A Predictive Probabilistic Model of Village Site Location
Within the Santa Ynez Valley, California

A thesis submitted in partial satisfaction of the
requirements for the degree Master of Arts
in Anthropology
by
Mark Lynton Neal

Committee in charge:
Professor Michael Glassow, Chair
Professor Michael Jochim
Professor Keith Clarke
Professor John R. Johnson

December 2007
The dissertation of Mark Lynton Neal is approved.

Michael Glassow, Committee Chair

Michael Jochim

Keith Clarke

John R. Johnson

December 2007
A Predictive Probabilistic Model of Village Site Location

Within the Santa Ynez Valley, California

Copyright © 2007

by

Mark Lynton Neal
ACKNOWLEDGEMENTS

It is surprisingly difficult to write a list of those people who have helped me during my time at UCSB. Virtually everyone I have met has helped me in some way and I feel a deep sense of gratitude every time I walk onto campus.

First and foremost I want to express my deepest gratitude for the unfltering love and support of my wife, Sheila Holland and my two daughters, Rachel and Vivian Cirricione. Whatever I have accomplished in the last ten years would not have been possible without the love of my family.

I would like to thank Mike Glassow for his wisdom, knowledge, and his always-open door; Keith Clarke for his guidance in the early phases of this research and for his consistent and infectious enthusiasm; Mike Jochim for sharing his insight into anthropological theory, and for his approachable good-nature; and John Johnson for his feedback and support.

I would also like to thank some of the numerous individuals who have helped in so many ways; Terry Joslin, Nico Tripcevich, Nathan Craig, Sean Bennison, Dustin McKenzie, Kristina Gill, Amy Gusick, Lisa Hosale, and Eric Marsh.

I would like to thank my fellow members of the Cultural Resource Defense Team of the Santa Ynez band of Chumash Indians; Kathy Conti, Sam Cohen, Karen Keever, and Freddie Romero; as well as the Elders Council for their support, understanding, and accommodation during the many busy times.
I would also like to thank the Department of Anthropology as a whole. The open hearts and brilliant minds of the faculty, staff, and students have created an atmosphere of cooperation and camaraderie that fosters learning and research to a degree that is uncommon in my experience.
A Predictive Probabilistic Model of Village Site Location
Within the Santa Ynez Valley, California

by

Mark Lynton Neal

A predictive probabilistic model is created of archaeological sites in the Santa Ynez River watershed. Twenty-one ethnohistorically documented Chumash village sites are selected as training points in order to assure that all sites are contemporaneous and of similar function. Various working hypotheses based on anthropological theory and past research are constructed regarding potential environmental influences on site location. A GIS is created of datasets constructed to provide data to test those hypotheses. Based on the results of those tests, it is determined that villages within the study area were located so as to be within 800 meters of perennial water, within 200 meters of ecotones, in locations with less than 15º ground slope, and in regions of relatively greater habitat diversity. A map is produced showing relative probability of regions to contain village sites in five ranks ranging from “Low” to “High.” The highest two sensitivity ranks combined correctly predict 80.9% of the test sites within 23.1%
of the study area. The highest rank predicts 71.4% of the test sites within 6.6% of the study area.
# TABLE OF CONTENTS

1.0—Overview ........................................................................................................ 1
   1.1—Inductive Versus Deductive Approaches to Modeling...................... 2

2.0—Cultural context of this study ................................................................. 5
   2.1—Information Sources .................................................................. 5
   2.2—Settlement Patterns ................................................................. 7
   2.3—Economic, Social, Political, and Ceremonial Organization .... 14

3.0—Research Objectives and Methodological Goals .................................... 18
   3.1—Methodological Considerations and Goals................................. 18

4.0—Prior Models of Settlement Patterns in the Santa Ynez Valley ............ 23

5.0—The Research Area ............................................................................. 28

6.0—The Study Sites .................................................................................. 35

7.0—Hydrography ....................................................................................... 37
   7.1—Background .............................................................................. 40
   7.2—Previous Research ................................................................. 40
   7.3—Hypotheses and Test Implications ......................................... 42
   7.4—Methods ................................................................................. 46
   7.5—Results .................................................................................... 62
   7.6—Conclusions ............................................................................. 65

8.0—Habitat Regions ................................................................................ 66
   8.1—Proximity to Ecotones ................................................................. 68
8.2—Proximity to Boundaries of Grassland/Chaparral Habitats ..... 69
8.3—Ecological Diversity ................................................................. 71
8.4—Hypotheses and Test Implications............................................ 72
8.5—Data and Methods..................................................................... 74
8.6—Methods .................................................................................... 78
8.7—Discussion of the Methods for Creating the Relative
    Environmental Diversity Index Grid .......................................... 80
8.8—Results ...................................................................................... 87
8.9—Conclusions .............................................................................. 88
9.0—Slope ........................................................................................................... 93
  9.1—Hypothesis ................................................................................ 94
  9.2—Results ...................................................................................... 94
10.0—The Model ................................................................................................ 96
  10.1—The Process of Mapping Archaeological Sensitivity Using
         Weights of Evidence.................................................................. 97
11.0—Conclusions ............................................................................................ 111
  11.1—Future Research Directions .................................................. 112
LIST OF FIGURES

Figure 5.1—The study area ...........................................................................34
Figure 7.1—Perennial and intermittent streams within the study area ..........55
Figure 7.2—Distance to perennial streams in the study area .......................58
Figure 7.3—Distance to intermittent streams in the study area ....................59
Figure 7.4—Distance to springs in the study area .......................................60
Figure 7.5—Distance to confluences in the study area ..................................61
Figure 8.1—Gap analysis dataset within the study area ................................77
Figure 8.2—Distance to ecotones within the study area ...............................90
Figure 8.3—Distance to Grassland/Chaparral boundaries within the study area..91
Figure 8.4—Relative Environmental Diversity Index (REDI) in the study area..92
Figure 9.1—Slope within the study area .......................................................95
Figure 10.1—Map of relative potential for primary habitation sites within the
              Santa Ynez Valley .............................................................................109
Figure 10.2—Map of relative potential for primary habitation sites within the
              Santa Ynez Valley showing all unique conditions ..............................110
LIST OF TABLES

Table 6.1—Village sites included in this study.....................................................36
Table 7.1—Distances between water sources and the village sites included in this study...................................................................................................57
Table 8.1—Mean Diversity Index and various radii.............................................83
Table 8.2—Results of unsigned residual testing ...................................................87
Table 10.1—Final model weights .......................................................................100
Table 10.2—Weights associated with Relative Environmental Diversity Index 101
Table 10.3—Weights associated with distance to ecotones dataset...............101
Table 10.4—Weights associated with distance to intermittent streams dataset..102
Table 10.5—Weights associated with distance to boundaries of grassland/chaparral regions ............................................................102
Table 10.6—Weights associated with distance to perennial streams..............103
1.0—OVERVIEW

Predictive modeling is a research method that has gained in prominence over the past 20 to 30 years in the science of archaeology. Predictive modeling in archaeology can be defined as a

Simplified set of testable hypotheses, based either on behavioral assumptions or on empirical correlations, which at minimum attempts to predict the loci of past human activities resulting in the deposition of artifacts or alteration of the landscape (Kohler 1988:33).

The fundamental research question in predictive modeling is; “How and for what purpose are decisions regarding location made?” (Kohler and Parker 1986).

Though predictive models are often identified with new technologies, they are not a new concept in the science of archaeology. They have their roots in regional settlement studies conducted in the 1950s and 1960s. Willey (1953) studied regional data from the Virú Valley in Peru in an effort to understand the processes underlying settlement patterns. His work provided the impetus for additional studies and is credited with establishing the value of regional studies
for understanding regional settlement systems (Trigger 1989). As a result, in the following decades some archaeologists shifted their methodological focus from the study of single sites to that of regions (Dalla Bona 1994; Trigger 1989).

One of the early trends in settlement analysis is the adoption of ecologically based explanations of cultural patterns. In the 1930s, Julian Steward studied the interaction between the environment and culture in the southwestern United States (Trigger 1989). Perhaps one of the most significant changes associated with the advent of regional ecologically based archaeological studies was the increased interaction that it brought between archaeologists and other disciplines of earth science. A research project by MacNeish in the 1960s in the highlands of Mexico funded by the United States National Science Foundation brought together archaeologists, zoologists, geologists, and other specialists (Trigger 1989)—a practice now relatively common in archaeological research.

The advent of computer-based geographical information systems (GIS) in the 1990’s fostered a significant increase the number of predictive modeling efforts. The impact of GIS was really twofold. It greatly increased the ease with which archaeologists could manage and analyze large geographic datasets. It also initiated a fluorescence of publicly available geographic data, especially high-resolution environmental datasets, which are available to archaeologists.

1.1—Inductive Versus Deductive Approaches to Modeling

Most researchers identify two general approaches to predictive modeling, inductive and deductive. Deductive models are based entirely on
“anthropological theory, previous findings, or ethnographic analog as a way to define variables relative to past locational behavior” (Kvamme 2006). Inductive approaches, sometimes termed “correlative” or “data-driven” approaches, are based solely on empirical observations of the archaeological record (Wheatley and Gillings 2002). It is important to note that whereas in concept it may be possible to create strictly inductive or deductive models, it is not necessary to separate these two approaches. Wheatley and Gillings (2002:166) have written that

> It should be recognized that a distinction between data and theory is not universally recognized, and most archaeologists accept that the two are not independent—data is collected within a theoretical context, and so may be regarded as theory-laden, while theories are based to some extent on empirical observations.

There are significant advantages to combining these approaches. The explicit use of anthropological theory and past research to inform the process of model building extends the value of the model beyond that of mere prediction to explanation of human behavior. It has been asserted that predictive models without explanatory power are of little value (Dalla Bona 1994). It can be argued
that the principal value of predictive probabilistic models is their ability to test hypotheses regarding site location and settlement patterns. Archaeological sensitivity maps are most accurately defined as one of the products of a modeling effort that can be a valuable tool for those involved in cultural resource management.
2.0—CULTURAL CONTEXT OF THIS STUDY

2.1—Information Sources

This research focuses on Chumash villages that thrived during the Late Period (A.D. 1170 – 1769) and into the Mission Period (A.D. 1769 – 1834). Researchers have used a variety of ethnohistoric, ethnographic, and archaeological data to describe this fascinating culture group. The story of the Chumash at the time of contact with Europeans is one of political alliances, social complexity, and artistic and technological achievements.

Our earliest ethnohistoric information regarding the Chumash comes from Spanish explorers who made brief intermittent visits to the California coast between the mid-16th century and the late 17th century. Early on, the purpose of those visits was exploring the frontiers of New Spain with the hope of discovering a new route to China and later, the goal was to find suitable ports for the Manila Galleon to stop at after its journey across the Pacific on its way to New Spain (Landberg 1965). The early Spanish explorers considered the Chumash to be truly exceptional among California native groups (Landberg 1965). Noted in journals left by members of various expeditions were the impressive size and permanence of Chumash settlements, the complexity and political integration of Chumash society, their focus on trade and commerce, and their considerable artistic and technologic achievements (Johnson 1988).

The missionization of Alta California began in A.D. 1769 which, over a period of less than 70 years, radically altered and virtually decimated Chumash
culture and lifeways. The Franciscans did maintain records of the births, marriages, and deaths of Chumash neophytes and gentiles, and provided answers to questionnaires from the civil authorities in Mexico (Landberg 1965). Together, these records have provided a wealth of information regarding specific family histories that can be traced to specific villages. In addition, post marital residence patterns have been reconstructed using village of origin information (Johnson 1988).

Beginning in the late 19th century, a few individuals in the Santa Barbara area took interest in Chumash culture and gathered ethnographic information during interviews with Chumash residents (Blackburn 1975; Landberg 1965). During that period, the first systematic archaeological research began. Of the early examples of archaeological investigation in the Santa Barbara channel region, perhaps the best-known was D. B. Rogers’ *Prehistoric Man of the Santa Barbara Coast*, in which he outlined a three-phase cultural sequence to describe variations in coastal sites along the Santa Barbara Channel (Rogers 1929).

J.P. Harrington began his ethnological research among the Chumash in 1912. Harrington produced an enormous body of notes concerning the Chumash, most of which were unpublished when he died in 1961 (Blackburn 1975:6). His notes have since supported a great deal of research into aspects of Chumash culture such as basketry, economic interactions, social integration, oral narrative, ceremonial integration, and ethnobotany (Blackburn 1975). It is noteworthy that Harrington’s informants were mostly recounting oral tradition at a time when
their culture had been significantly disrupted, giving rise to some discussion of the detail and accuracy of the information. Landberg stated that “there were Chumash Indians alive in the late 19th and early 20th century who had considerable knowledge of native customs handed down to them by oral tradition by their elders” (Landberg 1965:21). However, Hildebrandt noted confusion between the reports of various informants about the location of the Inezeńo village of Jonjonata (Xonxon’ata), leading him to conclude that “the memory of the place was nearly lost by the turn of the century” (Hildebrandt 2004:13). In other cases the locations of villages seem to been accurately handed down through oral history. For example, Harrington was able to locate and visit the site of Sotonocmu (Soxtonokmu’), though he may have had some prior knowledge of the site through the earlier expedition of Cessac, who was guided to the site by Rafael Solares sometime between 1877 and 1879 (McRae 1999:43).

Since Harrington’s time, archaeological inquiry has been the major source of information regarding Cumash subsistence and settlement patterns. Though a great deal more archaeological research has taken place at coastal sites, work at inland sites has also added significantly to our knowledge of Chumash culture.

2.2—Settlement Patterns

At the time of contact with Europeans, the Chumash people were comprised of several distinct groups connected in a widespread, well-articulated trade network. The Chumash geographic area at the time of contact with
Europeans was comprised of at least six linguistic regions, each of which had a
distinct dominant language: Obispeño, Purisimeño, Ineseño, Barbareño,
Ventureño, and Island. Considered together, the entire Chumash region included
the northern Santa Barbara Channel Islands, the mainland coast from Malibu to
Morro Bay, and extended inland to the edge of the San Joaquin Valley
(Hildebrandt 2004). The region also spanned a wide variety of ecological zones
and resources. Trade relationships between settlements in different ecological
zones may have served to provide a buffer against shortfalls created by seasonal
and cyclical variations in resource abundance (Johnson 2000).

The total Chumash population at climax is estimated to have been
approximately 18,500 (Glassow 1996:13). Chumash settlements were referred to
by the Spanish as rancherías. The term ranchería was never precisely defined by
the Spanish, but can be taken to mean a fairly sedentary population and a group of
surrounding camps that were seasonally occupied for the purpose of resource
extraction (Landberg 1965). In areas of high resource density, such as the Goleta
Slough, several large residential communities were located close together and
may have been part of the same ranchería (Landberg 1965). The largest villages
were located along the coast of the Santa Barbara Channel with populations
numbering between 500 and 800 people (Glassow 1996:14). Settlements in the
Purisimeño and Ineseño regions were generally smaller, with populations ranging
from 200 to perhaps as low as 30 (Hildebrandt 2004)
Previous researchers have hypothesized that subsistence and settlement practices in the Santa Barbara Channel and surrounding areas, gradually changed from that of foragers, who move their residences to follow food sources, to that of collectors, who deploy smaller logistical groups with the purpose of bringing back supplies to permanent or semi-permanent settlements (Glassow 1996; Woodman, et al. 1991).

Collectors produce five types of sites— residential bases, field camps, locations, stations, and caches (Binford 1980). Residential bases are habitation where the population lives at least part of the year. Residential bases can vary from large, permanent villages occupied year-round, to more mobile habitations that may only be occupied seasonally. Field camps are places where task groups eat, sleep and otherwise maintain themselves when away from the residential base. Stations are information-gathering posts like hunting blinds. Caches are where surpluses of collected material are stored prior to transport back to the residential base (Binford 1980).

Because of the change over time from forager to collector subsistence strategies over 8,000 plus years of occupation in the Santa Ynez Valley, it would be expected that the area would possess a wide variety of site types. Archaeological and ethnohistoric evidence supports that expectation. Site types that have been discovered include rock art locations, quarries, locations, field camps, villages, and sacred sites.
Rock art sites are generally assemblages of pictographs located within rockshelters in bedrock outcrops. Rock art sites are not habitation sites, but they are sometimes found near sources of fresh water and can be associated with bedrock mortars. Ephemeral ethnohistoric information indicates that rock art locations were associated with large villages and were places where ceremonial equipment was kept. The rock art panels served to instill a sense of awe and respect among the village members who were not educated in ceremonial practices. For example, a ceremonial dance skirt made of eagle and crow feathers was found in one rock shelter near a rock art site (Grant 1965).

Quarry sites are places where materials for tool and ornament manufacture are extracted. Generally, quarries are identified by the presence of extensivedebitage, but typically have little evidence of habitation.

Locations are perhaps the most commonly found sites in the Santa Ynez Valley. Following Binford’s description, they appear as thin scatters of food remains or lithic flakes (1980). They seem to represent single events of occupation or tool production.

Field camps represent episodes of generally seasonal occupation that were more intensive or longer than those of locations. Archaeological evidence of field camps differ from that of locations in that field camps tend to be larger and have more developed middens.

The village sites represented in this study were longer-term, multi-season or year-round occupation sites. Features of villages might include a cluster of
houses, a cemetery, a ceremonial/dance area enclosed by a wind break constructed of poles and woven mats, a semi-subterranean sweatlodge, a menstrual house, and storage and drying facilities (Gamble 1995). The locations of the village sites in this study were determined by John Johnson (1988:95) from a combination of mission records, ethnographic information regarding Chumash placenames collected by J.P. Harrington, archaeological evidence of historic period occupation, information included in land grant diseños, and field reconnaissance.

Though ethnohistoric records include primarily descriptions of coastal villages, the features mentioned in those descriptions have also been documented at sites in the Santa Ynez Valley. One of the first California archaeologists, Rev. Stephen Bowers, traveled to the Santa Ynez River between June and August of 1878. He documented the spatial arrangement of features at the villages of Snihauj (Shniwax) (CA-SBA-823) in Los Prietos Canyon and Snajalayegua (Shnaxalyi wi) (CA-SBA-1309) (Benson and Bowers 1997). Though Bowers did not employ modern techniques, he is said to have been a keen observer and also had the advantage that the sites he investigated were still at a relatively high degree of preservation. Evidence of many of the features described to ethnohistorians was still visible on the ground and therefore he was able to map much that has since been lost to the plow zone.

Ethnohistoric accounts of houses from the coastal area along the Santa Barbara Channel may be found in the journals of Fages and Costanso from the
first land expedition into the area in 1769 and from Font, who visited the region in 1776. In addition, Harrington collected ethnographic information. As synthesized by Gamble (1995), according to ethnohistoric and ethnographic information, houses were oval or circular in shape, ranged from 4 to 16 meters in diameter, had hearths in the center of the floor, a smokehole in the roof, and were constructed of arched poles covered by interwoven grasses. Houses were clustered together and sometimes organized in rows.

The map Bowers made of Snihauj (Shniwax) shows a cluster of circular features that he identified as house floors. His map of Snajalayegua (Shnaxalyiwi) (CA-SBA-1309) shows two such circular features (Benson and Bowers 1997). In 1965 and 1966, Donald Miller documented three structures at the village of Eljman (H’elxman) (CA-SBA-485) in the central Santa Ynez Valley. Each of these measured from 15 to 20 feet (4.6 to 6.0 meters) in diameter, were circular to oval in shape, had features identified as postholes around their circumference, and possessed central hearths. Macko (1983) hypothesizes that at least one of these structures has aspects that match ethnohistoric descriptions of sweatlodges. He also states that it may be true that houses at Eljman (H’elxman) were arranged in rows but further excavation would be required to test that hypothesis.

Ethnohistoric and ethnographic accounts of sweatlodges indicate that they were semi-subterranean, semi-spherical, earth-covered, with entrances through smoke holes in the roofs. Sweatlodges were ceremonial structures where fires
were built to create great heat to cleanse and purify men’s bodies and souls (Gamble 1995). Bowers’ maps show sweatlodges at both Snajalayegua 
(Shnaxalyiwi) and Snihausj (Shniwax) (Benson and Bowers 1997). More recent 
investigations at the coastal village of Mikiw (Gamble 1995) and at the inland 
village of Eljman (H’elixman) (Macko 1983) have uncovered remains with 
hardened clay floors, post holes in the center, and hearths with fire cracked rock 
that seems to match ethnohistoric accounts of sweatlodges.

Ethnohistoric and ethnographic accounts report that ceremonial dance 
areas and gaming areas were two other common features of Chumash villages. 
Both can be described as flat, open areas at least partially surrounded by 
windbreaks constructed of poles and mats. There are no recent accounts of 
arCHAeological investigations that have uncovered evidence of dance floors or 
gaming areas. However, during the late 19th century Bowers did map features at 
village sites in the Santa Ynez Valley that correspond closely to ethnohistoric and 
ethnographic accounts. His map of Schniwax shows a rock enclosure measuring 
75 x 100 feet (22.9 x 30.5 meters) that he concluded was “used for their religious 
rites and served them as a dance floor” (Benson and Bowers 1997:179). He also 
reported a large, subterranean floor measuring about 90 feet (27.4 meters) in 
diameter that could have been a dance or gaming area (Benson and Bowers 1997).

Ethnohistoric accounts describe the Chumash as having cemeteries that 
were associated with villages but were separate and usually just outside the 
village. Bowers did document the presence of separate cemeteries in the Santa
Ynez Valley. His map of Schniwax, for example, shows such a cemetery located just outside the village. Bowers’ basic mode of fieldwork included the discovery and excavation of cemeteries for the purpose of recovering artifacts, so it can be assumed that the sites he was drawn to excavate in the Santa Ynez Valley likely had cemeteries with artifacts.

2.3—Economic, Social, Political, and Ceremonial Organization

Trade alliances were maintained and served to buffer subsistence stress induced by climatic fluctuations by distributing resources between different ecological zones (Johnson 2000). The Chumash region can be divided into three basic ecological zones— island, coastal, and inland. People living in the island and coastal ecological zones developed efficient ways to exploit marine resources. Technological innovations such as the shell fishhook and the tomol, a seaworthy and highly advanced canoe made from sewn planks, led to expanded use of the marine environment and specifically pelagic resources at coastal sites. People living in the inland ecological zone, which includes the Santa Ynez Valley, were primarily dependent on terrestrial resources. Those include anadromous fish (steelhead), deer, and vegetal resources such as small seeds and acorns.

The variations in resource bases between these areas fostered trade alliances that may have included fiestas and ceremonial gatherings where resources were exchanged. Spanne (1975) argued that coastal populations moved to the interior when seasonal conditions made coastal fishing difficult. Tainter
(1975) argued that rather than moving populations, trade was facilitated by a seasonal cycle of fiestas, which would be held at inland or coastal sites based on resource availability.

Chumash trade fluoresced as a fully monetized exchange system based on beads made from *Olivella biplicata* shells. *Olivella* shells were only available in quantity on the islands, which created a monopoly over manufacture that established *Olivella* beads as a prestige good. Island villages traded *Olivella* beads with mainland villages for food, goods, and other valued items (Hildebrandt 2004; King 1976).

Social organization among the Chumash achieved a remarkable level of complexity (Blackburn 1975). The social organization of Chumash villages at the time of contact with Europeans was stratified into three general levels: the elites, craft specialists, and commoners (Glassow 1996). The elite class included the *wot*, who was the political leader of the village, his family, and the *paxa*, or ceremonial leader. The position of *wot* was hereditary and generally inherited by the eldest son or daughter, if no son was available (Blackburn 1975), but also required the general support of the village (Landberg 1965). In addition, dances and ceremonies were performed by a powerful cult organization whose members were referred to as ‘*antap*’. The ‘*antap*’ were members of the elite class (Blackburn 1975).

Certain manufactured goods were produced by occupational specialists who belonged to sodality-like organizations known as “brotherhoods” (Blackburn
Ethnographic sources report that the tomol, a sewn-plank canoe, was solely manufactured by the Brotherhood of the Tomol, whose members, in some cases, came from the wot’s own family (Harrington et al. 1978).

Postmarital residence was matrilineal for most Chumash, but was patrilocal for wots (Johnson 1988, 2007). Unlike most Chumash, wots also practiced polygamy, which may have facilitated alliances through marriage with other villages.

Chumash villages were largely autonomous, but did form multivillage alliances. Wots of certain large political centers held at least some degree of limited authority over surrounding villages (Johnson 1988, 2007). These alliances, however, seem to have been tenuous, based on either economic exchange or military alliance (Glassow 1996; Johnson 1988). Landberg argued that village populations would increase to the greatest extent supported by resources within the village resource catchment and would result in aggressively defended territories (1965). Intervillage conflict is well documented among the Chumash, mostly in the form of raids by small groups. Portolá, for example, noted many villages that were abandoned due to warfare (Landberg 1965).

Though an ethnohistorically documented cause of conflict was trespass, all hostilities between Chumash villages cannot be simply attributed to territorialism. Witchcraft was sometimes considered to be the cause of unexplained deaths, which would be retaliated with revenge killings. Economic motives cannot be
discounted, either. The status of *wot* could extended to sons and grandsons of *wots* if they showed particular prowess in warfare (Johnson 2007).
3.0—RESEARCH OBJECTIVES AND METHODOLOGICAL GOALS

The main objective of this research is to derive and test hypotheses concerning the environmental factors that influenced Chumash settlement patterns in the Santa Ynez Valley and to use the results of that hypothesis testing to create a map depicting the relative potential of regions within the valley to contain undiscovered archaeological sites. That objective has been further refined based on methodological goals and considerations.

3.1—Methodological Considerations and Goals

During the past 25 years of more intensive work in archaeological site location modeling, it has become apparent that several issues remain at the forefront and continue to affect the power and precision of models (Kvamme 2006). One goal of this research is to create a more powerful model by attempting to employ refined methods and data sources in an effort to address some of those issues. Kvamme has suggested that predictive modeling has entered a “second age” and offers a number of ideas and suggestions as to how new methods might be used to improve the process of modeling (2006:21). The problems and addressed in this study are listed below

1. Archaeological site types.

   Much of the data regarding the location of archaeological sites are based on surface reconnaissance of a few artifacts or
handfuls of food remains and are insufficient to provide accurate estimates of the function or age of sites or the intensity of settlement. That has often forced modelers to throw all sites into one large model even though ethnography and common sense indicate that sites of differing function and intensity of habitation will necessarily be located according to different criteria. In addition, the inclusion of sites of widely varying antiquity presents the possibility that even sites of similar function may have been inhabited at times when the environment was very different; therefore, even if the criteria of site location were similar, the resulting locations would necessarily be different.

Kvamme has stated that “defining meaningful site types and modeling each as a separate class is probably the greatest potential improvement in the quality of archaeological models” (2006:18). To that end, a subset of archaeological sites within the Santa Ynez Valley representing the locations of ethnohistorically documented year-round or nearly year-round long-term habitation sites will be used. The use of such a set of locations should improve the resulting model by assuring that all sites were occupied at roughly the same time and all conformed to a similar function. It should be noted that the goal is not to predict all site
locations within the study area, but to create an accurate model of only a single class of sites.

2. The Natural Environment

Relatively recent developments of GIS tools and datasets allow more accurate and finer-resolution representations of the natural environment. A particularly promising area is in the use of drainage runoff algorithms to derive flow patterns based on landform, allowing researchers to improve on often subjective and arbitrary blue-line features on topographic maps (Kvamme 2006).

Another area of potential lies in the use of more advanced landcover datasets now available. The advent of remote sensing has allowed improvements in the accuracy and resolution of vegetation and habitat maps, which can be used when appropriate as a proxy for past landcover.


Certain modeling methods, such as logistic regression, require that two classes of geographic locations be assumed—sites and non-sites. It has been argued that two-class approaches are appropriate for modeling archaeological phenomena because of the widely accepted concept that there are some locations that show evidence of past human activity (sites) and others that do not (non-sites). However, it is impractical, and perhaps in some respects
impossible, to perform a level of reconnaissance that might
guarantee that any particular spot is not an archaeological site nor
was ever the location of past human behavior. Therefore, the
further use of methods that rely only on a single “site-present”
class may be more appropriate to archaeological contexts of
predictive modeling.

4. Model testing.

Perhaps the most commonly reported measure of model
performance is the number of sites correctly predicted. That
statistic alone does little to describe the utility of a given model.
Appropriate measures of performance should not only consider the
proportion of sites identified correctly, but also the proportion of
the study area that the predicting class occupies. In addition, the
significance of correlations and confidence intervals associated
with those correlations can be used not only to test a model, but
also to guide the modeling process.

5. Full disclosure.

Much of the progress in archaeology is the result of
learning from the perspectives and methods employed by past
researchers. Archaeologists working with traditional research
methods such as field survey and excavation are particularly
thorough in reporting their methods, and we must be just as
thorough in describing our methods and the provenience of our
data when using GIS to construct models and perform analyses.
4.0—PRIOR MODELS OF SETTLEMENT PATTERNS IN THE SANTA YNEZ VALLEY

Valuable to this research are past efforts to construct models of subsistence and settlement patterns and systems within the Santa Ynez Valley and between the Santa Ynez Valley and coastal regions (Horne 1981; Tainter 1975:1; Woodman et al. 1991). Compared with the coastal regions of the Santa Barbara Channel and the Channel Islands, the Santa Ynez Valley has been the subject of considerably less research, which has been cited as a significant challenge that has hindered efforts at creating a detailed and well-articulated model of inland Chumash settlement patterns (Horne 1981; Spanne 1975). Past research into the prehistoric lifeways of the Ynezeno Chumash have been based on a combination of ethnohistoric, ethnographic, and archaeological data that are often sparse, requiring tremendous effort and creativity to tease out our current knowledge of inland Chumash prehistory.

Though no prior works have specifically intended to create a predictive model of archaeological sites within the Santa Ynez Valley, there are several that are valuable resources to this research. For example, John Johnson’s *Chumash Social Organization: An Ethnohistoric Perspective* provides data in the form of a comprehensive list of ethnohistorically documented villages (1988). Works by Stephen Horne (1981) and Joseph Tainter (1975) provide valuable insight into environmental influences on prehistoric village location and therefore serve to
inform the process of creating testable hypotheses by which data are selected for inclusion in the final predictive model.

A relatively recent and comprehensive study by Woodman et al. (1991) centered on Burton Mesa in the western Santa Ynez Valley proposes that sites in their study area may represent only a part of a larger settlement system that included coastal and inland sites. Based on the results of excavations of 23 different inland sites, researchers reported that virtually all project sites conform to a single generalized settlement and subsistence pattern—that of the dispersal of small residential family groups from villages in other areas. The researchers concluded that small family groups seasonally spanned out into the western plains area of the valley to gather resources both for sustenance and transport back to their home villages along the coast.

The evidence that Woodman et al. presented, combined with the relative infrequency of large village sites on the western alluvial plains when compared with the central and eastern Santa Ynez Valley, may indicate that two different settlement systems were in place in these two regions. That hypothesis may also be supported by ethnohistoric references that place the western boundary of the Ineseño linguistic territory as a roughly north to south line aligned with Nojoqui Summit. Land within the Santa Ynez Valley that lies to the west of that boundary is thought to have been part of the Purisimeño linguistic territory (Horne 1981; King 1975).
An earlier and less extensive study by Tainter (1975) conducted in the central and eastern Santa Ynez Valley suggests that residents there may have employed a more flexible settlement pattern whereby during dry years residents aggregated at large villages near confluences of rivers and perennial streams. During wetter years reliable water would have been more abundant, allowing them to disperse along tributaries at higher elevations. Based on the presence of shellfish at certain sites he proposed an additional hypothesis describing a fiesta system that linked coastal settlements with inland ones in the Santa Ynez Valley. During years when anadromous fish were plentiful, fiestas would be held in the valley and coastal residents would travel to them. During years where coastal resources were more plentiful fiestas would be held at coastal villages (Tainter 1975). According to Tainter’s model, it would be expected that the locations of year-round habitations should be biased towards proximity to confluences where water resources may have been more reliable year to year. That expectation lends itself well to testing using regional environmental data and to incorporation into a predictive model.

Steven Horne’s (1981) research area included part of the central portion of the Santa Ynez Valley, extending from the Santa Ynez River at its southern boundary beyond the Santa Ynez Watershed to the northern edge of the Cuyama River watershed. Using a combination of ethnohistoric and archaeological data, Horne sought to reconstruct the subsistence/settlement system within his research area. Reminiscent of Binford’s collector-forager spectrum (1980), Horne predicts
site types including Base Villages, Summer Villages, Food Processing Sites, Pine Nut and Acorn Camps, and Food and Water Caches. He also described an idealized seasonal round whereby resources in grassland and chaparral ecological zones would be exploited during late spring and early summer. During late summer and early fall, smaller task-oriented groups of adults would disperse to take advantage of mast harvests of acorns and piñon nuts that would be transported back to the residential base. During winter and early spring, groups would once again aggregate at residential bases, subsisting on stored foods.

Though the prediction of the locations of archaeological sites is beyond the scope of his work, Horne does provide hypotheses regarding the locations of base villages that are testable and lend themselves to the process of predictive modeling. He states that residential bases should be located with reference to “four major considerations: major trails, permanent water, access to food resources, and moderate winter weather” (1981:152).

He also proposes that during winter stored foods would have been supplemented with fresh foods that could be foraged within a short radius and that during summer residential bases would have been occupied by the elderly, infirm, and children to whom foraging areas close to the village would have been important. Therefore one expectation of Horne’s work would be that residential bases would have been located with a bias towards areas that possessed a variety of forage areas within a short distance.
Horne also points out that hunting camps would be expected on the boundaries of chaparral and grassland areas where hunters would deploy in search of game such as deer. That may mean that village sites near those areas would have been selected in order to reduce the amount of energy required to transport kills back to the residential base.
5.0—THE RESEARCH AREA

The Santa Ynez Valley lies in the central and western parts of Santa Barbara County in southern California. The valley is bounded on the south by the Santa Ynez Mountains, which are part of the Transverse Ranges, which are unusual in that they trend east-west rather than the north-south which is more typical of the coast ranges (Norris 2003). The valley is bounded on the north by the San Rafael Mountains, which are part of a transitional zone connecting the coast ranges to the Transitional Ranges. Together, they define the eastern and central portions of the valley, consisting of narrow bottoms through which flows the Santa Ynez River and its tributaries (Carpenter 1931). Near the town of Santa Ynez in the central region, the valley broadens rapidly to the west into a relatively flat and wide alluvial plain that narrows again slightly where the Santa Ynez River flows into the Pacific Ocean. Defined as such, the valley stretches approximately 104 kilometers from east to west, measures just over 22 kilometers north to south at its widest point, narrowing to less than one kilometer wide at its eastern extremity and approximately eight kilometers wide at its western terminus at the Pacific ocean. Elevations in the western alluvial plains range from sea level to about 600 feet and a few peaks within the bordering mountain ranges attain elevations over 4,600 feet (Norris 2003; Woodman et al. 1991).

By far the largest drainage in the region is the Santa Ynez River basin. Water flow is distinctly seasonal, following the Mediterranean climate pattern of cool, dry summers and wet, mild winters. Rainfall in the area averages
approximately six to eight inches per month during the winter and falls to less than one-tenth of an inch per month during the summer. Water flow in the Santa Ynez River varies considerably season-to-season and year-to-year. During the summer months water flow is usually less than 7 cubic feet per second (cfs). During most winters water flow varies from 100 to 1,000 cfs though periodic weather events can cause extremely heavy rainfall that in years prior to the building of flow regulation structures in the 1920s caused widespread flooding. For example, in 1907 during the largest flood on record, the peak flow of the Santa Ynez River was estimated to be 120,000 cfs (Woodman et al. 1991)—over 120 times what the maximum flow would be in a typical winter.

The Santa Ynez River Valley is a mosaic of ecotones with a wide range of habitats. Among the coastal terraces of the western alluvial plains, Coastal Dune plant communities predominate that could have supported many potentially exploitable species. Widespread distribution of weedy forbs, herbs, and shrubs combined with interspersed open areas provides an ideal habitat for jackrabbits (*Lepus californicus*), which are abundant in that part of the valley (Woodman et al. 1991).

Dry slopes and terraces north of the Santa Ynez River support the Coastal Sage Scrub plant community. Plant species in this community that might have provided resources to prehistoric people include soap plant (*Chenopodium californicum*), chia (*Salvia columbariae*), and red maids (*Calandrinia ciliata*) from which small seeds could have been harvested (Timbrook 2007), as well as
blue dicks (*Dichelostemma capitatum*) that could have provided edible corms (Anderson 2005). Terrestrial mammals that could have been exploited include mule deer (*Odocoileus hemionus californicus*), desert cottontail (*Sylvilagus audubonii*), brush rabbit (*Sylvilagus bachmani*), and jackrabbit (Woodman et al. 1991). In addition, portions of the valley lie beneath the Pacific Flyway and marshland habitats attract both resident and migratory birds such as ducks, geese, and rails (Woodman et al. 1991).

Areas that receive more water such as those at somewhat higher elevations support Chaparral plant communities. Potentially useful plant species that populate the chaparral include live oak, manzanita (*Arctostaphylos purissima*), chia, blue dicks, soap plant, islay (*Prunus ilicifolia*), and toyon (*Heteromeles arbutifolia*). Mule deer, fox (*Urocyon cinereogenteus*), birds, and jackrabbits are among the faunal species that could have been important to prehistoric foragers (Woodman et al. 1991). Other plant communities that occur within the valley include Bishop Pine, Grasslands that could have provided small seeds and corms, and Wetlands, Marshlands, and Riparian Woodlands with numerous resources such as willow saplings for building houses and basketry (Timbrook 2007), *Chenopodium macrospremum* and *Amaranthus californicus*, both sources of small seeds. Oak Woodlands were very important to Later Period Chumash as a source of acorns and pine nuts were harvested from Grey Pine (*Pinus Sabiniana*), that is a member of the Blue Oak—Digger Pine habitat (California. Dept. of Forestry and Fire Protection et al. 1988).
The watershed of the Santa Ynez River forms the geographical bounds of the research area of this study (Fig 5.1). The watershed area is 233,387 hectares and is defined by the spatial database CalWater 2.2.1 that was authored and is maintained by the California Interagency Mapping Committee (2004).

An important aspect of GIS-based modeling studies is the use of available datasets as appropriate in order to avoid the often labor-intensive and therefore costly process of digitizing and data capture. Another benefit of using public spatial datasets is that mapping projects conducted by large government agencies often have budgets far beyond even the most extravagant archaeological research budget, and therefore error detection and correction is often accomplished to a level that cannot be replicated by the average individual researcher.

The CalWater 2.2.1 database is such a dataset. As the name of the authoring committee would imply, the CalWater 2.2.1 dataset is the product of several revisions as a cooperative effort between several government agencies. The original dataset was generated in the 1970s and 1980s based on 1:500,000 scale Hydrologic Basin Planning Maps from the State Water Resources Control Board (SWRCB) in Sacramento, California. In the 1990s the California Department of Forestry and Fire Protection contracted to have the maps improved based on 1:24,000 scale SWRCB Hydrologic Basin Planning Maps. Currently, the USDI Bureau of Reclamation, USDA Natural Resources Conservation Service, USDA Forest Service, U.S. Geological Survey and other federal agencies continue to work with staffs of National Forests, BLM, Counties, Resource
Conservation Districts (RCD), Regional Water Boards, the State's Interagency Watershed Mapping Committee (IWMC), and other interested parties to correct errors and to revise California watershed boundaries for conformance to national standards (California Interagency Watershed Mapping Committee 2004).

The CalWater 2.2.1 database was downloaded from the California Spatial Information Library (http://gis.ca.gov/casil/hydrologic/watersheds/calwater/). The features that comprise the Santa Ynez River watershed were extracted, and the resulting feature class was reprojected to the project coordinate system using the transformation recommended by Environmental Science Research Institute, Inc. (Environmental Systems Research Institute 2006a). The reprojected data was visually checked against features such as peaks and ridges to assure its spatial alignment with other layers in the project. Please note that the term “feature class” can is defined here as a collection of geographic features with the same geometry type (such as point, line, or polygon), the same attributes, and the same spatial reference.

The choice of a watershed boundary is appropriate for this research for a number of reasons. It has been said that watersheds were “the most common boundary” of hunter-gatherer groups (Kroeber 1962:49). Surrounding mountain ranges would have presented a natural barrier that could have made interactions between settlements within the watershed more geographically attractive than interactions with villages beyond the valley. In addition, the rugged landscape of the surrounding mountains would have confined travel to major watersheds
(Horne 1981), possibly creating a trail network within the valley that may have roughly followed the Santa Ynez River and its tributaries.
Figure 5.1 – Study area.
The Chumash were perhaps some of the most socially complex prehistoric hunter-gatherer groups known. Some aspects of Chumash culture that most impressed early observers include high population densities; large, socially integrated settlements, an emphasis on trade, and highly developed arts and technology (Johnson 1988). Settlement within the Chumash regions has great time depth. Archaeological evidence along the Santa Barbara Channel dates back over 10,000 years and prehistoric land use in the Santa Ynez Valley extends back at least 8,000 years (Glassow 1996). During that time there is evidence of increasing social complexity and an articulated system of trade, subsistence procurement, and settlement systems that spanned both coastal and inland areas (Glassow 1996; Woodman et al. 1991).

A great deal of what we know about Chumash social organization and lifeways comes from ethnographic and ethnohistoric resources. The ethnographic information comes primarily from sources that can be divided into four categories: the Spanish Exploration period (1542-1770); the Mission period (1772-1834); a short period of ethnographic inquiry (1860-1900); and modern anthropological research (Blackburn 1975). Though researchers lament the paucity of written ethnohistoric records of the details of Chumash life, Spanish missionaries did leave a legacy of important information in the form of birth, death, and marriage records of Chumash who were recruited into the mission system. That data has allowed researchers to reconstruct a great deal of Chumash
social organization, population distribution, and marriage patterns (Johnson 1988).

One of the products of such research is the identification of archaeological sites that represent specific villages mentioned in the registers. John Johnson used the mission registers as well as other sources to reconstruct a list of 22 villages within the Santa Ynez Valley. One of the villages, Huelecmen, contributed only five marriages to the mission records and was eliminated from Johnson’s geographic analysis (Johnson 1988). The 21 remaining villages are those used for this study (Table 6.1). The availability of such a list of villages assures that the villages included in this study were occupied contemporaneously. Because all were principal settlements, they can be expected to have been located according to similar criteria, if those criteria were relative to site function. The feature class representing the locations of archaeological sites identified with those villages was supplied by the Central Coastal Information Center at the University of California at Santa Barbara.

<table>
<thead>
<tr>
<th>Trinomial</th>
<th>Village Name</th>
<th>Trinomial</th>
<th>Village Name</th>
<th>Trinomial</th>
<th>Village Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBA-123</td>
<td>Snojoso</td>
<td>SBA-1187</td>
<td>Sajuchu</td>
<td>SBA-485</td>
<td>Eljman</td>
</tr>
<tr>
<td>SBA-842</td>
<td>Miasap</td>
<td>SBA-1510</td>
<td>Ytiax</td>
<td>SBA-871</td>
<td>Huililic</td>
</tr>
<tr>
<td>SBA-865</td>
<td>Huisapa</td>
<td>SBA-235</td>
<td>Jojonata</td>
<td>SBA-1283</td>
<td>Sajcaya</td>
</tr>
<tr>
<td>SBA-809</td>
<td>Aquitsumu</td>
<td>SBA-1190</td>
<td>Jalama</td>
<td>SBA-219</td>
<td>Lompoc</td>
</tr>
<tr>
<td>SBA-1645</td>
<td>Stucu</td>
<td>SBA-572</td>
<td>Sipuc</td>
<td>SBA-167</td>
<td>Sotonocmu</td>
</tr>
<tr>
<td>SBA-516</td>
<td>Calahuasa</td>
<td>SBA-823</td>
<td>Sinjuaj</td>
<td>SBA-1309</td>
<td>Snajalayegua</td>
</tr>
<tr>
<td>SBA-1183</td>
<td>Najue</td>
<td>SBA-477</td>
<td>Tequeps</td>
<td>SBA-1800</td>
<td>Siguaya</td>
</tr>
</tbody>
</table>

Table 6.1—Village sites included in this study
Availability of fresh water was possibly the single most influential environmental factor affecting the locations of the village sites included in this study. The influence of water sources on settlement patterns is not unexpected. Fresh water is one of the most fundamental human biological needs. According to comprehensive studies conducted in the 1940s, an adult working in the sun requires between 7.5 and 11 liters of fresh water per day (The Science News-Letter 1943:193). Doing the math, that adds up to at least 375 kilograms of water per day that must be transported to maintain even a moderately sized village of 50 residents, so not only access to fresh water sources but also close proximity to those sources would have been critical.

Previous researchers working in the study area have noted the importance of reliable water sources as critical to long-term habitation in the study area (Horne 1981; Tainter 1971). Tainter cited proximity to water, specifically confluences and intermittent tributaries, as a basis of his settlement model (1971). He speculated that residents of the central Santa Ynez Valley congregated in the summer into large permanent villages during years of relatively low rainfall. In years of greater rainfall residents would disperse along less reliable streams and into higher elevations.

Locations of water sources may indicate that different settlement strategies may have been employed in the central as opposed to the eastern portions of the
Santa Ynez valley. It can be observed that the most reliable perennial water sources are only available in the western and central portions of the research area, yet there are two villages in the eastern part as well which are relatively distant from documented perennial water sources. It is possible to theorize various ways that prehistoric inhabitants of the Santa Ynez Valley may have strategically approached the problem of obtaining reliable supplies of water. Horne has suggested that perennial water was one of the resources most influential to settlement location and so settling near perennial streams may have been an effective strategy of adaptation to the dry and unpredictable environment of the valley. Tainter (1975) suggests that water sources may have been relatively more reliable at confluences and therefore locating long-term, year-round settlements near confluences might have been a successful adaptive strategy.

Each of those strategies would necessarily manifest as different evidence layers in a probabilistic predictive model. Each of those strategies can be expressed as testable hypotheses that can be used to inform the process of selecting and deriving datasets and to determine which of the resulting evidence layers are most appropriate for inclusion in the resulting predictive model. In order to determine the most appropriate way to include proximity to water in this model, the following steps were taken:

1) Multiple working hypotheses were constructed based on previous research and theory.
2) A wide array of sources were employed to construct a spatial database that, as accurately as possible, reflects hydrologic flow patterns in the Santa Ynez Valley during the most recent times when the study sites were occupied.

3) Attributes were defined for the hydrology dataset so as to differentiate between water sources of varying reliability (e.g. perennial, intermittent, etc.).

4) A spatial dataset was generated of confluences of perennial and intermittent streams within the study area.

5) Raster datasets were generated wherein each cell contains the distance in meters between that cell and the features of the hydrologic datasets within the study area. A raster is defined as a dataset where geographic space is divided into an array of equally sized cells arranged in rows and columns with each cell containing a value.

6) The distance rasters generated in step five were used to create tables of distances between villages and hydrologic features.

7) Statistical analyses were performed to test hypotheses.
8) Based on those results, a new raster dataset was generated so as to achieve the greatest accuracy, precision, and therefore the model of the highest predictive power.

7.1—Background

The climate of the Santa Ynez Valley is classified as Mediterranean, characterized by cool, dry summers and mild, wet winters (Woodman et al. 1991). Two of the most notable characteristics of the valley are its semi-arid environment and mountainous terrain (Carpenter 1931). Rainfall in the region is markedly seasonal, averaging only 17.25 inches per year with about 90% of precipitation falling during the months of November through March (Carpenter 1931). Over 96% of the soils in the study area are rated as well-drained to excessively drained (Soils Survey Staff 2006). As a result, there are relatively few sources of fresh water in the study area that persist throughout the year, which presented significant challenge and risk to prehistoric settlements within the study area.

7.2—Previous Research

Water has been listed among the six components that define a landscape from a human perspective (Stafford and Hajic 1992). Researchers have stated that water is a resource “so basic and so vital that the distance to obtain [it] must be minimized” (Roper 1979:121). Consistent access to fresh water was possibly one of the most immediate and fundamental challenges to the maintenance of permanent, year-round habitations within the Santa Ynez Valley. Of the several
pieces of past research dealing with settlement, subsistence, and social, trade, and
political relations within the study area (Horne 1981; Johnson 1988; Spanne 1975;
Tainter 1971, 1975), two in particular are environmentally based and feature
access to water as a significant influence on settlement patterns with regard to
year-round long-term habitations. Horne (1981) published a study that primarily
sought to reconstruct a seasonal round in the central Santa Ynez Valley, with
access to perennial water sources as one of the major criteria determining the
suitability of locations for year-round, long-term habitation. Tainter (1971)
proposed a model wherein during the dry season in dry years, settlements would
be expected to be tethered to places where water might be more reliably found,
such as springs, the confluences of drainages, or along perennial streams.

Researchers in other geographic regions have designated fresh water as a
critical factor influencing settlement patterns. In his analysis of catchments
centered on a geographic region to the southeast of the Santa Ynez Valley,
Tartaglia cited availability of fresh water as the primary factor influencing the
location of late period villages (Tartaglia 1980). Groups in arid or semi-arid
regions whose movements and settlement patterns are heavily influenced by the
availability of water have been termed “tethered foragers” (Kelly 1995:126). The
effect of being tethered to water is magnified by the positive correlation between
plant food resources and available water. In very arid regions the avoidance of
risk associated with lack of fresh water can drastically affect hunter-gatherer
mobility patterns. Foragers in Australia, for example, have been observed to
accept extreme caloric shortages and extend their foraging radius as much as 15 kilometers rather than relocate away from a reliable fresh water source (Kelly 1995). Therefore it is not surprising that proximity to water is possibly the most common dataset included in predictive probabilistic models (e.g. Brewster et al. 2003; Duncan and Beckman 2000; Madry 1986; Rosenthal et al. 2003; Warren and Asch 2000; Wescott and Kuiper 2000).

7.3—Hypotheses and Test Implications

Whereas proximity to water may seem to be a simple and rather perfunctory characteristic to model with respect to human habitation, there are a variety of strategies that native inhabitants of the Santa Ynez Valley may have employed in an effort to achieve the greatest possible reduction in risk associated with access to water while reducing the energy required to transport water to the habitation site. In order to assure that the final dataset included in the model has the greatest explanatory and predictive power, multiple working hypotheses must be considered. A brief description of those hypotheses and test implications is as follows:

1) *Proximity to perennial water sources was a major criterion of village site location.*

Because the habitation sites included in this study are thought to have been occupied year-round or nearly year round, access to perennial water sources would have been critical. This hypothesis is also derived
from aspects of both Horne’s (1981) and Tainter’s (1971) models.

The test implication of this hypothesis are:

If the mean distance between village sites and their closest perennial water sources is consistently shorter than that of a random distribution of points, then the null hypothesis that village site location was random with respect to perennial water would be rejected and the hypothesis would not be rejected.

Another possible explanation could be that the presence of perennial streams is masking some other coincident phenomena that had a more significant influence on village site location. Masking is a theoretical possibility with any dataset that may be included in a predictive model (Kvamme 2006), and therefore the possibility of masking within a predictive model can never be completely rejected. The possibility of masking is greatest in models that rely heavily on purely inductive methods whereby numerous datasets are included in an effort to discover correlations. By basing our hypotheses on anthropological theory and human behavior, the likelihood of masking is reduced to an acceptable level.

Masking is also an issue when determining whether confluences or perennial streams had a greater influence on village settlement distribution. Confluences are necessarily coincident with the locations of streams, so if a comparison shows, for example, that the mean distance
between village sites and perennial streams was significantly smaller than
the mean distance between a random set of points and perennial streams,
then the same comparison of means repeated for confluences would
invariably show that village sites were also significantly closer to
confluences. To address this issue, a two-step process has been used.
First, the mean distances between village sites and perennial streams will
be compared with a random distribution of points and then the distances
between village sites and confluences will be compared with the distances
between village sites and streams. This will be discussed further under
Hypothesis 3 below.

2) **Villages were located near intermittent streams if perennial water sources
were not available in the region.**

This hypothesis is one logical outcome if Hypothesis #1 above is
rejected. The test implication of this hypothesis is:

If the distances between village sites and perennial water sources is
less than the mean distance to intermittent water sources in regions where
perennial water sources exist and the mean distance between intermittent
water sources and village sites is less than that of perennial water sources
in regions where perennial water sources do not exist, then the hypothesis
would be supported. If village sites located away from perennial water
sources were also distant from intermittent water sources then the
hypothesis would be rejected.
3) Proximity to confluences of streams was the primary criterion of village site location with respect to proximity to water.

Test implications of this hypothesis are:

a. If the mean distance between village sites and streams were significantly greater than the mean distance between village sites and confluences, then the hypothesis would be supported.

b. If the mean distance between village sites and confluences is significantly smaller than half the mean distance between confluences then the hypothesis would be supported.

Note: this test implication addresses the possibility that villages may erroneously appear to be located close to confluences in areas where confluences are very spaced very close together. Due to the rugged terrain that exists throughout much of the study area, regions exist where confluences are spaced very closely together so this question must be addressed.

In order to illustrate this issue, consider a hypothetical situation where one priority of settlement was to be within 300 meters of a perennial stream. Imagine that a village is located according to that priority (within 300 meters of a perennial stream) between two confluences that are 500 meters apart. In that case, even if the true draw to settlement was proximity to perennial water, the village could have unavoidably been located even closer
to the nearest confluence. In such a case, the furthest that the village could be located from the nearest confluence would be approximately half the distance between the confluences, or about 250 meters.

4) Proximity to springs was a significant criterion of village site location with respect to proximity to water

Springs are another water source suggested by Tainter as being more reliable, and consequently having influenced settlement patterns during dry seasons of dry years (1971).

The test implications of this hypothesis are:

a. If the mean distance between the village sites and the nearest spring is significantly smaller than the mean distance between a random distribution of points and the nearest spring then the null hypothesis that the distribution of village sites was random with respect to the location of springs would be rejected and therefore the hypothesis would not be rejected.

7.4—Methods

Generation of a Hydrography Feature Class

Data collection is one of the most time-consuming and important of tasks in the creation of a Geographic Information System (Longley 2005). The creation of the hydrology feature class was no exception. Because the distances between villages and water sources are generally less than 300 meters, the flow paths must
be represented very accurately. In addition, the hydrology feature class must differentiate between intermittent and perennial streams. The most comprehensive hydrology dataset publicly available is the National Hydrology Dataset (NHD), which is publicly available through the United States Geological Survey (USGS) (U.S. Geological Survey 1995). One concern with hydrologic datasets is accuracy of flow path (Keith Clarke, personal communication, 2006). Kvamme also voiced this concern when he wrote that, “blue-line features on topographic maps are frequently arbitrary and unreliable indicators of water” (Kvamme 2006:7).

The published metadata for the NHD verifies that the cause for concern regarding accuracy is genuine. The reliability of attributes, such as the classification of streams as perennial and intermittent, is reported to be 98.5% (U.S. Geological Survey 1999d), so that aspect of accuracy is adequate. Horizontal accuracy (e.g., the east-to-west and north-to-south accuracy of a feature’s location on the surface of the earth) is somewhat less than optimal for determining such small distances. The NHD was originally transposed from earlier Digital Line Graphs, which were derived from 1:100,000 topographic maps. The accuracy standard of the DLGs is published as 90% of the points tested being within 0.02 inches at map scale. In addition, the accuracy of digitizing the finished dataset is stated as plus or minus 0.003 inches. Features are also sometimes purposely shifted if feature density makes maps hard to read (U.S. Geological Survey 1999d). At 1:100,000 scale, 0.023 inches error equals 58.4
meters ground measurement—over 51% of the mean distance between villages and perennial or intermittent water sources (113.9 meters).

A more accurate alternative is to generate flow paths of streams from drainages mapped using digital elevation models. The digital elevation models used in this study are datasets in a format commonly referred to as “raster.” Raster datasets can be visualized as grids, with each cell of the grid holding a number. In a digital elevation model, each cell of the raster contains the average elevation within that cell.

The highest quality digital elevation model that is currently available for public download is the National Elevation Dataset (NED) (U.S. Geological Survey EROS Data Center 1999). National Elevation Dataset data covering the study area is available in acceptably high resolution. The National Elevation Dataset dataset downloaded for this study was in a geographic coordinate system with a NAD83 horizontal datum, each cell representing 1/3 arc second or approximately 10 meters square of ground surface.

National Elevation Dataset horizontal accuracy is reported as 7 meters Root Mean Squared Error (RMSE) (U.S. Geological Survey 1999b). RMSE is defined as the square root of the average squared error and can be thought of as the average error in each observation, whether that error is positive or negative (Longley 2005:140-141). Other characteristics of National Elevation Dataset data imply that it is more suitable for the purposes of deriving drainages and hydrographic flow lines, though those characteristics are not necessarily
quantified by the USGS in terms of improved accuracy. For example, the elevation from the National Elevation Dataset is a bare-ground reading as opposed to “first return” or canopy-based readings of more current radar-derived elevation data such as the more current Shuttle Radar Topography Mission (SRTM) (U.S. Geological Survey 1999b). In addition, one of the effects of the National Elevation Dataset processing steps is a “much-improved base of elevation data for calculating slope and hydrologic derivatives” (U.S. Geological Survey 1999b:1).

A hydrologic dataset was derived using the hydrologic toolset that is supplied with ESRI ArcGIS 9.1 (Environmental Systems Research Institute, Inc. 2006b). Best practices as defined by ESRI entails a series of steps to derive runoff characteristics from digital elevation model data. Some of those steps are required and others are optional, depending on the nature of the specific input data. A brief description of the steps taken in deriving the potential flow paths within the study area is as follows:

a) *Creation of a depressionless Digital Elevation Model.*

The direction of flow from any given cell is determined to be the direction of maximum slope away from that cell. Digital elevation model datasets commonly contain minute depressions, termed “sinks.” Whereas sinks can reflect actual depressions in the landscape, they are often erroneous errors in the data. As flow direction can proceed into such a sink but cannot proceed out,
sinks can interrupt flow paths and impart errors into the final hydrographic dataset. The National Hydrography Dataset downloaded for this study was analyzed and a small number of sinks were identified. The depth of each sink was determined and each sink was filled using the ArcGIS Fill tool (Environmental Systems Research Institute, Inc. 2006c).

b) *Creation of a raster dataset representing the direction of potential water flow at each cell within the study area.*

The Flow Direction tool supplied with ArcGIS 9.1 was used to create a raster of flow direction for each cell in the depressionless Digital Elevation Model.

c) *Creation of a raster dataset representing the flow accumulation for each cell in the study area.*

The Flow Accumulation tool supplied with ArcGIS creates a raster that represents the accumulated weight of all cells flowing into each downslope cell.

d) *A raster was created by selecting all cells with a flow accumulation greater than 700 cells.*

Using the raster calculator supplied with ArcGIS, rasters were created representing cells with flow accumulations of 100 to 1,500. Those rasters were visually compared against the National
Hydrography Dataset and the raster of cells with flow accumulations ≥ 700 was selected as having the level of detail closest to the NHD network of perennial and intermittent water sources.

e) *The Stream to Feature tool supplied with ArcGIS was used to create a linear feature class from the raster created in step 4 above.*

The accuracy of the derived flow paths was assessed by visually comparing the resulting feature class to stream paths visible in USGS Digital Ortho Quarter Quadrangles (DOQQ) aerial photographs (U.S. Geological Survey (USGS) EROS Data Center 1997) and National Agricultural Imagery Program (NAIP) one-meter resolution aerial photography. (USDA - Farm Service Agency - Aerial Photography Field Office 2005). Whereas the apparent accuracy of the feature class was quite impressive, several problem areas were identified. Erroneous flow paths were discovered in areas where topographic relief was particularly low and were manually corrected to match aerial photographs. In addition, there are three reservoirs in the study area; Lake Cachuma, which was created when the Bradbury dam was built in 1951; the Gibraltar Reservoir which was created when the Gibraltar Dam was built in 1920; and Jameson Lake which was created when the Juncal Dam was built between the years 1921 and 1930 (Montecito Water District 1998). All three reservoirs were depicted as flat areas
in the digital elevation model and therefore created erroneous flow paths through those areas.

Lake Cachuma, being quite large, was of particular concern as it disrupted part of the Santa Ynez River where two of the village sites in this study are located. In an effort to reconstruct the original flow of the river prior to the construction of Bradbury Dam, historic aerial photographs were obtained from the Map and Imagery Laboratory at the University of California Santa Barbara. That imagery, which was acquired by Fairchild Aerial Surveys in 1938, has sufficient resolution that river channels are discernable (Fairchild Aerial Surveys 1938). The photographs covering the Lake Cachuma area were georeferenced using features such as buildings and trees that appear in both the Fairchild aerial photographs and the NAIP imagery. The flow path of the river was edited to match the path apparent in the aerial photographs.

The other two identified reservoirs, Gibraltar Reservoir and Jameson Lake, were created before the Fairchild imagery was acquired so the flow path in those areas were estimated based on the surrounding topography.

The above process resulted in a feature class that accurately represents flow paths as determined by drainages, but in order to generate the data needed to test the previously listed hypotheses, perennial and intermittent streams must be differentiated. Where permanent streams begin and whether those streams are perennial, intermittent, or ephemeral is affected not only by the contributing catchment but also by climate, slope, and soil characteristics (Environmental
Systems Research Institute, Inc. 2006c). Therefore, the required attributes cannot be derived by simply setting flow accumulation thresholds.

Even though the National Hydrography Dataset’s spatial accuracy is less than required by this study, the published accuracy of its attributes is 98.5%, clearly adequate for our purposes. The flow type attribute (perennial and intermittent) was transferred from the National Hydrography Dataset to the study hydrography feature class first through selecting features using buffers where possible, then by visually scanning the entire study area in small sections, matching features in the elevation-derived hydrologic data to the corresponding features in the National Hydrology Dataset, and transferring the needed data to indicate whether a particular stream is classified as perennial or intermittent (Figure 7.1).

Generation of the Confluence Feature Class

A point feature class was created representing the locations of confluences within the study area. The points were digitized manually during the process of applying flow type attributes to the hydrography feature class.

Generation of a Springs Feature Class

The springs within the study area were extracted from the National Hydrography Dataset geodatabase, as that was the only source of this data that was discovered while data was being collected for this study.
Generation of Distance Surfaces

Four raster datasets were generated using the Euclidian Distance tool supplied with ArcGIS 9.1 as part of the Spatial Analyst toolset (Environmental Systems Research Institute Inc. 2006b). The cell size of each raster was approximately 10 meters square and each cell contained the average Euclidian distance and the nearest intermittent stream, perennial stream, spring, or confluence (Figures 7.2—7.4).

Create attributes of distances to perennial streams, intermittent streams, springs, and confluences within the village site feature class.

The attributes were copied from each raster using the Intersect Point tool, of Hawth’s Analysis Tools for ArcGIS (Beyer 2004) (Table 7.1).
Figure 7.1 – Perennial and intermittent streams within the study area.
Creation of feature classes of random points.

The purpose of this step is to derive a dataset of the distances to the various water sources that would be derived from a random distribution so that null hypotheses can be efficiently tested. Hawth’s Analysis Tools for GIS was used to create a point feature class of 2,100 points randomly distributed throughout the study area and the various distance attributes were collected using the Intersect Point tool in the same toolset (Beyer 2004).

Statistical analysis

Comparisons of means were performed by independent samples t-tests using the statistical analysis package SPSS 15.0 for Windows (SPSS Inc. 2006).
<table>
<thead>
<tr>
<th>Trinomial</th>
<th>Village Name</th>
<th>Perennial Stream</th>
<th>Intermittent Stream</th>
<th>Spring</th>
<th>Confluence</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBA-123</td>
<td>Snojoso</td>
<td>85.2</td>
<td>723.9</td>
<td>3568.0</td>
<td>592.2</td>
</tr>
<tr>
<td>SBA-842</td>
<td>Miasap</td>
<td>52.4</td>
<td>456.2</td>
<td>2140.9</td>
<td>463.6</td>
</tr>
<tr>
<td>SBA-865</td>
<td>Huisapa</td>
<td>43.0</td>
<td>141.2</td>
<td>2526.9</td>
<td>165.6</td>
</tr>
<tr>
<td>SBA-809</td>
<td>Aquisumu</td>
<td>54.4</td>
<td>292.3</td>
<td>1061.5</td>
<td>170.4</td>
</tr>
<tr>
<td>SBA-1645</td>
<td>Stucu</td>
<td>48.7</td>
<td>169.6</td>
<td>1433.0</td>
<td>431.8</td>
</tr>
<tr>
<td>SBA-516</td>
<td>Calahuasa</td>
<td>115.8</td>
<td>129.1</td>
<td>6325.3</td>
<td>119.6</td>
</tr>
<tr>
<td>SBA-1183</td>
<td>Najue</td>
<td>78.0</td>
<td>1271.5</td>
<td>1997.3</td>
<td>1633.7</td>
</tr>
<tr>
<td>SBA-1187</td>
<td>Sajuchu</td>
<td>761.5</td>
<td>250.7</td>
<td>4270.2</td>
<td>792.7</td>
</tr>
<tr>
<td>SBA-1510</td>
<td>Ytiax</td>
<td>312.9</td>
<td>0.0</td>
<td>581.4</td>
<td>163.3</td>
</tr>
<tr>
<td>SBA-235</td>
<td>Jojonata</td>
<td>146.3</td>
<td>343.0</td>
<td>3738.2</td>
<td>458.2</td>
</tr>
<tr>
<td>SBA-1190</td>
<td>Jalama</td>
<td>129.1</td>
<td>54.4</td>
<td>2768.9</td>
<td>205.3</td>
</tr>
<tr>
<td>SBA-572</td>
<td>Sipuc</td>
<td>215.2</td>
<td>1308.7</td>
<td>4889.4</td>
<td>593.0</td>
</tr>
<tr>
<td>SBA-823</td>
<td>Snijuaj</td>
<td>243.0</td>
<td>200.2</td>
<td>5805.0</td>
<td>281.5</td>
</tr>
<tr>
<td>SBA-477</td>
<td>Tequeps</td>
<td>441.0</td>
<td>35.5</td>
<td>6879.6</td>
<td>727.7</td>
</tr>
<tr>
<td>SBA-485</td>
<td>Ejman</td>
<td>51.7</td>
<td>747.5</td>
<td>3054.4</td>
<td>697.5</td>
</tr>
<tr>
<td>SBA-871</td>
<td>Huililic</td>
<td>77.5</td>
<td>85.2</td>
<td>701.5</td>
<td>133.6</td>
</tr>
<tr>
<td>SBA-1283</td>
<td>Sajcaya</td>
<td>94.7</td>
<td>134.5</td>
<td>5187.1</td>
<td>134.5</td>
</tr>
<tr>
<td>SBA-219</td>
<td>Lompoc</td>
<td>1859.0</td>
<td>289.4</td>
<td>5298.0</td>
<td>2327.9</td>
</tr>
<tr>
<td>SBA-167</td>
<td>Sotonocmu</td>
<td>19.2</td>
<td>433.3</td>
<td>4324.2</td>
<td>439.8</td>
</tr>
<tr>
<td>SBA-1309</td>
<td>Snajalayegua</td>
<td>13122.4</td>
<td>85.2</td>
<td>3767.8</td>
<td>267.4</td>
</tr>
<tr>
<td>SBA-1800</td>
<td>Siguaya</td>
<td>11817.7</td>
<td>60.3</td>
<td>4016.6</td>
<td>594.4</td>
</tr>
</tbody>
</table>

Table 7.1—Distances between water sources and the village sites included in this study.
Figure 7.2 – Distance to perennial streams in the study area.
Figure 7.3 – Distance to intermittent streams in the study area.
Figure 7.4 – Distance to springs within the study area.
Figure 7.5 – Distance to confluences in the study area.
7.5—Results

The results of hypothesis testing are as follows:

1) *Proximity to perennial water sources was a major criterion of village site location.*

Results:

The 1960.8 m difference between the mean distance to perennial water of 2100 points distributed randomly throughout the study area (3378.4 m) and that of the village points (1417.6 m) is very significant ($t=5.518$, $p>0$). There were three extreme outliers in the data; CA-SBA-1309 *Shnaxalyiwi* at 13,122.4 m, CA-SBA-1800 *Siwaya* at 11,817.7 m, and CA-SBA-219 (Lompoc) at 1859 m. The first two, CA-SBA-1800 and CA-SBA-1309 are located in the eastern valley where no watercourses are identified as perennial in the NHD dataset. The lack of perennial streams in the eastern valley may be a deficiency of the National Hydrology Dataset rather than an accurate reflection of the reliability of water sources. Archaeologists who frequent that geographic region have observed that watercourses in the vicinity of these two sites have surface water much of the year (John Johnson, personal communication). In addition, the site of *Siwaya* is adjacent to the Ogalvie Ranch Homestead, which would likely have been located so as to have easy access to reliable water (John Johnson, personal communication).

The third outlier, CA-SBA-219 (*Lompo’*), is located in an area of intensive
irrigation and agriculture where the original stream course has likely been disrupted. The site record states that it “occupied a triangular piece of land at the junction of a creek and the Santa Ynez River or two branches of the R.” (Ruth 1950:1). Based on these test results, the null hypothesis that village site location was random with respect to perennial streams is rejected and therefore the hypothesis cannot be rejected.

2) A major criterion in village site location was proximity to the most reliable water available within the settlement region.

Results:

Based on analysis of village sites that are relatively far from perennial streams, the hypothesis cannot be rejected, but the number of unambiguous cases (n=2) is too small to facilitate meaningful statistical analysis.

Two village sites, CA-SBA-1309 (Snaxalyiwí) and CA-SBA-1800 (Siwaya) are in regions where no perennial water sources are noted in the National Hydrology dataset (Figure 6.2). They are 11,800 m and 13,100 m from the nearest marked perennial stream. Both are close to intermittent streams (50 and 85 meters). As noted before, this may be due to inaccuracies in the National Hydrology Dataset.

Two other village sites that are relatively far from perennial water sources may have ambiguous locations with respect to perennial water. CA-SBA-219 (Lompoc) as stated above is in an area where stream path has likely been
disrupted. CA-SBA-477 (Tequepsh) is 441 meters from the nearest perennial stream and only 35 meters to the nearest intermittent stream, but its location and the location of the watercourse were disrupted by the construction of the Bradbury Dam and the resulting creation of Lake Cachuma.

3) Proximity to confluences was a significant criterion when selecting village sites.

Results:

Village sites are significantly further from confluences than streams so the hypothesis is not supported. After applying a log transformation to the data to equalize variances, the 428.7 m difference between the mean distance to streams (113.9 m) and the mean distance to confluences (542.6 m) is highly significant (t = –5.784, .005 > p > 0).

The average distance to confluences is not significantly less than the mean distance between confluences; therefore, the null hypothesis cannot be rejected and the hypothesis cannot be supported. The 154.3 m difference between the mean distance between confluences (697 m) and the mean distance from village sites to nearest confluence (542.6 m) is not significant (t = 1.52, p = 0.857).

4) Proximity to springs was a significant criterion of village site location with respect to proximity to water

Results:

The 197.2 m difference between the mean distance between village sites
and the nearest springs (3539.8 m) and that of a random distribution of 2100 points (3737.0 m) is not significant ($t=-.492$, $p=.682$) therefore the null hypothesis that village location was random with respect to springs cannot be rejected and the hypothesis is not supported.

7.6—Conclusions

Based on the results of these tests, it can be said that proximity to perennial streams was a significant criterion of the location of village sites within the study area but it was not the only criterion. It is possible that suitable locations near intermittent streams were selected in regions where perennial streams were absent. It is unlikely that confluences or springs were significant draws to the location of village sites. Therefore, an evidence layer representing distances to both perennial and intermittent streams will be considered for the predictive model. Evidence layers representing proximity to confluences and springs will not be included.
The necessity of access to food resources has been a major influence on spatial organization and mobility among hunter-gatherer societies (Jochim 1976; Roper 1979). Theories of behavioral ecology propose that human fitness is in part a product of the energy efficiency of food resource procurement (Bird and O’Connell 2006; Jochim 1981; Smith 1979; Zeanah 2004). One outcome of such theories has been the generation of ecological models of hunter-gatherer settlement based on the principal that village location is strongly affected by the need to optimize the amount of energy expended in the procurement of resources necessary for the maintenance of villages (Binford 1980; Jochim 1976). Therefore, studies based on relationships of humans to the land have emphasized the importance of the temporal and spatial abundance of plant, animal, and mineral resources in determining site location (Roper 1979).

Hunter-gatherer societies often derive greater sustenance from plant foods than meat, leading some researchers to prefer the term “gatherer-hunter” to “hunter-gatherer” (Kelly 1995:65). Within the research area, one of the theoretical bases supporting Horne’s subsistence/settlement model was the primary importance of plant communities in forming the relationship of aboriginal groups to their environment (1981). Yet, despite the abundance of theory supporting the relationship between prehistoric people and plant communities, evidence layers describing those relationships are rare in GIS-based predictive
probabilistic models. One reason for that absence may be the formidable and daunting task of reconstructing maps of past vegetation communities. Over a time depth of several thousand years, factors that affect the distribution of plant communities, such as climate, soils, and landforms, can change significantly. In addition, the incursion of settlement by outside groups can bring land clearing to support agriculture and the introduction of domestic ungulates whose grazing habits reduce the ability of the land to resist erosion, which, in turn, can increase the volume and speed of river flow. With settlement often comes the introduction of non-native plants species that compete with native species and sometimes displace them.

However, in some cases, it is possible to derive meaningful datasets describing relative environmental diversity that can be used as evidential themes in predictive modeling. The creation and use of such themes derived from satellite imagery is one of the possible improvements that Kvamme calls for in the “Second Age of Modeling” (Kvamme 2006:21). It is possible to derive meaningful spatial data describing the relationship between village sites within the study area and plant communities. By limiting the dataset of site locations to ethnohistorically documented villages, it is likely that changes in plant communities have not been so extensive that their current spatial distribution cannot be used as a proxy for plant communities that existed at the time of occupation. Vegetation communities tend to persist where not purposely altered (Aschman 1959; Horne 1981; Woodman et al. 1991). All principal plant
communities found today were mentioned in the same ecological contexts by the earliest explorers (Aschman 1959). Non-native species may out-compete native species, but modern plant communities, such as grasslands and chaparral, will contain members that are analogous to those that existed when the village sites included in this study still thrived (Woodman et al. 1991). Landberg noted that maps covering the study area depicting vegetation communities as they existed prior to European settlement vary “in few details from a map of modern vegetation” except for areas designated as cultivated or urban which were mostly California Prairie (a grassland community of bunchgrasses) and coastal sagebrush (1965:46).

8.1—Proximity to Ecotones

There are aspects of the relationship between hunter-gatherers and plant communities that do not necessarily require detailed reconstruction of the presence, abundance, and resource value of the component plants that make up plant communities. A productive way of viewing the relationship between hunter-gatherers and their environment might be in terms of costs and benefits involved in exploiting locations rather than specific resources. For example, if people sought to minimize travel and resource transport costs, then settlement should favor proximity to the locations most often visited (Jochim 1981).

Ecotones, for example, have been hypothesized to be locations that were attractive to settlements within the Santa Ynez Valley (Breschini 1986; Glassow,
personal communication; Horne 1981). Where resource distributions are in close proximity or overlap, “procurement effort can be minimized by locating settlement in the regions of overlap or juxtaposition” (Jochim 1981:153). In such a case, hunter-gatherers may reduce risk by locating their base camps near secure, low-mobility plant resources and widen their catchments by establishing satellite extraction camps and sending out hunting parties for more mobile, high-prestige resources such as deer (Jochim 1976). In areas where the habitats of high-prestige animal resources do not coincide with productive plant communities it may be more efficient for hunter-gathers to locate settlements with proximity to both.

Proximity to ecotones has on occasion been considered in past settlement analyses and modeling efforts in other geographical regions, but it has not been found to be a significant draw to settlement. In his analysis of data from the Glenwood Project in western Colorado, Kvamme (1985) found that proximity to ecotones was not a critical factor in settlement location for any of the site types included in his research. Warren and Asch (2000) also considered proximity to ecotones and found the correlation to site location to be insufficient to justify retaining that data in their model.

8.2—Proximity to Boundaries of Grassland/Chaparral Habitats

Grasslands and chaparral were a likely focus of food-gathering activities at prehistoric times within the study area (Horne 1981). Those regions provided a wide variety of potential resources such as wild grasses, calindrinia, tarweed, red
maids, and chia from which small seeds could be harvested, yucca hearts, islay, which was one of the most highly-valued plant foods, and a host of others (Smith 1998; Timbrook 2007). Many of those resources are storable, making them particularly valuable for providing sustenance for year-round habitations during late winter and early spring when few plant resources are available for harvest. Chaparral and grasslands were also habitat for animals that were hunted for food such as rabbit, ground squirrels, and mule deer (Horne 1981).

Horne suggests that the resources of the chaparral and grassland habitats were “abundant but diffuse,” which accentuated the need to reduce the transportation costs of gathering (Horne 1981:132). To accommodate that, he suggests that village populations would disperse into smaller task groups, which would establish field camps while extracting resources. He also predicts that short-term hunters’ camps would have been established in mixed chaparral and near the edges of grasslands (1981:128). Based on that scenario, base villages may have been located as nearly as possible to the boundaries of combined grassland and chaparral habitats while also optimizing their proximity to reliable water sources. By settling near the boundaries of these regions, transport costs associated with deployment of task-oriented groups and transportation of extracted resources back to the base village for storage would have been minimized. In addition, such locations would also be in ecotones where other habitats could also be monitored and exploited.
8.3—Ecological Diversity

If ecotones were attractive to prehistoric settlement because they offered access to a greater variety of resources, then locations with close access to a wide variety of habitats may have been an even greater draw. Many parts of the Santa Ynez Valley are mosaics of habitats, each associated with different plant and animal communities. The settlement of areas high in diversity of habitats would have been advantageous if that strategy offered a reduction in risk and increased efficiency.

In general, gathering of plant food resources is a relatively low-risk, high-return economic strategy (Jochim 1976). Proximity of settlements to areas high in ecological diversity would have reduced risk by providing a resource base resistant to environmental fluctuations (Horne 1981). Researchers have attempted to describe possible seasonal rounds in the research area based on seasonal ripening schedules of plants (e.g., Horne 1981; Spanne 1975; Woodman et al. 1991). Being centrally located within a short distance of a variety of habitats would have allowed for the monitoring of more resources so that each could be harvested at the optimal season, thereby increasing efficiency and reducing the risks associated with seasonal environmental variability. In addition, many resources in the Santa Ynez Valley vary in productivity year-to-year depending on short-term fluctuations in rainfall, temperature and other environmental factors. By locating base villages in areas where multiple resources can be
efficiently monitored, the risk that failure of any single crop will have a significant impact on available food resources is reduced.

Close proximity to multiple resources could have been efficient in other ways as well. Even though much of the village population would have dispersed into task groups during harvest times, population segments such as the elderly, infirm, and children would have likely remained in the village (Horne 1981) and close proximity to resources would have allowed for gathering activities within a range accessible to people of relatively restricted mobility. Horne asserts that mobility may have also been decreased by weather during winter when populations would have aggregated (Horne 1981). Close proximity to diverse habitats could have provided more varied and reliable sources of fresh food to supplement stored resources during times of inclement weather.

8.4—Hypotheses and Test Implications

1. *Proximity to ecotones was a significant criterion of village location.*

   Test Implication:

   If the mean distance between village sites and ecotones is significantly less than the mean distance between a random set of points and ecotones then the null hypothesis that village sites were randomly located with respect to ecotones would be rejected and therefore the hypothesis would not be rejected.
2. *Proximity to the boundaries of grassland/chaparral habitats was a significant criterion of village location.*

Test Implication:

If the mean distance between grassland/chaparral habitat boundaries is less than the mean distance between a set of random points and grassland/chaparral habitat boundaries then the null hypothesis that village site location was random with respect to grassland/chaparral habitat boundaries would be rejected and the main hypothesis would not be rejected.

Because grassland/chaparral boundaries are also ecotones, it is necessary to construct additional hypotheses and test implications in order to determine if grassland/chaparral boundaries had significantly greater influence on settlement location than ecotones in general. Hypothesis three below has been constructed to address this issue.

3. *Proximity to ecotones had a significantly greater influence on settlement location than proximity to grassland/chaparral habitat boundaries.*

Test Implication:

If the mean distance between village sites and ecotones is significantly less than the distance between village sites and grassland/chaparral habitat boundaries, then the hypothesis would not be rejected.
4. Settlement locations were favored that had a high level of habitat diversity.

Test Implication:

If the study area is divided into zones of greater and lesser habitat diversity and the mean habitat diversity in the regions around village sites is significantly greater than that of a set of random points, then the null hypothesis that the village sites are distributed randomly with respect to habitat diversity would be rejected and therefore the hypothesis could not be rejected.

8.5—Data and Methods

The data selected to represent habitat regions within the study area is the product of the California GAP Analysis Project published 1998 by the Biogeography Lab at the University of California at Santa Barbara (Figure 7.1). The California GAP Analysis project grew out of the need to identify landcover types and wildlife species that are inadequately represented in existing biodiversity management areas (i.e., the “gaps”). The California GAP Analysis project is part of a larger National GAP Analysis project that is administered by the U.S. Geological Survey (Davis et al. 1998).

The California GAP dataset was produced using a variety of datasets including remote sensing and traditional ground survey. The data sources used and steps taken to compile the maps include:
1. Polygon boundaries were derived from photointerpretation of 1990 Landsat Thematic Mapper digital images and supplemented by 1990 National High Altitude Photography and large scale vegetation maps (Biogeography Lab University of California at Santa Barbara 1998).

2. The vegetation communities of the polygons were identified using Weislander Vegetation Type Maps created in the 1930s and 1940s by a U.S. Forest Service ground survey project spearheaded by Albert Everett Weislander. In this massive project known as the Weislander Mapping Project, regional vegetation types were mapped and classified using actual ground survey (University of California at Berkeley and University of California at Davis 2006).

3. Aerial photo interpretation was used to confirm and enhance attributes.

4. National High-Altitude Photography and NASA-JPL color infrared transparencies were viewed stereoscopically to identify vegetation type, percent coverage, canopy closure, and disturbance.

5. The NHAP photos were at a scale of 1:58,000 and were captured between 1980 and 1984. The NASA imagery was at a scale of 1:65,000 and was captured during the late 1980s and early 1990’s (Biogeography Lab University of California at Santa Barbara 1998).
Another noteworthy aspect of the GAP landcover dataset that makes it particularly useful for inter-disciplinary research is that its habitat regions have attributes for various landcover classification systems such as Jepson bioregion, Wildlife Habitat Relationship (WHR) used by the USDA, and the California Natural Diversity Database or “Holland” system. The USDA has also published a “crosswalk” to allow cross-referencing of WHR classifications into a wide variety of other classification systems (California. Dept. of Forestry and Fire Protection et al. 1988).
Figure 8.1– Gap analysis dataset within the study area
8.6—Methods

In order to extract the data necessary to test the hypotheses and to create the final data required for predictive modeling, three grids were created. The grids and methods by which they were created are as follows:

1. Grid of distance to ecotones.
   a. ET Geowizards toolset was used to convert Gap Analysis polygons to polyline outlines (Tchoukanski 2006). These lines represent the center of ecotones between habitats.
   b. The ArcGIS 9.1 Euclidian Distance tool was used to create a grid of distances to the polylines (Environmental Systems Research Institute Inc. 2006b).
   c. Hawth’s Tools Intersect Point tool was used to attribute distances to ecotones to both the village site feature class and to the feature class of 2100 points distributed randomly throughout the study area (Beyer 2004).

2. Grid of distance to grassland/chaparral boundaries
   a. Annual Grassland, Chamise-Redshank Chaparral, and Mixed Chaparral habitats from the Gap Analysis feature class were extracted and all features were combined (dissolved) to make one polygon of combined grassland and chaparral regions.
b. ET Geowizards toolset was used to the grassland/chaparral polygon to polyline outlines (Tchoukanski 2006).

c. The ArcGIS 9.1 Euclidian Distance tool was used to create a grid of distances to the polylines (Environmental Systems Research Institute Inc. 2006b).

d. Hawth’s Tools Intersect Point tool was used to attribute distances to grassland/chaparral boundaries to both the village site feature class and to the feature class of 2100 points distributed randomly throughout the study area (Beyer 2004).

3. Grid of Relative Environmental Diversity Index (REDI).

   a. Hawth’s Tools was used to create a staggered sampling grid of points 100 meters apart covering the entire study area (Beyer 2004).

   b. The number of habitats within a two-kilometer radius of each point was counted.

      i. A two kilometer buffer of each point was created using ArcGIS 9.1 (Environmental Systems Research Institute Inc. 2006b). Note: The term “buffer” is defined as a shape that is drawn at a specified distance around a given point, line, or polygon.
ii. The “Polygon in Polygon Analysis” tool, which is a part of Hawth’s Tools, was used to tabulate the number of unique Gap Analysis habitats that each buffer overlapped (Beyer 2004). That created an attribute with values ranging from one to nine.

iii. ET Geowizards “Polygon to Point Conversion” tool was used to create a point feature class of the centroids and attributes of each buffer (Tchoukanski 2006).

iv. The Nearest Neighbor spatial interpolation method was used to create the final 30-meter Relative Environmental Diversity Index raster.

8.7—Discussion of the Methods for Creating the Relative Environmental Diversity Index Grid

In determining how to best calculate an appropriate measure of regional ecological diversity, certain methodological choices were made that warrant further discussion. Three of the parameters chosen, specifically the shape and size of the search radii and the use of spatial interpolation significantly impact the magnitude and range of the values included in the diversity raster. The specific parameters used were chosen so that the most accurate Relative Environmental Diversity Index raster could be created in the most efficient way possible.
Regional studies that involve diversity of resources fall generally under the heading of catchment analysis. Catchment analysis is a process whereby the availability, abundance, spacing, and seasonality of plant, animal, and mineral resources are analyzed for the entire area from which a site population derived those resources (Roper 1979). The basic principles of site catchment analysis deal with the balance between energy investment and economic return—the further a resource patch is from a settlement, the greater the energy investment in travel to and from that patch such that, as patches are considered that are further and further from the settlement, eventually an economic boundary will be reached beyond which exploitation is not sensible from a standpoint of energy return (Gaffney and Stancic 1991). Though the methods used to create the Relative Environmental Diversity Index raster may in some ways resemble catchment analysis, the individual buffers are not intended to represent catchments. It is beyond the scope of this study to attempt to reconstruct the catchment potential from any given point within the Santa Ynez River watershed. The hypotheses as stated require the creation of a raster whereby at any given point the relative diversity of habitats available in the immediate vicinity can be quantified.

The decision to use two-kilometer radii as the basis for deriving habitat counts was based on testing various radii ranging in size from 500m to 10km. In addition, discussions of catchment size in regions adjacent to the study area were consulted in order to assure that the distance selected was not significantly greater
than a reasonable foraging radius of hunter-gatherers in the environmental context of the study area.

Sources that discuss catchment size within the entire Chumash region are few and generally in disagreement. In the Malibu Creek sub-region east of the study area, Tartaglia wrote that the number of village sites per five-kilometer radius was less than half that expected, leading him to hypothesize that catchments in that region were considerably smaller than five kilometers (1980:189). From another perspective, the most appropriate radius may be that of the shortest foraging distance to which inhabitants might be restricted at various times of the year. Horne writes that during winter, it would have been most advantageous for food resources to have been within a twenty minute walk (Horne 1981) which, based on personal experience can represent a kilometer or less. Horne’s statement seems to have been based perhaps on the most mountainous region of his study area whereas most of our study area is considerably more clement and therefore it would be expected that winter travel times would be little affected by weather if at all.

A somewhat similar concept to the one being proposed here was employed by Schermer and Tiffany wherein a measure of environmental diversity was determined by counting environmental variables (such as landform, elevation, exposure, water, and vegetation) within fixed radii of 108 Woodland period archaeological sites in the Iowa Valley (Schermer and Tiffany 1985). Within their study area, they compared the use of 100 m, 500 m, and 1 km radii and
determined that even though the mean number of resources varied directly with the radius used, the hypothesis that sites were located in areas of greater environmental diversity was supported regardless of which radius was used.

In a similar test, radii of 500 m, 1 km, 2 km, 3 km, 5 km, and 10 km were compared within our study area. The mean Relative Environmental Diversity Index calculated for the village sites within our study using each of those radii were compared with those of 200 points distributed randomly throughout the study area. Whereas in all cases the mean Relative Environmental Diversity Index of the village sites was greater than that of the random points, at radii greater than 2 km the significance of that difference was not great enough to reject the null hypothesis (Table 8.1). It seems that this is due to local variations in diversity being obscured when relatively large radii are used. The 2 km radius selected for this study reflects the largest radius that yields a sufficiently articulated raster and is within the foraging range proposed by past researchers working in the Chumash region.

<table>
<thead>
<tr>
<th>Radius</th>
<th>Mean Diversity Index</th>
<th>t value</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Village Sites</td>
<td>Random Points</td>
<td></td>
</tr>
<tr>
<td>500 m</td>
<td>2.43</td>
<td>1.75</td>
<td>-2.821</td>
</tr>
<tr>
<td>1 km</td>
<td>3.05</td>
<td>2.47</td>
<td>-2.402</td>
</tr>
<tr>
<td>2 km</td>
<td>4.38</td>
<td>3.66</td>
<td>-2.290</td>
</tr>
<tr>
<td>3 km</td>
<td>5.19</td>
<td>4.62</td>
<td>-1.500</td>
</tr>
<tr>
<td>5 km</td>
<td>6.48</td>
<td>6.12</td>
<td>-0.763</td>
</tr>
<tr>
<td>10 km</td>
<td>9.86</td>
<td>9.30</td>
<td>-1.438</td>
</tr>
</tbody>
</table>

Table 8.1—Mean Relative Environmental Diversity Index and various radii.
The use of spatial interpolation was adopted in order to more efficiently produce the final Relative Environmental Diversity Index raster at a finer resolution so as to more closely match the other evidence layers in the study. The calculation of counts of habitats that overlap buffers is a particularly processor-intensive computational task. The 100-meter point feature class contained over 270,005 points and calculating the necessary data required over 230 hours of computer time. The calculation of the number of points that would have been necessary to yield directly measured data at the resolution of the rest of the data in the study (approximately ten meter cells) would have been both untenable and unnecessary.

Spatial interpolation is a technique pervasive in GIS whereby the value of unmeasured points are estimated based on the value of measured points using spatial statistics. It has been used in a wide variety of applications with continuous field data (Longley 2005). Even though the original Gap Analysis feature class is vector and therefore not continuous, the data concerning the relative number of habitats is continuous, so the use of spatial interpolation is appropriate. In addition, the depiction of habitats as polygons with sharp, immediate boundaries is a necessary simplification of the gradual and variable transition between habitats that is more common in nature. Therefore, though it is not within the scope of this project to model transitions between habitats, it may
be said that using interpolation techniques that create more gradual transitions between habitats is appropriate.

The fundamental tenet of spatial interpolation is based in Tobler’s First Law of Geography: “Everything is related to everything else, but near things are more related than distant things” (Longley 2005:65). Numerous methods have been developed using a variety of formulas used to determine the value of intermediate, unmeasured points based on the value of measured points by applying greater weight to nearer values that more distant ones. Three widely used interpolation methods are considered here—Inverse Distance Weighting (IDW), Kriging, and Natural Neighbors.

Inverse Distance Weighting (IDW) determines the value of an unmeasured point by first measuring the distances to surrounding known points and assigning weights to each based on the reciprocal of the distance squared (1/D^2), therefore the closer points are weighted more heavily than relatively distant ones. IDW is best suited for data sets where known data do not follow strict regional trends (Berry 2006b).

Kriging is a mathematically more complex technique that is firmly grounded in geostatistical theory (Longley 2005). In essence, the technique develops custom weights based on trends in the values of known points. Kriging is most appropriate for dense, systematically sampled data that exhibits regional trends (Berry 2006b).
The Natural Neighbors technique uses only the values within a subset of samples that surround a query point. Thiessen polygons are created for all known data points. A new Thiessen polygon is created for the unknown point being interpolated, and the values of neighboring points are weighted based on the amount of overlap between the polygons. Natural Neighbors is a robust technique particularly suitable for large datasets (Environmental Systems Research Institute Inc. 2006c).

Rasters were created using each of the three techniques and variations of parameters within the techniques and tested to determine the relative accuracies of each. Testing was done by comparing the actual values of a test set of known points to their interpolated values. The difference at a given point between its known value and its interpolated value is termed the “residual.” The unsigned average (“Average Unsigned Residual”) is a measure of the average magnitude of error. Comparing Average Unsigned Residuals is a widely accepted method of assessing interpolation accuracy (Berry 2006a).

The Natural Neighbor method was chosen because it had one of the lowest Average Unsigned Residuals (0.03) and also yielded the greatest percentage of points where the interpolated exactly matched the actual value (Table 8.2). The final datasets representing distance to ecotones, distance to grassland/chaparral boundaries, and Relative Environmental Diversity Index values are depicted in Figures 8.2 – 8.4.
8.8—Results

The results of hypothesis testing are as follows:

1. *Proximity to ecotones was a significant criterion of village location.*

   After transforming the data to equalize variances, the 279 m difference between the mean distance to ecotones of a test set of 2100 points distributed randomly throughout the study area (512.3 m) and that of village sites (233.3 m) is very significant \( (t = -3.205, 0 < p < .005) \). Therefore the null hypothesis that village sites are distributed randomly with respect to ecotones is rejected and the hypothesis is not rejected.

2. *Proximity to the boundaries of grassland/chaparral habitats was a significant criterion of village location.*

   After transforming the data to equalize variances, the 432.9 m difference between the mean distance to grassland/chaparral boundaries of a test set of 2100 points distributed randomly throughout the study area (919.27 m) and that of village sites (486.37 m) is very significant \( (t = -3.454, p = .002) \). Therefore the null hypothesis that village sites are distributed randomly with respect to grassland/chaparral boundaries is rejected and therefore the hypothesis cannot be rejected.

---

Table 8.2 - Results of Unsigned Residual testing.

<table>
<thead>
<tr>
<th>Method</th>
<th>Average Unsigned Residual</th>
<th>Maximum Unsigned Residual</th>
<th>% Exact</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDW:</td>
<td>0.03</td>
<td>0.04</td>
<td>85.5%</td>
</tr>
<tr>
<td>Power = 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius=Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Points=6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDW:</td>
<td>0.78</td>
<td>0.71</td>
<td>74.0%</td>
</tr>
<tr>
<td>Power = 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius=Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Points=12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krigging:</td>
<td>0.04</td>
<td>0.75</td>
<td>89.0%</td>
</tr>
<tr>
<td>Method=Ordinary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semivariogram=S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius=Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Points=4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Krigging:</td>
<td>0.05</td>
<td>0.50</td>
<td>82.0%</td>
</tr>
<tr>
<td>Method=Ordinary</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Semivariogram=S</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spherical</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radius=Variable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Points=8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural Neighbor</td>
<td>0.68</td>
<td>0.68</td>
<td>88.0%</td>
</tr>
</tbody>
</table>

87
3. Proximity to ecotones had a significantly greater influence on settlement location than proximity to grassland/chaparral habitat boundaries. After transforming the data to equalize variances, the 253.1 m difference between the mean distance between village sites and grassland/chaparral boundaries (486.4 m) and the mean distance between village sites and ecotones (233.3 m) is only moderately significant (t = –1.662, p = 0.104). Therefore the null hypothesis that differences in the two means could be due to the vagaries of sampling cannot be rejected.

4. Settlement locations were favored that had a high level of habitat diversity.

    The .69 difference between the mean Relative Environmental Diversity Index of villages (4.31) and that of a set of 2100 points distributed randomly throughout the study area (3.62) is significant (t = –2.287, p = .022) therefore the null hypothesis that village sites are located randomly with respect to habitat diversity is rejected and the hypothesis cannot be rejected.

8.9—Conclusions

Based on the results of these tests it can be stated that proximity to ecotones, grassland/chaparral boundaries, and regions having relatively greater habitat diversity all may have influenced village site location within the study area. Therefore, each of these layers of evidence will be incorporated into the model-building phase. The difference in the relative influence on village site
location of ecotones versus grassland/chaparral boundaries is not significant, so the two layers of evidence will be compared in the modeling phase of the study in order to select one over the other for inclusion in the final model.
Figure 8.2 – Distances to ecotones within the study area.
Figure 8.3 – Distances to Grassland/Chaparral boundaries within the study area.
Figure 8.4 – Relative environmental Diversity Index (REDI) within the study area.
9.0—SLOPE

Slope is a very common layer of evidence in predictive probabilistic models (Dalla Bona 2000; Duncan and Beckman 2000; Kvamme 1985; Mink et al. 2006; Pilgram 1987; Warren and Asch 2000; Wescott and Kuiper 2000). Kvamme, for example, found slope to be a significant factor affecting site location. In the site location model he created using data from the Glenwood Project area in western Colorado, he found that sites showing evidence of extended activity (which are most closely analogous to the village sites included in this study) were found on landforms with an average of 16.1% grade (9º) slope and were always at slopes less than 40% grade (21.8º) whereas general slopes in the area were as steep as 100% (Kvamme 1985:233).

A slope raster is quite easy to create from a digital elevation model using modern GIS software. For the purposes of this study, a raster was created of slope in degrees using the Slope tool that is supplied with ArcGIS 9.1 (Environmental Systems Research Institute Inc. 2006b) and the National Elevation Dataset described earlier (U.S. Geological Survey (USGS) EROS Data Center 1999) as a digital elevation model. Slope is calculated for a given cell by fitting a plane that passes through the eight cells that surround it and deriving the slope based on that plane’s rise and run (Environmental Systems Research Institute Inc. 2006c).
9.1—Hypothesis

Areas of low slope were preferred for village site location within the study area.

Test Implication

The mean slope of the village sites will be significantly less than that of a set of points distributed randomly throughout the study area.

9.2—Results

The 11.41° difference between the mean slope of the village sites (5.58°) and that of a set of 2100 points distributed randomly throughout the study area (16.99°) is highly significant ($t = -12.279, .005 < p > 0$). Therefore, the null hypothesis that village sites within the study area are distributed randomly with respect to slope can be rejected and the hypothesis cannot be rejected. The dataset representing slope in the study area is depicted in Figure 9.1.
Figure 9.1 - Slope within the study area.
10.0—THE MODEL

One of the potential benefits of a predictive probabilistic model is the production of maps that can be used to guide Cultural Resource Managers as to which regions are most likely to contain undiscovered sites. To that end, the Weights of Evidence method was used to combine the data layers in such a way as to reflect the relative influence of each layer on settlement distribution. Though other methods such as Logistic Regression have also been applied to predictive probabilistic models (Agterberg et al. 1993; Warren and Asch 2000), the Weights of Evidence technique was chosen for this study because the weights and statistics generated through the technique are relatively easy to interpret (Raines 2000). In addition, a comprehensive toolset for use with ArcGIS, entitled Spatial Data Modeler (SDM), has been created to support the method and is publicly available at no charge (Sawatzky et al. 2004).

One of the earliest applications of Weights of Evidence was in health care as a decision-making tool to relate complex constellations of symptoms to causal diseases (Raines 2000). Beginning sometime in the late 1980s Weights of Evidence drew wider attention in the geologic communities for its ability to map mineral potential (Agterberg et al. 1993; Bonham-Carter et al. 1988; Harris et al. 2001; Raines 1999). The technique has since been used in other arenas including prediction of the spread of pollutants in aquifers (Masetti et al. 2007), prediction of the viability of crop-specific agriculture as a decision-making tool for farmers in developing countries (O'Brien 2004), and in the prediction of regional potential.
for the discovery of archaeological sites as a tool for research and cultural resource management (Dalla Bona 2000; Ford and Clarke 2005; Rosenthal et al. 2003).

10.1—The Process of Mapping Archaeological Sensitivity Using Weights of Evidence

The process used in this study to map the relative potential of regions to contain undiscovered village sites can be grouped into five steps.

Step One—Collection and evaluation of datasets.

This step has been thoroughly discussed in previous sections. It is undoubtedly the most important step in the generation of a predictive probabilistic model. A set of points is also generated representing locations where the phenomena being modeled, in this case village sites, are located. It is generally preferable that this set, termed “training points,” is a randomly selected subset of a larger group of occurrences so that the model can be tested using a set of observances independent of the training points. In the case of this study the total number of known points (n=21) is sufficiently small that the entire set was used as training points, and alternative methods of model testing are employed that do not depend on independent observations.

Step Two—Generalization of datasets into binary layers of evidence.
It is considered preferable to generalize datasets into evidence layers containing only regions of presence or absence of the condition that is hypothesized to influence the spatial distribution of training points. For example, if it is hypothesized that the training points were located within 300 meters of drainages, then an appropriate evidence layer would be a raster where all cells that are within 300 meters of the closest drainage (the area where sites are predicted to be present) would have the value one, and all cells that are further than 300 meters from the closest drainage (the area where sites are predicted to be absent) would have the value zero.

One advantage of the Weights of Evidence method is that the weights and confidence statistics can be easily understood and interpreted (Raines 2000). Positive weight values for a particular layer of evidence indicate that more training points occur in the “presence” area of that layer than would occur due to chance. Negative values mean that fewer training points occur. Analogous weights are calculated for the “absence” areas. The difference in the presence and absence weight values is termed “contrast” and is useful as a measure of the importance of the theme to the overall model. Positive contrasts indicate that a layer of evidence has at least some predictive power (Raines 2000). Values between 0 and 0.5 are considered to be weakly predictive. Values between 0.5 and 1.0 are moderately predictive, and values greater than 1.0 are strongly predictive (Rosenthal et al. 2003).
The Spatial Data Modeler (ArcSDM) toolset contains a powerful set of tools to generate tables of weights and estimates of confidence that allow datasets to be analyzed so that the most influential ranges of values can be selected (Sawatzky et al. 2004). ArcSDM also has the ability to calculate the ratio of contrast to standard deviation, termed “Studentized Contrast” (Sawatzky et al. 2004), which serves as a measure of confidence levels of the association between the training points and the layer of evidence. This measure is also been referred to as the “Normalized Contrast” (Raines 1999:259) and is said to function similar to a Student’s T-Test. Values greater than 1.96 are associated with a 97.5% or greater confidence level (Raines 1999).

In order to facilitate generalization of the datasets included in this study into binary evidence layers each was first divided (“classified”) into many arbitrary ranges of values. For example, the dataset representing Ecotone Proximity was divided into 14 classes (Table 9.3). For each classified dataset, tables of weights and confidences associated with each class were generated (Tables 9.2 - 9.6).

The following thresholds were used as a guideline in reducing datasets into binary layers of evidence.

1. Classes that had the highest contrast grouped to make the “site presence” areas of the final binary evidence layers. If no combination of classes of a particular dataset could be grouped
such that the resulting evidence layer would have a contrast of at
least 1.0, then the layer was rejected.

2. So long as criteria one and two above are met, datasets will be
classified so as to contain the greatest number of village sites in the
smallest area.

Full tables of the classes used to analyze each dataset and the weights
associated with each can be seen Tables 10.2—10.6. The weights, contrasts, and
confidence measures associated with the final binary evidence layers are as
follows:

<table>
<thead>
<tr>
<th>Area Sq km</th>
<th>#Points</th>
<th>W+</th>
<th>W-</th>
<th>Contrast</th>
<th>stud(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>REDI 4 through 7</td>
<td>1194.70</td>
<td>16</td>
<td>0.39</td>
<td>-0.72</td>
<td>1.11</td>
</tr>
<tr>
<td>&lt; 200m to Ecotones</td>
<td>788.01</td>
<td>17</td>
<td>0.87</td>
<td>-1.25</td>
<td>2.12</td>
</tr>
<tr>
<td>&lt;800m to Perennial Streams</td>
<td>747.28</td>
<td>18</td>
<td>0.98</td>
<td>-1.56</td>
<td>2.54</td>
</tr>
<tr>
<td>&lt;15º Slope</td>
<td>995.54</td>
<td>18</td>
<td>0.69</td>
<td>-1.39</td>
<td>2.09</td>
</tr>
</tbody>
</table>

Table 10.1—Final model weights.

It is noteworthy that the Relative Environmental Diversity Index dataset
had somewhat anomalous weights. Classes four and seven showed sufficient
contrast but classes five and six did not. No logical justification could be
determined to exclude categories five and six, so the binary dataset was generated
using all four classes. Weights were calculated for the resulting layer of evidence,
which passed all criteria so it was retained in the model.
Grassland/chaparral boundaries are coincident with ecotones so it was decided that only one of the two datasets could be retained. The distance to ecotones dataset had higher contrasts and so was retained.

It was not possible to classify the distance to intermittent streams dataset so as to create a layer of evidence with sufficient contrast so that dataset was not included in the final model.

<table>
<thead>
<tr>
<th>Relative Environmental Diversity Index Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>REDI</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

Table 10.2—Weights associated with Relative Environmental Diversity Index.

<table>
<thead>
<tr>
<th>Ecotone Proximity Weights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance in Meters</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>0 - 50</td>
</tr>
<tr>
<td>50 - 100</td>
</tr>
<tr>
<td>100 - 200</td>
</tr>
<tr>
<td>200 - 300</td>
</tr>
<tr>
<td>300 - 400</td>
</tr>
<tr>
<td>400 - 500</td>
</tr>
<tr>
<td>500 - 600</td>
</tr>
<tr>
<td>600 - 700</td>
</tr>
<tr>
<td>700 - 800</td>
</tr>
<tr>
<td>800 - 900</td>
</tr>
<tr>
<td>900 - 1000</td>
</tr>
<tr>
<td>1000 - 1500</td>
</tr>
<tr>
<td>1500 - 2000</td>
</tr>
<tr>
<td>2000 - 3476.47</td>
</tr>
</tbody>
</table>

Table 10.3—Weights associated with distance to ecotones dataset.
Table 10.4—Weights associated with distance to intermittent streams dataset.

<table>
<thead>
<tr>
<th>Distance in Meters</th>
<th>Class</th>
<th>Area Sq km</th>
<th>#Points</th>
<th>W+</th>
<th>W-</th>
<th>C</th>
<th>stud(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 50</td>
<td>1</td>
<td>207.4576</td>
<td>2</td>
<td>0.0690</td>
<td>-0.0070</td>
<td>0.0760</td>
<td>0.1023</td>
</tr>
<tr>
<td>50 - 100</td>
<td>2</td>
<td>410.8514</td>
<td>6</td>
<td>0.4844</td>
<td>-0.1429</td>
<td>0.6273</td>
<td>1.2985</td>
</tr>
<tr>
<td>100 - 200</td>
<td>3</td>
<td>806.0486</td>
<td>10</td>
<td>0.3213</td>
<td>-0.2230</td>
<td>0.5443</td>
<td>1.2456</td>
</tr>
<tr>
<td>200 - 300</td>
<td>4</td>
<td>1170.4211</td>
<td>14</td>
<td>0.2848</td>
<td>-0.4026</td>
<td>0.6873</td>
<td>1.4848</td>
</tr>
<tr>
<td>300 - 400</td>
<td>5</td>
<td>1476.5063</td>
<td>15</td>
<td>0.1214</td>
<td>-0.2515</td>
<td>0.3729</td>
<td>0.7719</td>
</tr>
<tr>
<td>400 - 500</td>
<td>6</td>
<td>1712.1689</td>
<td>17</td>
<td>0.0985</td>
<td>-0.3356</td>
<td>0.4341</td>
<td>0.7811</td>
</tr>
<tr>
<td>500 - 600</td>
<td>7</td>
<td>1882.2067</td>
<td>17</td>
<td>0.0038</td>
<td>-0.0161</td>
<td>0.0199</td>
<td>0.0359</td>
</tr>
<tr>
<td>600 - 700</td>
<td>8</td>
<td>2002.2317</td>
<td>17</td>
<td>-0.0580</td>
<td>0.2927</td>
<td>-0.3507</td>
<td>-0.6310</td>
</tr>
<tr>
<td>700 - 800</td>
<td>9</td>
<td>2086.3891</td>
<td>19</td>
<td>0.0121</td>
<td>-0.1079</td>
<td>0.1200</td>
<td>0.1614</td>
</tr>
<tr>
<td>800 - 900</td>
<td>10</td>
<td>2148.5126</td>
<td>19</td>
<td>-0.0173</td>
<td>0.1810</td>
<td>-0.1983</td>
<td>-0.2667</td>
</tr>
<tr>
<td>900 - 1000</td>
<td>11</td>
<td>2194.4938</td>
<td>19</td>
<td>-0.0385</td>
<td>0.4659</td>
<td>-0.5044</td>
<td>-0.6785</td>
</tr>
<tr>
<td>1000 - 2000</td>
<td>12</td>
<td>2328.2436</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2000 - 2477.26</td>
<td>13</td>
<td>2333.8682</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 10.5—Weights associated with distance to boundaries of grassland/chaparral regions.

<table>
<thead>
<tr>
<th>Distance in Meters</th>
<th>Class</th>
<th>Area Sq km</th>
<th>#Points</th>
<th>W+</th>
<th>W-</th>
<th>C</th>
<th>stud(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 50</td>
<td>1</td>
<td>112.8936</td>
<td>6</td>
<td>1.7765</td>
<td>-0.2869</td>
<td>2.0634</td>
<td>4.2709</td>
</tr>
<tr>
<td>50 - 100</td>
<td>2</td>
<td>224.4888</td>
<td>8</td>
<td>1.3767</td>
<td>-0.3785</td>
<td>1.7552</td>
<td>3.9055</td>
</tr>
<tr>
<td>100 - 200</td>
<td>3</td>
<td>434.4847</td>
<td>11</td>
<td>1.0347</td>
<td>-0.5360</td>
<td>1.5707</td>
<td>3.5946</td>
</tr>
<tr>
<td>200 - 300</td>
<td>4</td>
<td>623.7243</td>
<td>12</td>
<td>0.7601</td>
<td>-0.5364</td>
<td>1.2965</td>
<td>2.9400</td>
</tr>
<tr>
<td>300 - 400</td>
<td>5</td>
<td>796.2174</td>
<td>12</td>
<td>0.5159</td>
<td>-0.4301</td>
<td>0.9460</td>
<td>2.1452</td>
</tr>
<tr>
<td>400 - 500</td>
<td>6</td>
<td>952.7815</td>
<td>14</td>
<td>0.4905</td>
<td>-0.5740</td>
<td>1.0646</td>
<td>2.2996</td>
</tr>
<tr>
<td>500 - 600</td>
<td>7</td>
<td>1092.0791</td>
<td>14</td>
<td>0.3541</td>
<td>-0.4677</td>
<td>0.8218</td>
<td>1.7752</td>
</tr>
<tr>
<td>600 - 700</td>
<td>8</td>
<td>1218.6311</td>
<td>15</td>
<td>0.3134</td>
<td>-0.5144</td>
<td>0.8278</td>
<td>1.7137</td>
</tr>
<tr>
<td>700 - 800</td>
<td>9</td>
<td>1330.5866</td>
<td>16</td>
<td>0.2901</td>
<td>-0.5909</td>
<td>0.8810</td>
<td>1.7195</td>
</tr>
<tr>
<td>800 - 900</td>
<td>10</td>
<td>1431.4500</td>
<td>16</td>
<td>0.2170</td>
<td>-0.4850</td>
<td>0.7020</td>
<td>1.3701</td>
</tr>
<tr>
<td>900 - 1000</td>
<td>11</td>
<td>1522.9300</td>
<td>16</td>
<td>0.1550</td>
<td>-0.3781</td>
<td>0.5332</td>
<td>1.0406</td>
</tr>
<tr>
<td>1000 - 1500</td>
<td>12</td>
<td>1852.7752</td>
<td>20</td>
<td>0.1821</td>
<td>-1.4656</td>
<td>1.6477</td>
<td>1.6080</td>
</tr>
<tr>
<td>1500 - 2000</td>
<td>13</td>
<td>2055.4020</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>2000 - 3000</td>
<td>14</td>
<td>2255.4169</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>3000 - 5177.50</td>
<td>15</td>
<td>2333.8682</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
### Distance to Perennial Streams Weights

<table>
<thead>
<tr>
<th>Distance in Meters</th>
<th>Class</th>
<th>Area Sq km</th>
<th>#Points</th>
<th>W+</th>
<th>W-</th>
<th>C</th>
<th>stud(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 50</td>
<td>1</td>
<td>56.8328</td>
<td>3</td>
<td>1.7697</td>
<td>-0.1295</td>
<td>1.8992</td>
<td>3.0448</td>
</tr>
<tr>
<td>50 - 100</td>
<td>2</td>
<td>111.4475</td>
<td>9</td>
<td>2.1951</td>
<td>-0.5107</td>
<td>2.7059</td>
<td>6.1350</td>
</tr>
<tr>
<td>100 - 200</td>
<td>3</td>
<td>216.3113</td>
<td>12</td>
<td>1.8194</td>
<td>-0.7501</td>
<td>2.5695</td>
<td>5.8264</td>
</tr>
<tr>
<td>200 - 300</td>
<td>4</td>
<td>316.4416</td>
<td>14</td>
<td>1.5930</td>
<td>-0.9530</td>
<td>2.5460</td>
<td>5.4995</td>
</tr>
<tr>
<td>300 - 400</td>
<td>5</td>
<td>412.6360</td>
<td>16</td>
<td>1.4611</td>
<td>-1.2406</td>
<td>2.7017</td>
<td>5.2729</td>
</tr>
<tr>
<td>400 - 500</td>
<td>6</td>
<td>504.0370</td>
<td>17</td>
<td>1.3216</td>
<td>-1.4150</td>
<td>2.7366</td>
<td>4.9242</td>
</tr>
<tr>
<td>500 - 600</td>
<td>7</td>
<td>590.0705</td>
<td>17</td>
<td>1.1640</td>
<td>-1.3668</td>
<td>2.5308</td>
<td>4.5539</td>
</tr>
<tr>
<td>600 - 700</td>
<td>8</td>
<td>671.4871</td>
<td>17</td>
<td>1.0347</td>
<td>-1.3190</td>
<td>2.3537</td>
<td>4.2353</td>
</tr>
<tr>
<td>700 - 800</td>
<td>9</td>
<td>747.2850</td>
<td>18</td>
<td>0.9849</td>
<td>-1.5601</td>
<td>2.5449</td>
<td>4.0809</td>
</tr>
<tr>
<td>800 - 900</td>
<td>10</td>
<td>817.6752</td>
<td>18</td>
<td>0.8948</td>
<td>-1.5147</td>
<td>2.4095</td>
<td>3.8637</td>
</tr>
<tr>
<td>900 - 1000</td>
<td>11</td>
<td>882.3159</td>
<td>18</td>
<td>0.8187</td>
<td>-1.4711</td>
<td>2.2899</td>
<td>3.6719</td>
</tr>
<tr>
<td>1000 - 2000</td>
<td>12</td>
<td>1311.2970</td>
<td>19</td>
<td>0.4765</td>
<td>-1.5263</td>
<td>2.0028</td>
<td>2.6941</td>
</tr>
<tr>
<td>2000 - 3000</td>
<td>13</td>
<td>1530.5975</td>
<td>19</td>
<td>0.3219</td>
<td>-1.2849</td>
<td>1.6068</td>
<td>2.1614</td>
</tr>
<tr>
<td>3000 - 5000</td>
<td>14</td>
<td>1785.7137</td>
<td>19</td>
<td>0.1677</td>
<td>-0.9029</td>
<td>1.0706</td>
<td>1.4401</td>
</tr>
<tr>
<td>5000 - 10000</td>
<td>15</td>
<td>2138.4634</td>
<td>19</td>
<td>-0.0126</td>
<td>0.1282</td>
<td>-0.1408</td>
<td>-0.1894</td>
</tr>
<tr>
<td>10000 - 15798.69</td>
<td>16</td>
<td>2333.8682</td>
<td>21</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Table 10.6—Weights associated with distance to perennial streams.

### Step Three—Generation of the Response Raster.

The four evidence layers listed in Table 10.1 were combined to generate a response raster (Table 10.1). Each unique combination of layers (termed “unique condition”) is given a number. In the response raster each cell contains the number of the unique condition it falls on. A table of data is generated listing the number of each unique condition along with the probability that a given cell within that unique condition will be located on a site (termed “posterior probability”). In order to generate a usable sensitivity map, the raster was ranked (classified) by relative posterior probability. In this context the term “class” is defined as a group of cells within a certain range of values that are grouped so as
to create geographic zones. The final sensitivity map was generated by classifying the response raster by posterior probability into five classes using the Jenks Natural Breaks method, representing areas ranging from “Very High” to “Very Low” relative probability to contain undiscovered village sites (Figure 10.1).

Step Four—Conditional independence testing and assessment.

Conditional independence is a concern when using Bayesian techniques such as weights of evidence. Probabilities for each unique combination of layers of evidence are derived through adding the probabilities associated with each layer and therefore conditional dependency between layers exaggerates those probabilities. Bonham-Carter (1994) sets a threshold of 0.85 as being the lowest acceptable Conditional Independence Ratio.

Violation of conditional independence thresholds is not universally held to invalidate weights of evidence models. Gary Raines, a notable expert in weights of evidence modeling and one of the authors of the ArcSDM toolset has stated:

Because conditional independence is never satisfied completely, the posterior probabilities are usually overestimated in absolute terms. However, the relative variations in posterior probability (as seen in spatial patterns on the response map) are usually not much affected by violations of this assumption (2000:45).
Posterior probabilities of models that violate conditional dependence thresholds can still be used in relative terms to rank geographic regions and are therefore still valid (O'Brien 2004; Raines 2000).

The Conditional Independence Ratio of the final model with all four datasets was 0.628, which is below the threshold set by Bonham-Carter (1994). Pairwise tests were conducted of all possible combinations of datasets, and whereas some conditional dependence was present in all combinations, no combination was found which had a Conditional Independence Ratio less than 0.85. In addition, every combination of three datasets was tested and none of them showed sufficient conditional independence to exceed the Conditional Independence Ratio threshold. Further model testing demonstrated that the most precise and accurate model resulted from the inclusion of all four datasets as separate layers of evidence, so it was decided that the final model was valid despite its violation of conditional independence thresholds so long as it is used to rank relative sensitivity of regions, which is the goal of this model, rather than to determine absolute probability of an undiscovered village site being located at a particular piece of land.

Step Five—Response Raster Testing

Perhaps the most commonly reported statistic describing the power of a predictive probabilistic model is that of accuracy. That statistic is usually
reported as the percentage of sites that are correctly identified by the response raster. Accuracy stated that way has little meaning unless the amount of area that is covered by the predicting classification is considered as well. In that sense, all response rasters can be said to have two qualities that affect their overall utility—accuracy and precision. Accuracy is defined as the percentage of points correctly predicted by a class of the response raster. Precision is the proportion of the study area that is covered by that class. Neither accuracy nor precision alone are adequate measures of a response raster’s utility. For example, if a hypothetical class covers the entire study area, it will always contain all points and therefore will be 100% accurate, but it will be of little utility. Similarly, if a class covers only one 10 m-square cell that contains one site then it can be said to have maximum precision, but a very low percentage of sites are predicted and again the model is of little utility.

A simple and effective measure of response raster utility suggested by Kvamme (1988) is generally referred to as Kvamme’s Gain Statistic (KGS). It is used to measure the gain in effectiveness of a given response raster over a “purely random model of no predictive capacity” (Kvamme 1988:329). The formula is expressed as:

\[
Gain = 1 - \left( \frac{PA}{PS} \right)
\]

Where \( PA \) = Percentage of total area covered by a particular class within the response raster and \( PS \) = Percentage of total sites within that class (Kvamme 1988:329).
Positive values have some gain over a purely random model, and as values approach 1.0 the response raster has increased predictive utility. Negative values actually would have reverse predictive value (Kvamme 1988). One particularly valuable aspect of this formula is its utility in comparing response rasters.

All response rasters that were generated during the conditional independence testing phase were compared using Kvamme’s Gain Statistic (KGS), and the final response raster that was generated using all four layers of evidence (Figure 9.1) was found to yield the highest KGS value (0.714). Therefore, despite conditional dependence, the response raster is held to be the response raster of highest utility. The final response raster correctly predicted 80.9% of the sites in 23.1% of the study area, which is the combined area of the top two classes, titled “Very High” and “High” sensitivity. The highest-ranking class, “Very High,” had the highest utility with a KGS of 0.907. It correctly predicted 71.4% of the sites in 6.6% of the study area.

It can be argued that the two most eastern sites, Siguaya and Snajalegua, represent geographic outliers that may have been located according to different environmental criteria. That possibility may be supported in that they are the furthest from perennial water (11,817.7m and 13,122.7m respectively), two of the three that are furthest from ecotones (805.2m and 554.0m respectively), and are in a region of relatively low habitat diversity (Relative Environmental Diversity Index =2). Even if the data concerning perennial streams is inaccurate, as
previously discussed, it still remains that these two sites are in very different environmental contexts from the other village sites. If those two sites are removed from the model, then the KGS would be 0.741 with 89.5% of the sites being correctly predicted in 23.1% of the area. The KGS of the highest-rank

The KGS of the highest-ranking class would be 0.916 with 78.9% of the sites correctly predicted in 6.6% of the area.
Relative Potential for Primary Habitation Sites
Within the Santa Ynez Valley

Legend
Relative Potential
- Red: Very High
- Orange: High
- Green: Moderate
- Light Blue: Low
- Blue: Very Low

Kilometers

Figure 10.1—Map of relative potential for primary habitation sites within the Santa Ynez Valley
Relative Potential for Year-Round Habitation Sites in the Santa Ynez Valley Showing All Unique Conditions

Legend
Each unique condition in order of sensitivity

1 (highest) 6
2 7
3 8
4 9
5 10
11
12
13
14
15
16
17
18
19 (lowest)

Figure 10.2—Map of relative potential for primary habitation sites within the Santa Ynez Valley showing all unique conditions.
11.0—CONCLUSIONS

Based on the results of this research, a number of statements are supported regarding the criteria by which village sites were located in the Santa Ynez Valley. Locations with less than 15º slope, within 800 meters of perennial streams, within 200 meters of an ecotone, and in areas with a relatively high diversity of habitats were greatly preferred for settlement. It can also be said that proximity to perennial water had the greatest influence on village site location, with proximity to ecotones and low slope being the next criteria in level of importance. Relative environmental diversity, though still a significant influence, was less so than any of the other three.

It has also been shown that these criteria can be combined into a map that accurately shows which regions most effectively meet these criteria and therefore have the greatest potential to contain the remains of undiscovered village sites. The usefulness of such a map may seem intuitive but the question remains; “how is it to be used?”

One use of such a map is in support of archaeological research. The goal of archaeology is to reconstruct human behavior based on material remains. Village sites were locations of more intense habitation than other sites within the settlement systems of the inland Chumash. The greatest number and diversity of activities took place at village sites and they are valuable repositories of a great wealth of information about past lifeways. Sensitivity maps such as the one
produced here could be useful in the process of stratified sampling so as to assure that areas of greater and lesser probability are sampled equally. If a research design is focused only on discovery of sites, perhaps for the purposes of comparison between sites, then ground survey could be focused only on those areas where probabilities of site location are highest.

Perhaps the greatest potential utility of sensitivity maps is as a decision making tool for city and county planning departments. Planning departments are charged with the task of determining which specific requirements to impose on a development project in order to prevent destruction of undiscovered cultural resources. Village sites such as those modeled in this study were places central to the lives of the ancestors of living Native Americans. Village sites serve to anchor tribal identity to the landscape and are therefore extremely important to the continuance of cultural identity. In addition, village sites in the Santa Ynez Valley often have adjacent, virtually unmarked cemeteries that are held sacred by living Chumash people. Therefore, the areas marked most sensitive can be thought of as most likely to contain sites that arguably are the highest priority archaeological cultural resources and the measures required prior to ground disturbance should be appropriate and comprehensive.

11.1—Future Research Directions

The purpose of this study has been to model the distribution of a single class of archaeological sites and therefore can be seen as one part of a larger, more comprehensive study. It would be a valuable course of research to create
training point themes of sites of other functional types, such as short-term
habitation and resource extraction sites. Each functional site class could be
modeled separately and accurate models could be created for each. The resulting
sensitivity maps could be used as an even more detailed and comprehensive tool
for cultural resource managers.

There are multiple approaches whereby researchers might be able to work
towards the creation of site function classes. One would be to randomly select a
subset of known sites and perform testing at each, applying some criteria, such as
artifact richness, to determine site classes. Another approach might be to
construct multiple hypotheses regarding the criteria according to which expected
site types might be located and to create evidence layers to provide the data
needed to test those hypotheses. Based on the results of that testing, sites might
be tentatively separated into functional classes and those classes could be verified
by other methods such as limited subsurface testing of a subset of sites.

In addition to classifying sites within a model by function, it would also be
valuable to classify sites by antiquity. It is worthy to note that, while this study
included sites restricted to a specific, relatively modern time span, it is likely that
the sensitivity map also applies to older sites, as long as similar resource patterns
existed and similar adaptive strategies were used. However, throughout the past
millennia, environmental conditions within the study area have changed, and
Chumash social and economic systems changed as well. Therefore, not only did
the location and area of habitat regions likely change throughout time, but the
strategies employed to extract those resources and adapt to environmental conditions likely changed as well. It would be valuable to create response rasters that would be more accurate for sites of greater antiquity by creating maps of paleoenvironments, generating testable hypotheses concerning adaptive strategies that would have been employed under those conditions, and, guided by the test implications of those hypotheses, create response rasters and sensitivity maps that would more accurately predict sites of greater antiquity.

Another benefit of paleoenvironmental reconstruction is that it could allow the generation and testing of hypotheses concerning the influence of other important resources on settlement. For example, we know from ethnohistoric sources that acorns that were harvested from oak woodlands were important to the Chumash. It is likely that much of those areas was subsumed by agriculture as a result of settlement and are therefore underrepresented in modern datasets (John Johnson, personal communication). The reconstruction of the original extent of those habitats would allow the testing of hypotheses regarding the influence those habitats had on settlement.

Another way to create sensitivity maps of greater accuracy and precision might be to divide the study area into smaller geographic regions and model each separately. For example, the results of this study indicate at least some support for dividing the study area into eastern, central, and western subsections. It could be productive to create multi-class site distributions within each of these areas.
separately. The resulting sensitivity maps of each of those areas could be brought together into one tool for planners.

An alternative might be to overlay a grid onto the study area and model all sites in each grid square separately. Although there might be issues with abrupt changes in sensitivity zones where the squares meet, the overall accuracy might be improved in that models of smaller geographic areas are generally more accurate than larger ones.

A map such as the one generated by this study should not be seen as an endpoint, but a starting point. It is noteworthy that the model that is presented in this study predicts 80.9% of the village sites, the fact that it does not predict 100% can be taken to mean that there are other factors that influenced village site location. An advantage of predictive modeling is that additional layers can be added to this framework or changed as more powerful layers become available.

Closing Statements

It is my hope that this study has served multiple purposes. First, it can serve to test various hypotheses concerning environmental influences on inland Chumash settlement distribution. It has demonstrated that settlements were located so as to be close to perennial water, ecotones, regions of relatively high habitat diversity, and on locations with low ground slope. In addition, this study may be seen to support certain generalizations about the thought processes and values of the Chumash of the Santa Ynez Valley. First and foremost, this study has shown that relatively few environmental factors are required to create a model
that predicts a majority of the known ethnohistorically documented villages in the valley. From that, we can surmise that even though non-ecological factors such as political alliances and enmities were almost certainly important to the maintenance of those villages, the choice of location was clearly based on the natural environment.

The results of this study also may imply that, whereas hunter-gatherers in the Santa Ynez Valley possessed very detailed knowledge of their environment, their villages were located with regards to a much smaller subset of the specific resources—those most critical to the survival of a mostly sedentary population whose members included young, old, healthy, elderly, and infirm. In a geographical region that in different years might experience torrential rains or extreme drought, reduction of risk was critical. To that end, villages were located only in areas not with access to the best overall mix of a wide variety of non-essential resources, but with constant and reliable access to each one of a small set of critical resources.

It is also my hope that this research serves to demonstrate ways in which Geographic Information Systems can be used by archaeologists to test hypotheses. A tool, such as GIS, that can be used to analyze data both statistically and spatially is a natural fit to modern archaeological research. I hope that in some way this might serve to further the use of this powerful tool.
References Cited

Agterberg, F. P., G. F. Bonham-Carter, Q. Cheng and D. F. Wright
1993 Weights of Evidence Modeling and Weighted Logistic Regression for
Mineral Potential Mapping. In Computers in geology—25 years of
University Press, New York.

Anderson, M. K.
2005 Tending the Wild: Native American Knowledge and Management of
California’s Natural Resources. University of California Press, Berkeley,
California.

Archaeological Site Survey Record: CA-SBA-101
1987. On file at the Central Coast Information Center, University of
California Santa Barbara, Santa Barbara, CA.

Aschman, H.
1959 The Evolution of a Wild Landscape and its Persistence in Southern
California. Annals of the Association of American Geographers 48(3, part
2).

Benson, A. and S. Bowers
1997 The noontide sun : the field journals of the Reverend Stephen Bowers,

Berry, J.
2006a Justifiable Interpolation. GeoWorld 10(2).

Beyer, H. L.
2004 Hawth's Analysis Tools for ArcGIS.

Binford, L. R.
1980 Willow Smoke and Dog’s Tails: Hunter-Gatherer Settlement Systems
Biogeography Lab University of California at Santa Barbara

Bird, D. and J. O’Connell

Blackburn, T. C.

Bonham-Carter, G.

Bonham-Carter, G. F., F. P. Agterberg and D. F. Wright

Breschini, G. S.
1986 Preliminary Cultural Resources Reconnaissance of a Proposed Underground Transmission Line, Santa Ynez Peak, Santa Barbara County, California. Central Coastal Information Center, University of California at Santa Barbara CHRIS ID SBR00972.

Brewster, A., B. F. Byrd and S. N. Reddy

California Interagency Watershed Mapping Committee
2004 California Interagency Watershed Map of 1999 (CalWater 2.2.1). California Spatial Information Library (CaSIL).

California. Dept. of Forestry and Fire Protection., Pacific Southwest Forest and Range Experiment Station (Berkeley Calif.), W. F. Laudenslayer and K. E. Mayer
Carpenter, E. J.

Dalla Bona, L.


Duncan, R. B. and K. A. Beckman

Environmental Systems Research Institute, I.
2006a ArcGIS Projection Engine version 9.x Datum transformations available and geographic areas for which each transformation method should be applied.

Environmental Systems Research Institute Inc.


Fairchild Aerial Surveys
1938 Santa Barbara County Flight ID C-4950. Map and Imagery Laboratory, Santa Barbara, CA, Santa Barbara, CA.
Ford, A. and K. C. Clarke
2005 Modeling Settlement Patterns of the Late Classic Maya Civilization with Weights-of_Evidence and GIS. University of California at Santa Barbara, Santa Barbara, CA.

Gaffney, V. and Z. Stancic

Gamble, L.
1991 Organization of activities at the historic settlement of Helo': a Chumash political, economic, and religious center. In UMI Dissertation Services, Ann Arbor, MI. Obtained through HRAF, 2000, Ann Arbor, MI <http://ets.umdl.umich.edu/cgi/e/ehraf/ehraf-idx?q1=chumash%20village%20organization;rgn=paragraphs;owc=NS50;type=boolean;c=ehraf;a=view;start=1;size=25;subview=ocm;id=NS50-008;section=cite;sectbyte=20760013>.

Glassow, M. A.

2005 Personal Communication, edited by M. Neal, Santa Barbara, CA.

Gleick, P. H.

Grant, C.

Harrington, J. P., T. Hudson, J. Timbrook and M. Tempe
1978 Tomol : Chumash watercraft as described in the ethnographic notes of John P. Harrington. Ballena Press anthropological papers, no. 9. A Ballena Press/Santa Barbara Museum of Natural History cooperative publication, Socorro, N.M.
Harris, J. R., L. Wilkinson, K. Heather, S. Fumerton, M. A. Bernier, J. Ayer and R. Dahn

Hildebrandt, W. R.
2004 Xonxon'ata, in the tall oaks : archaeology and ethnohistory of a Chumash village in the Santa Ynez Valley. Contributions in anthropology (Santa Barbara, Calif.) ; no. 2. Santa Barbara Museum of Natural History, Santa Barbara, Calif.

Horne, S. P.

Jochim, M. A.


Johnson, J. R.


Kelly, R. L.

King, C.

Kohler, T. A.

Kohler, T. A. and S. C. Parker

Kroeber, A. L.
1962  Two papers on the aboriginal ethnography of California. Reports of the University of California Archaeological Survey, no.56. University of California Archaeological Survey Dept. of Anthropology, University of California, Berkeley.

Kvamme, K. L.


Landberg, L. C. W.
Longley, P.
2005 Geographical information systems and science. 2nd ed. Wiley, Chichester, West Sussex.

Macko, M. E.
1983 Beads, Bones, Baptisms, and Sweatlodges: Analysis of Collections from “Elijman” (CA-SBA-485), a Late Period Ynezeño Chumash Village in the Central Santa Ynez Valley, California., University of California Santa Barbara.

Madry, S. L. H.

Masetti, M., S. Poli and S. Sterlacchini

McRae, K. S.
1999 Soxtonokmu' (CA-SBa-167): An analysis of artifacts and economic patterns from a late period Chumash village in the Santa Ynez Valley, University of Texas at San Antonio.

Mink, P. B., B. J. Stokes and D. Pollack

Montecito Water District
1998 Sources of Water

Norris, R. M.
2003 The geology and landscape of Santa Barbara County, California, and its offshore islands. Santa Barbara Museum of Natural History monographs; no. 3. Santa Barbara Museum of Natural History, Santa Barbara, Calif.
O'Brien, R. A.  

Pilgram, T.  

Raines, G. L.  


Rogers, D. B.  
1929  *Prehistoric Man of the Santa Barbara coast.* Santa Barbara Museum of Natural History, Santa Barbara, Calif.

Roper, D. C.  

Rosenthal, J., J. Meyer, W. Hildebrandt and J. King  

Ruth, C.  
1950  Archaeological Site Survey Record, CA-SBA-219. Unpublished - Available at the Central Coastal Information Center, University of California at Santa Barbara.

Schermer, S. J. and J. A. Tiffany
1985 Environmental variables as Factors in Site Location: An Example from the Upper Midwest. *Midcontinental Journal of Archaeology* 10(2).

Smith, C. F.
1998 *A flora of the Santa Barbara region, California: an annotated catalog of the native, naturalized, and adventive vascular plants of mainland Santa Barbara County, adjacent related areas, and four nearby Channel Islands*. 2nd ed. Santa Barbara Botanic Gardens & Capra Press, Santa Barbara, Calif.

Smith, E.

Soils Survey Staff

Spanne, L. W.

SPSS Inc.
2006 SPSS 15.0 for Windows. SPSS, Inc., Chicago, IL.

Stafford, C. R. and E. R. Hajic

Tainter, J. A.

Tartaglia, L. J.

Tchoukanski, I.

Timbrook, J.

Trigger, B. G.

U.S. Geological Survey


U.S. Geological Survey (USGS) EROS Data Center


U.S. Geological Survey EROS Data Center

University of California at Berkeley and University of California at Davis
Warren, R. E. and D. L. Asch
2000 A Predictive Model of Archaeological Site Location in the Easter
Prairie Peninsula. In Practical Applications of GIS for Archaeologists: A
Taylor and Francis, Philadelphia, PA.

Wescott, K. L. and J. A. Kuiper
2000 Using a GIS to Model Prehistoric Site Distributions in the Upper
Chesapeake Bay. In Practical Applications of GIS for Archaeologists: A
Taylor and Francis, Philadelphia, PA.

West, G. J.
1987 Bureau of Reclamation Mid-Pacific Regional Office Archaeological
Site Record CA-SBA-477, edited by B. o. Reclamation. Unpublished -
Available at the Central Coastal Information Center, University of
California at Santa Barbara.

Wheatley, D. and M. Gillings
2002 Spatial technology and archaeology : the archaeological applications
of GIS. Taylor & Francis, London ; New York.

When Thirsty – Drink!

Willey, G. R.
Print. Off., Washington,

Woodman, C. F., J. L. Rudolph and T. P. Rudolph (editors)
1991 Western Chumash Prehistory: Resource Use and Settlement in the
Santa Ynez River Valley, sponsored by the Unocal Corporation. . Science
Applications International Corporation, Santa Barbara, CA.

Zeanah, D. W.
2004 Sexual division of labor and central place foraging: a model for the
Carson Desert of western Nevada. Journal of Anthropological