APPENDIX L. DEVELOPMENT AND TESTING OF THE CULTURAL RESOURCE SITE DENSITY PREDICTIVE MODEL

A model of cultural resource site density was developed as a means of estimating the general densities sites involved in management decisions that involve landscapes. This model was developed following techniques utilized by other researchers in the region for estimating site densities. The goal of the model is to be able to estimate whether large or moderate numbers of sites might be expected within a given area of the landscape. The model is not designed to predict specific site locations (or non- locations). Nor is the goal of the model to determine that certain portions of the landscape may or may not be used in any particular way. The goal is to have a mechanism for assessing relative site densities. The model supplements, but does not replace, the existing knowledge held by FO specialists, who make land use decisions based on site-specific knowledge. Furthermore, it is important to note that the model is not 100% accurate; no archaeological site prediction model can achieve perfect accuracy. For the purposes of assessment it was determined that if the developed model could predict site densities with at least a 70% correct classification rate it would be acceptable for the purposes of the impacts analysis. A 70% correct classification rate can be considered conservative. This rate was obtained in a similar study (Tipps et al. 1988:125), and is higher than the rate achieved in a variety of other studies (see Tipps et al. 1988:125). This rate should be sufficient for the general purposes of this analysis in comparing the *relative* impact of one alternative versus another. Importantly, consistent application of the same model ensures that impacts analysis is replicable and consistent. This document provides detailed information about the development and testing of the model.

L.1 DEVELOPING THE MODEL

The model for predicting relative cultural resource site densities was developed by employing discriminant function analysis of environmental variables to develop a prediction of site density for a given block of land. This technique has been employed in multiple similar studies (see Tipps et al. 1988). Notably, given recent developments in GIS technologies, it was possible to utilize GIS data to produce information on environmental data. The model was developed to predict locations of the most common types of sites in the area, predominantly prehistoric sites, but including historical sites other than linear resources such as roads, railroads, or canals.

Discriminant function analysis is a statistical procedure that utilizes variables to produce linear functions that result in the maximum separation on a statistical basis between two or more groups defined by the user (see Tipps et al. 1988:115-118). Although there are many variations on the statistical procedure, in essence, discriminant function analysis is designed to determine which variables, selected by the user, can be used to separate two or more groups and produce a linear function that can be used to assign new sets of variables into the same groups.

In the case of the cultural resources density prediction model, the goal was to distinguish areas of the landscape on the basis of site density. Groups were defined on the basis of numbers of sites within a given unit of land. A large number of environmental variables were entered into the discriminant function analysis program. The program first determines which variables can be used to account for variation between the groups, eliminating those variables that do not contribute to differences between the groups. Then, the program produces a formula consisting of coefficients multiplied to relevant variables and added together with a constant. This formula produces a value for each group, and in a given area, the highest value is used to assign a group. In other words, the formula can be used to place new units of land into one of the site density categories on the basis of the variables determined to be relevant.

Following a procedure developed for an area north of the planning area (Tipps et al. 1988), discriminant function analysis was utilized to place blocks of land into "medium" and "high" site density areas. In previous studies, 160-acre quadrats were used as the land unit. This size was selected because previous studies have indicated that it is a viable and useful size for classifying site density (Tipps et al. 1988:118-121). One of the crucial aspects of discriminant function analysis is the choice of original groups -in this case defined site densities -used in the analysis. In previous studies, three groups were defined—quadrats without sites (or low site density), quadrats with sites (medium site density), or quadrats with two or more sites (high site density).

While we attempted to maintain this tripartite distinction, it was determined that there were almost no 160-acre quadrat areas within the planning area with no sites (low site density). Attempts were made to model different definitions of low, medium, and high site density, but these attempts consistently failed to meet the 70% classification success desired for the model. In part, the failure to be able to accurately define site density groups results from high variation in numbers of sites per 160-acre quadrat in the planning area. Study data from previous cultural resource surveys indicated a mean of 6 sites per quadrat, with a standard deviation of nearly 6. In other words, two-thirds of the quadrats have between 1 and 11 sites. In one sense, 1-11 sites could be considered "average" or "medium" site density in the planning area. However, any more than 4 sites in a 160-acre quadrat could be considered to present management challenges that are above average. Because the goals of the model were to provide a management tool and a relative means of comparing alternatives, it was decided that the model could be acceptably used with a definition of "medium" density as one site per 160-acre quadrat and "high" as 2 or more sites per 160-acre quadrat, as with previous studies. The model is designed to distinguish between areas that will have few to minimal cultural resource management issues and areas that will clearly have some, and potentially many, management issues.

The model was developed by applying environmental variables (Table L.1) developed through GIS data to areas that had been previously inventoried for cultural resources. The areas selected for model development were taken from 14 cultural resource inventories ranging in acreage from 85 acres to 1180 acres, and located within the broader Moab and Monticello FO area encompassing environments similar to the entire field office area (Table L.2). Survey dates for these inventories ranged from 1979 to 2005, with the majority inventoried in the late 1980s and 1990s. A total of 101 quadrats were present in these areas and used for the analysis. The variables were selected to include those previously used in similar studies (see Tipps et al. 1988:120-121) as well as other environmental variables (e.g. vegetation types, fauna, etc.) that might potentially be correlated with resources attractive to prehistoric humans and therefore be a potential correlate with the majority of archaeological sites in the area. The SPSS computer program was used to run the discriminant function analysis. Stepwise variable entry (52 maximum steps), with F-to enter (minimum partial F of 3.84), F-to remove (maximum partial F of 2.71), and Rao's V used as the selection criterion to enter variables (minimum of 0) was used following previous studies (Tipps et al. 1988:116).

Table L.1. Environmental Variables used in Developing the Cultural Resources Discriminant Function Model

Variable	Source of Data
Relief-Quadrat relief in meters. Defined as the difference between the maximum and minimum elevations within the quadrat	GIS query
Elevation -Sum of the maximum and minimum elevations in the quadrat divided by 2	GIS query
Distance to River -Distance to the nearest river in kilometers, measured from center of quadrat	GIS query
Distance to Water -Distance to nearest permanent water in kilometers, measured from center of quadrat	GIS query
Quadrat Cover -Percent of quadrat covered by pinyon-juniper vegetation	GIS query
Drainages – Number of drainages in the quadrat	GIS query
Count of springs in quadrat	AGRC Springs Database
Acres of Cottonwood-Willow Community in quadrat	vgripn_new
Acres of Tamarisk Community in quadrat	vgripn_new
Acres of Box Elder Community in quadrat	vgripn_new
Acres of other lotic ecosystems in quadrat	vgripn_new
Acres of sagebrush in quadrat	Utah GAP Vegetation
Acres of grassland in quadrat	Utah GAP Vegetation
Acres of desert scrubland in quadrat	Utah GAP Vegetation
Acres of oak in quadrat	Utah GAP Vegetation
Acres of aspen in quadrat	Utah GAP Vegetation
Acres of Agrid Soils in quadrat	NRCS Utah Statsgo (soils)
Acres of Orthid Soils in quadrat	NRCS Utah Statsgo (soils)
Acres of Fluvent Soils n quadrat	NRCS Utah Statsgo (soils)
Acres of Orthent Soils n quadrat	NRCS Utah Statsgo (soils)
Acres of Borolls Soils n quadrat	NRCS Utah Statsgo (soils)
Acres of Xerolls Soils n quadrat	NRCS Utah Statsgo (soils)
Acres of mule deer habitat (winter or summer) in quadrat	BLM data
Acres of Rocky Mtn. Elk habitat (winter or summer) in quadrat	BLM data
Acres of Pronghorn habitat in quadrat	BLM data
Acres of Bighorn (either Rocky Mtn. Or Desert) habitat in quadrat	BLM data
Acres of sage grouse habitat in quadrat	BLM data
Acres of 100-year floodplain in quadrat	SGID024_Floodplains

Table L.2. Survey Areas Utilized to Develop the Discriminant Function Analysis Model

Name of Area	State Project #(s)	Year
Maverick Point	U-85-FS-026	1985
South of Moqui Canyon	U-90-Al-525	1990
Cedar Mesa-near Johns Canyon	U-80-BL-322	1980
Near Bluff	U-85-BL-019 U-86-BL-654 U-05-BL-112	1985 1986 2005
Big Bench	U-93-AS-110	1993
Lower Montezuma Creek	U-88-CH-645	1988
Near Hovenweep	U-90-Al-461	1990
South Cottonwood/ WhiteMesa	U-79-UC-233	1979
South of Monticello	U-88-GB-417 U-89-GB-662 U-92-GB-619 U-93-GB-502	1988 1989 1992 1993
Lisbon Valley	U-87-BL-244 U-88-AS-104 U-88-BL-255	1987 1988 1988
Dry Valley	U-93-WN-199	1993
Harts Point	U-91-LA-441	1991
Indian Creek	U-83-UD-239	1983
Lime Ridge	U-90-CH-552	1990

The discriminant function analysis ultimately indicated that four variables were relevant to distinguishing medium and high site density areas (Table L.3). These included elevation, percent of pinyon-juniper cover, number of acres classified as sagebrush vegetation, and the number of acres classified as orthid soils. The standardized function coefficients can be interpreted to indicate that percent of pinyon juniper and sagebrush acres account for much of the differences between groups. The variables differ from those selected in similar discriminant function analyses (Tipps et al. 1988:128), in that variables related to water (e.g. distance to nearest river, distance to nearest permanent water, number of drainages in the quadrat, etc.) are not included in the final discriminant function. Elevation and percent pinyon-juniper are, however, included in a similar fashion. The role of orthid soil acres is somewhat perplexing. Orthid soils are generally poor for agriculture, generally occurring in poorly drained basins and having high salt or calcic components. Their presence in the model may reflect a preference for not situating occupations on land that could be used for agricultural purposes, or they may reflect occupation of basin interiors, where orthid soils typically occur.

Table L.3. Variables Determined Relevant by Discriminant Function Analysis for Predicting Medium (1 site/l acre) and high (2+ sites/160 acres) Density Areas with Associated Coefficients and Constants (Eigenvalue =.560, Canonical correlation=.599)

	Standardized Function Coefficient	Discriminant Function Coefficient/Constant*	
Variable		Medium	High
Elevation	-.651	.048	.043
Percent Pinyon-Juniper	1.097	-.207	-.136
Sagebrush Acres	1.153	-.085	-.041
Orthid Soil Acres	-.748	.101	.069
Constant	n/a	-43.384	-36.428

*The formula for determining site density prediction is applied in the following manner: Medium site density = (.048xElevation)+(-.207xPercent PJ)+(-.085 x Sagebrush Acres)+(.101xOrthid Soil Acres)+(-43.384). High site density = (.043xElevation)+(-.136xPercent PJ)+(-.041x Sagebrush Acres)+(.069xOrthid Soil Acres)+(-36.428). After calculating the values for medium and high site density, whichever value is greater is used to assign the quadrat to that group. In other words, if the value for medium is 13.87 and the value for high is 8.63, the quadrat would be defined as having "medium" site density.

L.2 Testing the Model

The discriminant function equation generally had very good classification success (Table L.4). The model correctly predicted 81% of medium site density areas, and 82% of high site density areas. It incorrectly predicted 3 out of 16 medium site density areas as high-density areas and 15 out of 85 high-density areas as medium-density areas. The overall classification success was 83 out of 101 quadrats, or 82%.

Table L.4. Actual and Predicted Site Density Values with Percentages of Correct and Incorrectly Classified Sites for the 101 Quadrats used to Develop the Discriminant Function Model

		Predicted		Actual Totals
		Medium	High	
Actual	Medium	13 (81% correct)	3 (18% incorrect)	16
	High	15 (17% incorrect)	70 (82% correct)	85

The true value of a classification formula, however, is in predictive success on new quadrats and areas other than the ones used to develop the formula. In order to test the formulae, they were applied to a second data set consisting of 82 quadrats from different survey areas (Table L.5). The areas selected for model development were taken from 16 cultural resource inventories ranging in acreage from 160 acres to 1474 acres, and located within the broader Moab/Monticello FO area encompassing environments similar to the entire field office area. Survey dates for these inventories ranged from 1981 to 2005, with the majority inventoried in the late 1980s and 1990s. A total of 85 quadrats were present in these areas and used for the analysis.

Table L.5. Survey Areas Utilized to Test the Discriminant Function Analysis Model.

Name of Area	State Project #	Year
Recapture	U-90-AS-730	1990
Horsehead Point	U-90-CH-246	1990
West edge Bug Point	U-85-LA-738	1985
Devil's Canyon Campground	U-92-FS-653 U-94-FS-540	1992 1994
Mouth of Squaw Canyon	U-84-LA-803 U-86-LA-754 U-86-LA-755	1984 1986 1986
Hovenweep	U-90-AI-461 also NPS in NM	1990 unknown
Blue Hogan	U-96-CH-470	1996
Allen Canyon West	U-86-WC-909	1986
Allen Canyon East	U-86-WC-909	1986
Allen Canyon NW	U-86-WC-909	1986
Wooden Shoe Canyon	U-86-NH-836 U-87-WN-553	1986 1987
Peters Point	U-90-FS-262 U-90-FS-422	1990 1990
Lower Indian Creek	U-99-BL-565	1999
Lockhart Basin West	U-98-BL-460	1998
Lockhart Basin East	U-98-BL-460	1998
Natural Bridges	U-81-UC-439 U-87-NA-038	1981 1987

The formulae for medium and high site densities were then applied to the same variables derived in the discriminant function analysis within the new and independent survey areas. The independent test yielded lower, but still valuable, predictive success (Table L.6). A total of 58 of 82 quadrats were correctly predicted for an overall classification success of 71%. Notably, however, the model did incorrectly predict 14 (or 66%) of the 21 medium-density quadrats as high-density quadrats. However, this error is actually conservative for the purposes of management, as it ultimately predicts higher site density in areas that have medium site density.

Table L.6. Actual and predicted site density values with percentages of correct and incorrectly classified sites for the 85 quadrats used to test the discriminant function model.

		Predicted		Actual Totals
		Medium	High	
Actual	Medium	7 (33% correct)	14 (66% incorrect)	21
	High	10 (16% incorrect)	51 (84% correct)	61

L.3 SUMMARY

Overall, while the site density prediction model is by no means a perfect predictor of site density, it is sufficiently accurate to be utilized as a tool for analyzing potential relative involvement of cultural resource sites in management decisions. It has between a 70 and 80% success rate in defining 160-acre quadrats with 1 or 2 or more cultural resource sites. It is therefore utilized in analyses in the RMP as a means of gauging whether a particular alternative will involve more acres of high site density land than another, or whether an alternative will involve more acres of medium site density land. It is not utilized to predict numbers of sites involved in decisions. Furthermore, the model should not be considered a replacement for full inventory. As noted, in compliance with Section 106 of the National Historic Preservation Act, all specific actions with potential to involve cultural resources will be subject to intensive identification efforts such as cultural resources inventory. The discriminant function model is only used here to provide a means of supplementing the existing knowledge held by BLM and other resource specialists regarding known resources and high-density areas with a means of assessing relative site density in unknown or unsurveyed areas. The model is developed simply to provide a consistent and replicable means of assessing relative acres of high and medium site density areas involved in management decisions.

THIS PAGE INTENTIONALLY LEFT BLANK