1
2800-3899.99		46.01-51					
3900-5450		51.01-56					

Distance (m) Roads	Distance (m) Communities	Distance (m) Schools
12	14	14
0-49.99	0-399.99	0-499.99
50-99.99	400-699.99	500-1099.99
100-149.99	700-999.99	1100-1599.99
150-199.99	1000-1299.99	1600-1899.99
200-249.99	1300-1499.99	1900-2299.99
250-349.99	1500-1799.99	2300-2599.99
350-449.99	1800-2099.99	2600-2899.99
450-549.99	2100-2499.99	2900-3299.99
550-649.99	2500-2899.99	3300-3599.99
650-749.99	2900-3299.99	3600-4099.99
750-999.99	3300-3699.99	4100-4799.99
>999.99	3700-4399.99	4800-5699.99
	4400-5599.99	5700-7999.99
	>5599.99	>7999.99

Figure 14. Environmental Variable Intervals.

The need for a relatively simple technique to use in the evaluation of significant correlation influenced the decision to implement the 1-sample chi-squared test in determining areas of high and low probability (Figure 15). Observed site numbers were obtained for each variable interval from the spatial join process described above, while expected site numbers were calculated by dividing the interval area by the total area surveyed and multiplying by the total number of prehistoric or historic sites. Based on the significance of the calculated chi-squared values, areas of high and low probability could be determined in conjunction with simple Observed/Expected calculations. Values significant at a level less than or equal to 0.05 with Observed/Expected ratios greater than 1 were considered to be high probability while those with Observed/Expected ratios less than 1 were considered to be low probability. To measure the relative strength of each environmental variable, chi-squared values from each interval were summed and divided by the total number of prehistoric or historic sites. These weights were used in the raster analysis to determine the final probability models (Figure 16).

$$X^2 = \sum \frac{(O - E)^2}{E}$$

X^2 = Chi-Squared
O = Actual # of Sites
E = Expected # of Sites

Figure 15. Chi-Squared Formula.

Variable	Prehistoric Weight
Soil Permeability	0.3387
Elevation (m)	0.1941
Slope (deg)	0.1692
Local Relief - 50 m	0.1663
Local Relief - 150 m	0.0981
Distance (m) Water	0.0336
Aspect (deg)	0
Local Relief - 250 m	0

Variable	Historic Weight
Soil Permeability	0.4244
Distance (m) Roads	0.1676
Elevation (m)	0.0902
Distance (m) Communities	0.0882
Distance (m) Schools	0.0515
Local Relief - 250 m	0.0427
Slope (deg)	0.0370
Local Relief - 150 m	0.0367
Local Relief - 50 m	0.0329
Distance (m) Water	0.0170
Aspect (deg)	0.0119

Figure 16. Environmental Variable Probability Weights.

The final raster analysis consisted of the differential weighting of environmental variable intervals, and the adding of these variables together to create zones of high, medium, and low probability. This was accomplished by taking the original rasters of each environmental variable, including a raster representing buffered intervals of distance to water, roads, communities, and schools, and reclassifying intervals with the new calculated weights using Spatial Analyst. High probability zones were given positive weights, low probability zones were given negative weights, and zones with no significant positive or negative correlation were not weighted. The reclassified rasters were then added together in the raster calculator to yield raw probability values for both the prehistoric (Figure 17) and historic model (Figure 18). Based on natural breaks within the distribution of raw values, each model was divided into areas of high, medium, and low probability for containing sites.

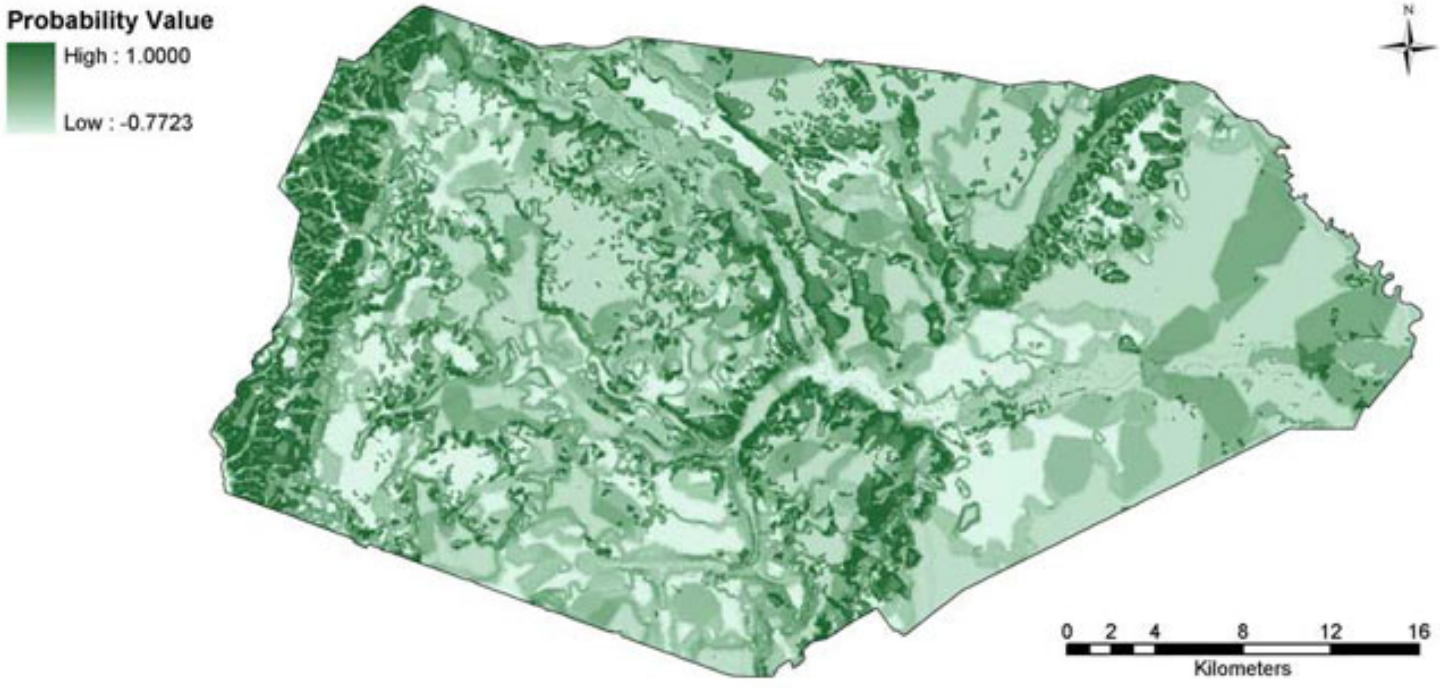

Figure 17. Prehistoric Probability Model.

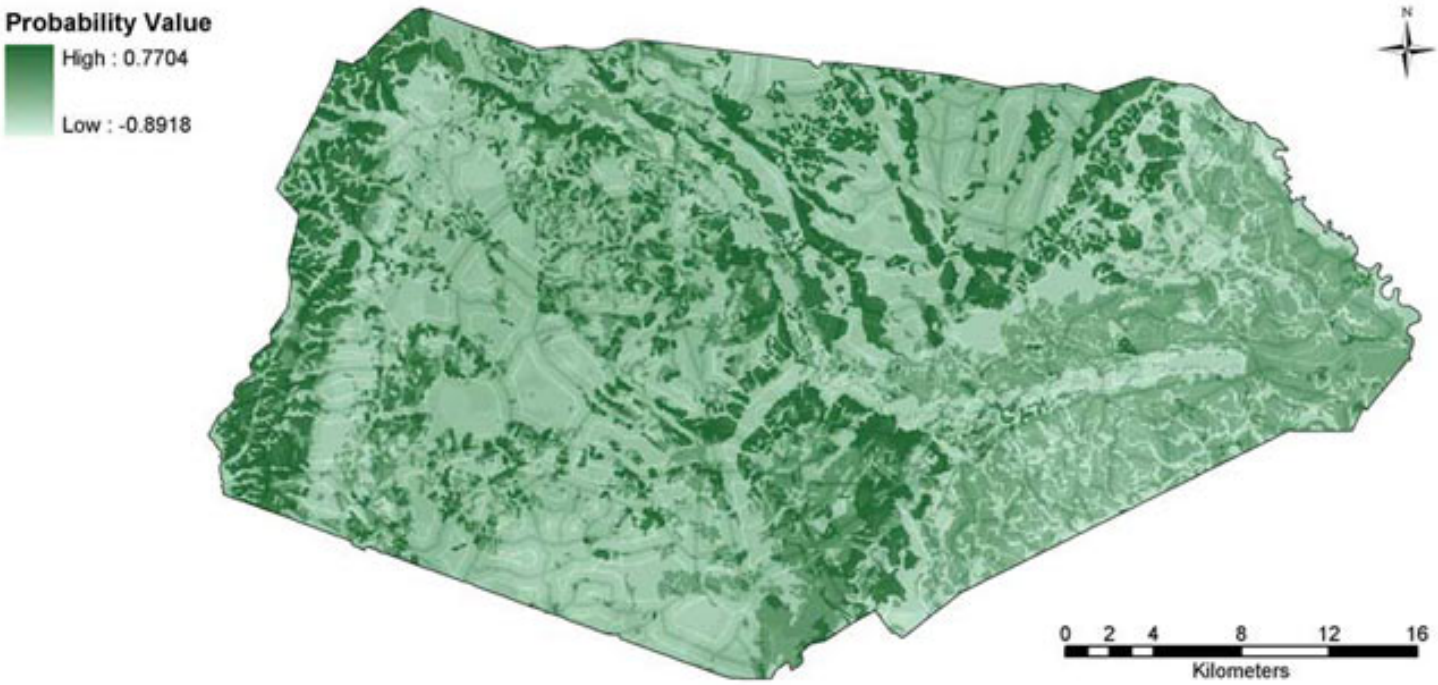

Figure 18. Historic Probability Model.

Due to the necessity of creating a probability model that accounts for both prehistoric and historic materials, the two models needed to be combined into a single depiction of probability areas at Fort Stewart. To accomplish this, probability zones in each were grouped using the classifications shown in Figure 19. Areas to be surveyed as high probability in the final model included all zones that were high probability for containing either prehistoric or historic materials and zones that are medium for containing both prehistoric and historic materials. The remaining zones were considered low probability. Once again, the raster calculator was used to combine the reclassified values yielding the raw probability values depicted in Figure 20. The final probability map, overlain on elevation data exaggerated at a scale of 30X and viewed from a vertical angle of 30°, is shown in Figure 21. This final map was created using ArcScene, the illustration and animation component of the 3D Analyst extension.

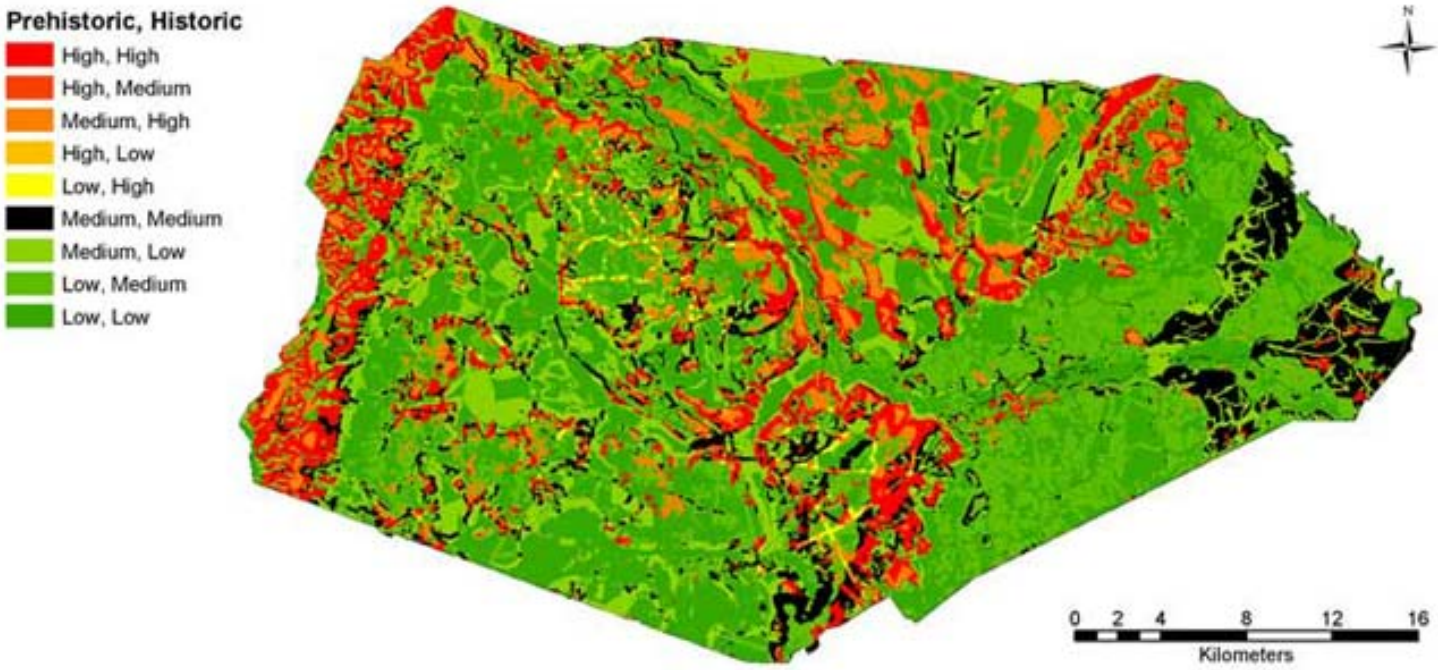

Figure 19. Final Probability Classifications.

Historic Probability			Prehistoric Probability: High (7)	Prehistoric Probability: Medium (4)	Prehistoric Probability: Low (1)
High	1		8 High 12845.23 0.1136	5 High 8635.62 0.0764	2 High 700.68 0.0062
Medium	0		7 High 441.5 0.0039	4 High 11859.57 0.1049	1 Low 20894.04 0.1848
Low	-1		6 High 439.28 0.0039	3 Low 18081.2 0.1599	0 Low 39187.01 0.3465

Weight (W)
Survey Probability
Area (ha)
Part Total

Figure 20. Final Probability Values.

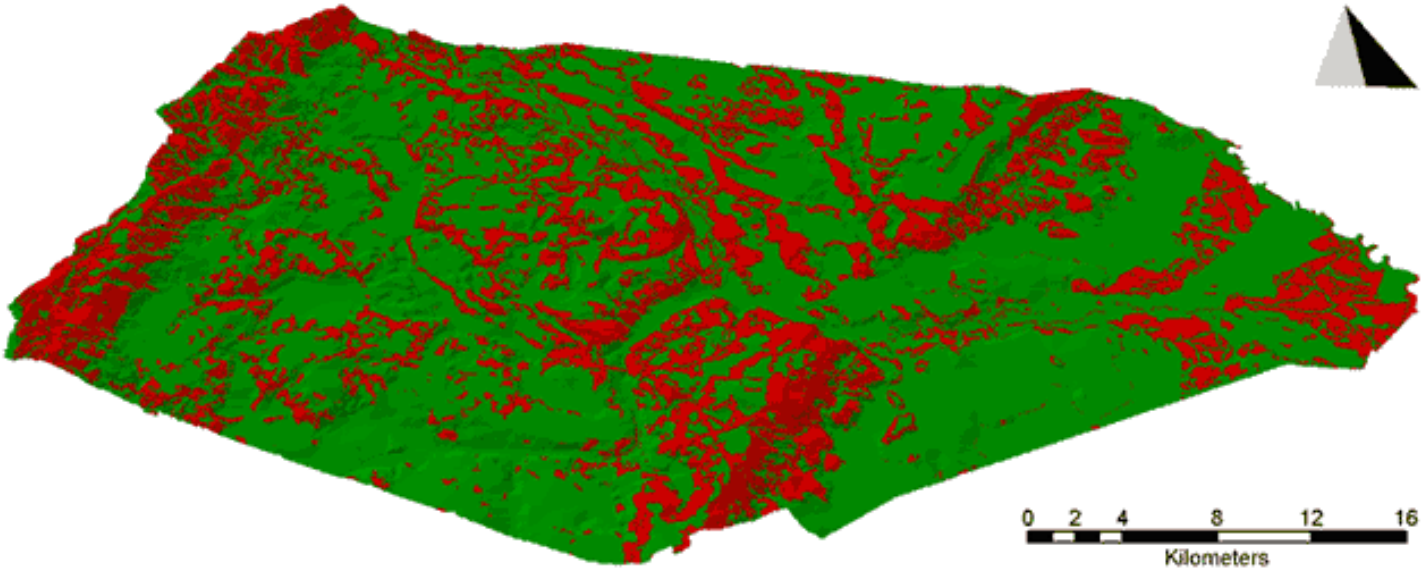

Figure 21. Final Probability Model (High Probability=Red).

It should be mentioned that all data was either imported into or created within a personal geodatabase. This includes the original Access database that contained specific site information, such as report reference, UTM coordinates, site size, temporal data, and survey data. By combining vector and tabular data within the geodatabase, the amount of available information for each archaeological site was increased significantly.

RESULTS

The final probability model described here will be used to determine the survey coverage for resource management units at the installation. Based on summary statistics obtained from ArcGIS, approximately 31% of the installation should be surveyed on a 30-x-30 m grid while the remaining 69% should be surveyed on a 45-x-45 m grid. This does not take into account areas that have already been investigated or areas that are off-limits due to the presence of unexploded ordnance. For both the prehistoric and historic models, soil permeability turned out to be the most important factor in influencing the presence or absence of archaeological sites (Figure 16). Areas that contained Moderately Well Drained to Excessively Well Drained soils were selected for, while areas with Somewhat Poorly Drained to Very Poorly Drained soils were selected against.

In the prehistoric model, elevation, slope, and local relief at a 50-m and 150-m radius also played an important part. Elevation was investigated due to its relation to five different marine terrace surfaces present at the installation (Figure 22), and there appears to be a general trend towards the location of sites along the edges of

these terraces. Slope and local relief also act as important indicators of zones along these marine terrace margins. Distance to water played only a minor role in determining site location, as the other variables discussed above appeared to be better indicators of whether or not sites were found. The one exception involves the correlation between distance to water and the feature class depicting prehistoric sites with diagnostic ceramic materials. In general, this feature class is representative of larger, multi-use sites that often show evidence of continuous occupation. The wide variety of activities that occurred at these locations contrasts sharply with activity specific areas represented solely by the presence of small undiagnostic lithic or ceramic distributions, for which distance to water does not play an important role. Although only briefly touched upon, the stratification of the prehistoric record into multiple feature classes for analysis made it possible to offer greater resolution regarding pre-European occupation at Fort Stewart and the Georgia Coastal Plain in general.

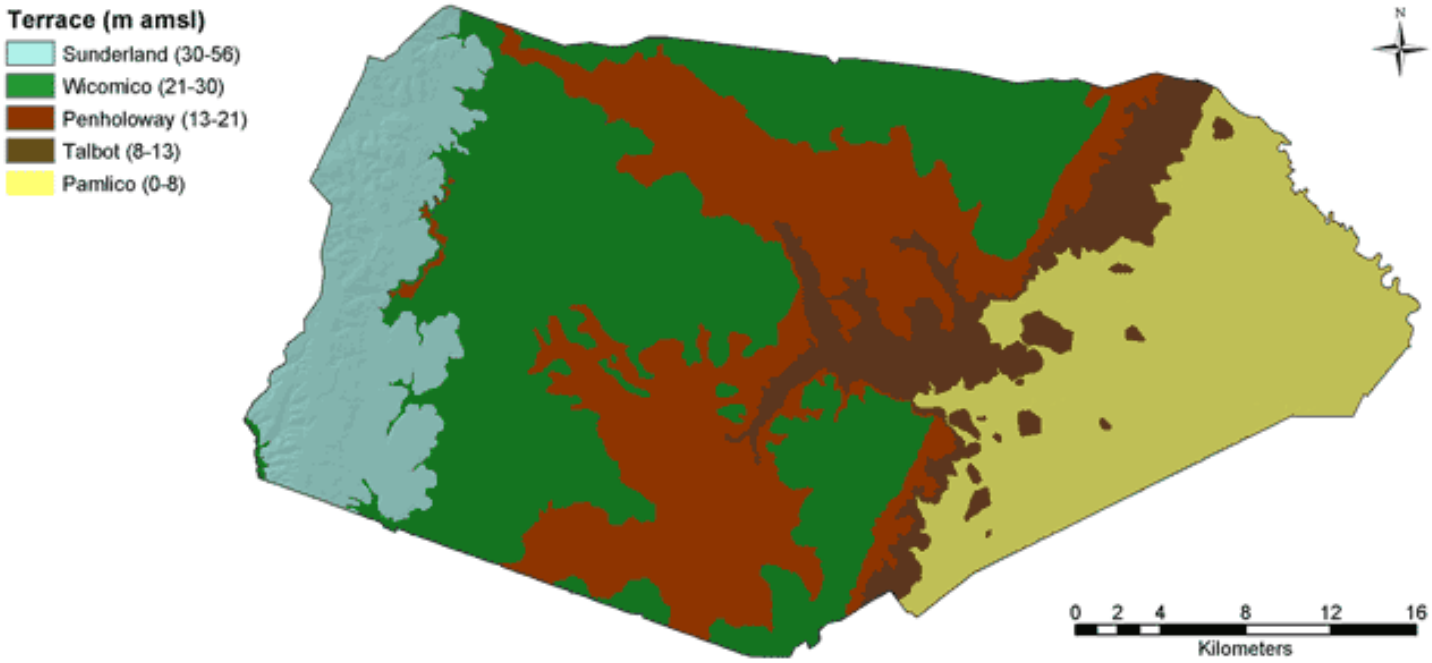

Figure 22. Marine Terrace Surfaces.

As mentioned previously, historic sites tend to be located on Moderately Well to Excessively Drained soils. In addition, they show some inclination towards areas close to roads, however little importance is placed in variables such as slope or local relief. Aspect plays no role in the determination of site locations for either model, which is to be expected in an area as flat and heavily wooded as Fort Stewart. When compared to prehistoric distributions, the greater importance placed on soil permeability by historic occupants makes sense in the context of the relative mobility of the two sample populations. Permanent settlers in a region would be expected to place a higher value on well drained soils outside of wetland areas and floodplains in order to protect their homes, schools, churches, cemeteries, communities, and livestock from seasonal flooding. Prehistoric site distributions are more representative of transient populations of hunter-gatherers, with limited evidence existing for permanent occupation. The secondary importance of roads to historic settlers suggests that transportation networks were integral to maintaining communication between regions both inside and outside the boundaries of Fort Stewart. In addition, proximity to these transportation arteries would be important to inhabitants who place a high value on the movement of goods to and from market areas.

CONCLUSIONS

The transition from a probability model based purely on a single environmental attribute to one that incorporates multiple environmental and cultural variables will provide cultural resource managers with a powerful tool for preserving the integrity of the archaeological record at Fort Stewart. Work in the near future will focus on the testing of the revised probability model, with preliminary indications suggesting a good fit with the actual archaeological record. The current project's reliance on the power and versatility of ArcGIS helped to bridge the

gap between data collected in the field and the process of explanation necessary for determining archaeological significance. It has also helped to provide a framework within which future investigations at Fort Stewart Military Installation can occur.

ACKNOWLEDGEMENTS

The author would like to thank Jennifer Grover and Brian Greer of Fort Stewart Military Installation, and David Anderson of the Southeast Archeological Center, all of whom provided the opportunity, background, and encouragement to complete this project.

AUTHOR INFORMATION

Seth Johnstone
Contracting Archeologist - All Points Logistics, Inc.
Southeast Archeological Center
National Park Service
2035 East Paul Dirac Drive
Johnson Building, Suite 120
Tallahassee, Florida 32310
850-580-3011 ext. 243 office
850-580-2884 fax
Seth_Johnstone@nps.gov

Return to the Top